Properties of induced subgraphs of random graphs with given degree sequences

Angus Southwell

Bachelor of Science (Hons), Monash University, 2016

A thesis submitted for the degree of

Doctor of Philosophy

at Monash University in 2022

School of Mathematics
Monash University
Melbourne, Australia
2022

Copyright notice

Abstract

In this thesis we study induced subgraphs of uniformly random graphs with given degree sequences. Specifically, if $G \sim \mathcal{G}(\boldsymbol{d})$ is a uniformly random graph with degree sequence $\boldsymbol{d}$, and $S \subset [n]$ is some subset of vertices (which can be fixed or random), we study the properties of the random graph $G[S]$. First we study the distribution of the degree sequence of $G[S]$, which we call the *induced degree sequence*. Under certain restrictions on both the maximum degree of $\boldsymbol{d}$ and the size of the set S, we show that the induced degree sequence is concentrated around some sort of average sequence defined deterministically in terms of $\boldsymbol{d}$ and S. Our main tool that we use to show this is the switching method. We then use this knowledge of the distribution of the induced degree sequence to determine properties of $G[S]$ using known results about random graphs with given degree sequences. Under the restrictions we impose on $\boldsymbol{d}$ and S, we determine thresholds for when $G[S]$ is connected using results of Gao and Ohapkin [56], as well the threshold for the existence of a giant component in $G[S]$ using the results of Joos et al. [83]. We also give results on the chromatic number using results of Gao and Ohapkin [56], and conditions for the (asymptotically almost sure) existence and non-existence of non-trivial automorphisms using results of Brick et al. [23]. This method of studying $G[S]$ is similar to the method of Fountoulakis [46] for studying site and bond percolation on sparse random graphs with given degree sequences. We then adapt our method to give thresholds for connectivity and existence of a giant component in $\mathcal{G}(\boldsymbol{d})$ after site percolation with survival probability p for constant p, which also applies for more dense graphs. We then focus on studying the threshold for the existence of giant components in $G[S]$ using a different, more direct approach. We modify the random graph process used by Joos et al. [83] in their proof of the threshold for the existence of giant components in $\mathcal{G}(\boldsymbol{d})$. We use this to determine the threshold for the existence of giant components in induced sub-pairings in the configuration model. We then apply the switching method to adapt this result to the random graph model for more restricted classes of degree sequences and vertex subsets. This result gives a more natural expression for the threshold, as well as being applicable for a wider range of degree sequences $\boldsymbol{d}$ and subsets S. We also show that the threshold in the configuration model does not translate to the random graph model for all possible degree sequences $\boldsymbol{d}$ and subsets S, and we give some examples where the configuration model gives the wrong intuition for the existence of giant components in induced subgraphs in random graph model.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Angus Southwell
1 July 2022

Publications during enrolment

- Counting triangles in power-law uniform random graphs, with P. Gao, R. van der Hofstad, and C. Stegehuis; The Electronic Journal of Combinatorics, vol. **27**, issue **3** (2020).

  This paper studies the local clustering coefficient (the probability that two random neighbours of a vertex are adjacent) and the number of triangles in two random graph models. The models studied are the erased configuration model and the uniform random graph model where the degree distribution follows a power-law exponent.

- The threshold of symmetry in random graphs with specified degree sequences, with L. Brick and P. Gao; preprint (2020).

  This paper studies the threshold for a uniform random graph with a given degree sequence to asymptotically almost surely have a non-trivial automorphism. In the case where the degree sequence is bounded, we precisely determine the threshold for the existence of non-trivial automorphisms. In the unbounded case, we do not completely characterise the threshold. However, we give necessary conditions under which a uniformly random graph asymptotically almost surely does not have a non-trivial automorphism, as well as conditions for which it asymptotically almost surely does have at least one non-trivial automorphism.

- Distribution of tree parameters by martingale approach, with M. Isaev and M. Zhukovskii; Combinatorics, Probability and Computing, pg. 1–28 (2022).

  This paper presents a general theorem for analysing the distribution of tree parameters on uniformly random trees. In particular, this result differs from other results about the limiting distribution of tree parameters in the conditions that we impose on the parameter. We then apply this theorem to prove the asymptotic normality of the number of occurrences of a small pattern, as well as the asymptotic log-normality of the number of automorphisms of a uniformly random tree.

This was fun.

Acknowledgements

First and foremost I would like to thank my supervisors, Nick Wormald, Jane Gao, and Mikhail Isaev. Thank you for all the support and guidance you've offered, and for all the fun and interesting problems you've brought to my attention. Thank you to Jane for taking me on initially and allowing me to visit the University of Waterloo, and thank you to Nick and Misha for providing support and supervision at Monash. I couldn't have gotten through this without all of your help. I would also like to thank Bruce Reed for a helpful discussion at an early stage of the research.

Secondly, I'd like to thank my family, who have loved and supported me through these hectic and uncertain times. You've helped and encouraged me through the difficult patches, and rejoiced in the successes with me. It's meant so much to me, and it's made so much of a difference to know that you're always there for me.

Thirdly, I'd like to thank my friends, who have made my candidature immeasurably better by their presence and their company. Thank you to Adam, Afsane, Billy, Dan, Kaustav, Michael, Robert, Rui, and Sophie, who made the office such an enjoyable place and have been a constant source of friendship and support. Thank you to Luci, Rachael, and Riley, who have always been around to catch up for lunch or coffee and to sympathise with the stresses of the PhD program. Thank you to all my friends outside of university, in particular to Wren and Jess, for your support and friendship, and for reminding me that there is a world outside of academia.

Finally, I would like to thank Norm Do, who encouraged me to pursue mathematics research and with whom I had my first foray into real mathematics all the way back in 2014. Thank you for all your support, and for giving me such a welcoming introduction to research.

Contents

Chapter 1

Introduction

Graphs and networks have become ubiquitous tools in all fields of modern research. As the world becomes more digital and interconnected, computational methods to analyse large and complex processes are rapidly becoming integral to all aspects of our lives. As such, there is a growing need to describe objects and processes as discrete systems, and graphs have proven to be a valuable tool for doing this. As a result, graphs and other network-like objects permeate all fields of study: computer science, neuroscience, biochemistry, linguistics, and economics to name just a few [37, 129, 87, 30, 8]. The different structures of graphs are as varied as their applications, and as the graphs grow in size their complexity explodes beyond the capabilities of naive enumeration and analysis. Brute force methods can give insights into basic properties of small graphs, but the networks that appear in many research or commercial applications can have millions and billions of nodes. Compounding this issue is the fact that the number of graphs grows staggeringly quickly. There are $2^{\binom{n}{2}}$ labelled graphs on n vertices, and the number unlabelled graphs (graphs up to isomorphism) on n vertices also grows incredibly quickly - for example, there are 64,001,015,704,527,557,894,928 unique unlabelled graphs on 16 vertices [121]. As a result, general methods and tools to study large graphs are invaluable for understanding these large networks and informing our applications of them.

One avenue for studying the properties of large graphs is to generate graphs via some random process, called a *random graph model*. By studying the distribution of a graph property in some random graph model, we can determine the typical behaviour of that class of graphs without explicitly generating and analysing each graph. Classical random graph models, such as the binomial random graph model $\mathcal{G}(n, p)$, have been studied extensively and given many insights into the properties of large graphs. However, many of these random graph models often assume that the vertices are homogeneous, in the sense that each vertex in the graph is more or less equivalent to every other vertex. For many networks that arise in real-world applications, this is not the case: it is common instead to see networks where some nodes are linked to only a few other nodes, while some nodes act as "hubs" and have many more links than other nodes [7, 115, 33]. These networks cannot be easily analysed by studying these classical random graph models, as many of their properties are a consequence of their inhomogeneity.

One tool for studying such networks is what is called a null model (see [90]). A null model is a random model (a random graph model, for our purposes) which satisfies a particular set of constraints but is otherwise sampled uniformly at random from all objects which satisfy the constraints. A particular instance of a graph, such as one obtained through empirical study of

some process, can then be compared with the null model to determine which of its properties are expected based on the observed constraints, and which properties are unexpected or non-trivial for such a network. If the null model for a set of constraints is likely to have a given property, then it is less interesting if the network itself has that property; however, if the network has some feature that the null model has with very low probability, then this feature is noteworthy, and likely a product of the underlying process being studied.

A natural constraint to construct a null model for a network is its *degree sequence*, which is a sequence $\boldsymbol{d} = (d(1), \ldots, d(n))$ where the elements are the number of links incident to each node in the network, or the number of edges incident to each vertex in the graph. For a given network, the degree sequence is relatively simple to determine, as it is a local property of the network. This gives rise to a family of random graph models called *uniformly random graphs with a given degree sequence*, where a graph is sampled uniformly at random from all graphs where each vertex has the required degree. Much less is known about this model (or family of models) due to the extra complexity that comes with fixing a particular degree sequence. Nevertheless, as we discuss in the following section, these models are still actively studied for a wide range of degree sequences because they give extremely helpful insights into these networks with inhomogeneous degree sequences. Due to the difficulties of working directly with this model, other related but simpler models have been studied to help give insights where the theoretical results about random graphs with given degree sequences fall short. One such model is the *configuration (or pairing) model*, which constructs a random (pseudo)graph from a random pairing of sets of points. This model was originally proposed to study uniformly random graphs with given degree sequences [10], but has since received a great deal of attention in its own right, particularly as a tool for studying networks [71].

A common question to ask about a graph is what subgraphs it contains. Many graph properties can be phrased in terms of subgraphs - for example, a graph is connected if and only if it contains a spanning tree, and a graph is bipartite if and only if it contains no odd cycle as a subgraph. This study of subgraphs has naturally followed in the field of random graphs, and the subgraphs of various models of random graphs have been extensively studied. This has applications for studying the properties of networks in the event that certain nodes or links are removed from the network. For example, Britton et al. [25] studied the spread of disease by modelling a social interaction network as a uniformly random graph with a given degree sequence. This graph model is the null model for the social interaction network under the constraint that each person has a given number of friends. They then studied vaccination strategies to minimise the number of infections: in this case, vaccinating a person corresponds to deleting that node from the network. Thus, studying the spread of the disease through the unvaccinated people in the network becomes a problem of studying the subgraph induced by the corresponding nodes. In their paper they also look at a random vaccination strategy: that is, studying the spread of the disease in a subgraph made by deleting vertices randomly according to some probability. This fits into a broader study of what is known as *percolation*, or more specifically site percolation. The study of percolation on large (sometimes infinite) fixed graphs, as well as on random graphs of various kinds, has been an active field of research since its inception, both from a purely mathematical perspective [26, 22, 107, 51, 82] and also for its applications to modelling problems [96, 27, 81].

Thesis aim. Study the distribution and properties of subgraphs of uniformly random graphs

with given degree sequences induced on a subset of the vertex set. In particular, we aim to study when this subgraph asymptotically almost surely is connected or contains a giant component.

In this thesis we study subgraphs of uniformly random graphs with given degree sequences induced on a fixed subset of the vertices. This is related to but notably distinct from the study of percolation, where the subset is random. We give particular focus to the connectivity structure of the induced subgraph, and under what conditions we can determine a threshold for connectivity or the existence of a *giant component*, a component with a linear number of edges and/or vertices, based on the degree sequence and the subset of the vertices in question. Connectivity is a fundamental graph property which has been studied in many random graph models [58, 42, 130, 45] and has been actively researched since the inception of the field. The study of giant components in random graphs has a similarly rich history [42, 15, 113], and recently Joos et al. [83] determined the threshold for the existence of giant components in random graphs with given degree sequences, for all possible sequences. Furthermore, it is a natural question when studying subgraphs of random graphs to ask about the connectivity structure of the remaining subgraph. We give two different approaches, one more general approach for studying a wide range of properties and one specifically for studying the existence of giant components. The approaches are:

(1) The *reduction method*, where we study the distribution of the degree sequence of the induced subgraph. We then combine this information with known results about random graphs with given degree sequences to determine various properties of the induced subgraph.

(2) The *exploration method*, where we define a modified breadth-first search process and study its evolution on a random graph to determine the threshold for the existence of giant components in the induced subgraph.

In Chapter 3, we analyse the distribution of the degree sequence of the induced subgraph, which we call the *induced degree sequence*. We use ideas similar to those used by Fountoulakis [46] in his study of giant components in percolated sparse random graphs with given degree sequences. We extend this method to study the case where the subset is fixed, as well as considering a much wider range of degree sequences: in particular, the results obtained in this thesis work for some degree sequences with much higher maximum degree and average degree. For sufficiently nice choices of degree sequence $\boldsymbol{d}$ and vertex subset S, we show that the degree sequence of the induced subgraph is "concentrated", in the sense that we can define a particular degree sequence and show that the induced degree sequence is asymptotically almost surely close to this. In Chapter 4, we use these results about the distribution of the induced degree sequence to study various graph properties of the induced subgraph. This method is general in that it works for a variety of different graph properties. By combining the knowledge of the distribution of the induced degree sequence with known results about thresholds for graph properties in random graphs with given degree sequences, we determine thresholds (in terms of $\boldsymbol{d}$ and S) for when the induced subgraph is connected or has a giant component, as well as results about the chromatic number of this graph or when it has non-trivial automorphisms. We call this method the reduction method as we are reducing the problem of studying the induced subgraph to studying pre-existing thresholds for properties of random graphs with given degree sequences. In Chapter 5, we extend the study of the distribution of the degree sequence of the induced graph to study site percolation, which is the case where S is constructed by choosing each vertex independently with probability p. For a

similarly nice choice of degree sequence, we give thresholds for when the site-percolated random graph with a given degree sequence is a.a.s. connected or contains a giant component in the case where p is a constant.

In Chapters 6 and 7, we study the problem of giant components in induced subgraphs using the exploration method. We adapt the proof of a recent result given by Joos et al. [83] describing the threshold for the existence of giant components in uniform random graphs with given degree sequences. Their method studies a breadth-first search exploration process which generates a random graph via the method of deferred decisions. Similar ideas have existed in the literature for some time (for example, [113]), but their method includes extra initial steps to handle the presence of high-degree vertices. We adapt this framework to study the existence of giant components in induced subgraphs. In Chapter 6, we apply this method in the configuration model. In this case, we determine the exact threshold for the existence of a giant component in the induced sub-pairing of a uniformly random pairing P with degree sequence $\boldsymbol{d}$, for all sequences $\boldsymbol{d}$, under the condition that both the set S and its complement are large with respect to the whole pairing. In Chapter 7, we adapt this result to the random graph model under tighter restrictions on the pair $(\boldsymbol{d}, S)$. Specifically, we introduce two restrictions on the contribution of high-degree vertices to S and its complement (the set of deleted vertices). The restriction on the contribution to the complement is significantly more strict. Under these tighter restrictions, we show that an analogous proof to the configuration model result carries over, and the same threshold as in the configuration model applies. The range of degree sequences for which this result applies is larger than the range for which the results given in Chapter 4 apply, but the method is much more specific to the study of giant components. We also showcase why the conditions we impose on $\boldsymbol{d}$ and S are meaningful: we give some examples of pairs $(\boldsymbol{d}, S)$ which show that the configuration model result cannot carry over in full generality.

In Chapter 8, we discuss possible directions for future research using either method. We describe various avenues to relax the conditions that we impose when studying $G[S]$ via the reduction method, as well as other results that could be proved using the current method and its generalisations. For the exploration method, we discuss some conjectures that slightly relax the conditions we impose in our results. We also give other possible avenues to explore. Many of these extensions centre around studying the neighbourhoods of vertices of high degree in $\mathcal{G}(\boldsymbol{d})$. This is a difficult problem, and progress in this direction would likely require novel methods that could be used to improve many current results about random graphs with given degree sequences. In the Appendix, we give a glossary of the notation that we use in the various sections, as well as some well known probability tools and bounds that we use throughout the thesis.

Chapter 2

Background

2.1 Basic definitions

2.1.1 What is a graph?

Here we give the most basic definition that sets up all the work in this thesis.

Definition 2.1.1. A (simple) graph is a pair $G = (V, E)$, where V is an arbitrary set and E is a set consisting of 2-element subsets of V.

This definition is correct and good, but it is not how we often imagine graphs. If we consider the set V as a set of nodes, or vertices, and the set E as connections between these nodes, then we can represent graphs pictorially. For example, the graph where $V = \{1, 2, 3, 4, 5, 6\}$ and $E = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 4\}, \{3, 4\}, \{4, 5\}, \{4, 6\}\}$ can be represented by the image in Figure 2.1.

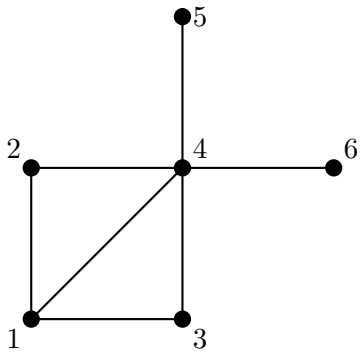

Figure 2.1: A visual representation of a graph.

This is much easier to conceptualise, but also hints at the motivation for studying these objects: a graph is like an abstracted form of a network. As such, the mathematical study of graphs can thus give insight in all sorts of surprising places in many different fields. On the other hand, from a purely mathematical perspective, graphs are incredibly interesting objects to tinker with, and over the last century or two graph theory has become an incredibly popular topic in mathematics and theoretical computer science.

Here we give some basic graph theory terminology for those less familiar with the field. For a more comprehensive introduction to graph theory, including some terminology that we use but

do not explicitly define, we recommend the classic textbook Graph Theory by Diestel [39]. We say that two vertices $u, v \in V(G)$ are *adjacent* or *neighbours* in a graph G if $\{u, v\} \in E(G)$. For brevity, we generally denote an edge $\{u, v\}$ simply by uv (or equivalently vu, since edges are not directed). A commonly used notation for adjacency is $u \sim v$, however we avoid this here to minimise notational clashes with asymptotic notation used later. The set of neighbours of a vertex $v \in V(G)$ is denoted $N_G(v)$. The number of vertices adjacent to v is the *degree* of v, and the maximum degree of a vertex in G is denoted $\Delta(G)$. A sequence of vertices $v_1 v_2 \dots v_k$ such that $v_i \neq v_j$ for all $i \neq j$ and $v_i v_{i+1} \in E(G)$ for all $i \in \{1, \dots, k-1\}$ is called a *path*. Two vertices are connected if there exists a path in G between u and v. A set of vertices which are all connected to each other in a graph is called a *component*. The *order* of a component is the number of vertices contained in the component, and the *size* is the number of edges between vertices in the component. A graph $H = (V', E')$ is called a *subgraph* of $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$. For some $V' \subset V$, we define the *induced subgraph on* V' to be the subgraph of G with vertex set V' and edge set E' consisting of all edges $uv \in E$ where $u, v \in V'$. Let $G[V']$ denote the induced subgraph on V'.

2.2 Random graphs and graph properties

Here we lay out the basic definitions of a general random graph model, as well as specifically the model of random graphs with a given degree sequence, which is the focus of this thesis. We also describe the asymptotic notation that we use to discuss properties of random graph models as the number of vertices grows arbitrarily large. We also give a history and literature review of the properties of random graphs that we study.

2.2.1 The basic concept

Definition 2.2.1. A random graph model is a probability space $(\Omega, \Sigma, \mathbb{P})$ where Ω is a set of graphs and $\Sigma = 2^{\Omega}$, the set of all subsets of Ω.

This definition is intentionally broad, as there are many, many different types of random graph models. One of the most classical and well-studied random graph models is the *binomial random graph*, also denoted by $\mathcal{G}(n, p)$. In this model, Ω is the set of all labelled, n-vertex graphs, and $\mathbb{P}(G) = p^{|E(G)|}(1-p)^{\binom{n}{2}-|E(G)|}$ for all $G \in \Omega$. Equivalently, and perhaps more intuitively, one can generate a random graph according to this distribution by taking n isolated vertices and adding each possible edge with probability p, independent of all other edges. This random graph model has been studied in formidable detail since its invention (see any or all of [16, 80, 50, 39] to merely scratch the surface), and despite its simple definition has yielded much insight into the behaviour of large graphs. In particular, when $p = \frac{1}{2}$, this model corresponds to a uniformly random labelled graph on n vertices. Thus, studying $\mathcal{G}\left(n, \frac{1}{2}\right)$ allows us to determine the unbiased distribution of graph properties across all n-vertex graphs. Another common random graph model is $\mathcal{G}(n, M)$, where Ω is the set of all graphs with exactly n vertices and M edges, and $\mathbb{P}$ is the uniform distribution. These are sometimes called Erdős-Rényi random graphs (although sometimes this is also ascribed to $\mathcal{G}(n, p)$) due to their early work studying this model. This model turns out to be extremely similar to $\mathcal{G}(n, p)$, for appropriate choices of p and M (for a more detailed discussion

on this see Section 1.1 of the book Random Graphs by Frieze and Karoński [50], in particular Theorem 1.4, originally given by Łuczak [99]).

Typically when studying random graphs, people are interested in the asymptotic behaviour of some graph property. That is, the focus is less on small graphs, and instead people are interested in what happens when the graphs are allowed to grow arbitrarily large. Because of this, many results in the field are expressed (sometimes implicitly) in terms of sequences of random graph models, indexed by the number of vertices in each graph (or some similar quantity). Suppose that $\{\mathcal{G}_n\}_{n\geq 1}$ is a sequence of random graph models. Let $\{E_n\}_{n\geq 1}$ be a corresponding sequence of events. These could be anything - for example, the event that the vertex labelled 1 is isolated (that is, has no neighbours), or the event that the graph sampled from $\mathcal{G}_n$ is connected, or countless other things. We say that an event occurs *asymptotically almost surely* (which we commonly abbreviate to a.a.s.) if the probability that the event occurs converges to 1 as n goes to ∞. That is, if $\mathbb{P}_n$ is the probability measure on $\mathcal{G}_n$, then $\{E_n\}_{n\geq 1}$ occurs a.a.s. if and only if $\mathbb{P}_n(E_n) \to 1$. If $\{\mathcal{G}_n\}_{n\geq 1}$ is a sequence of uniform probability spaces, this means that the fraction of graphs that do not have the property in question is vanishingly small. We discuss asymptotic concepts and notation in more detail in Section 2.2.4.

2.2.2 Random graphs with a given degree sequence

One of the most fundamental properties of a graph is its *degree sequence*, which is the sequence detailing the degree of each vertex (the number of neighbours of each vertex).

Definition 2.2.2. The degree sequence of a (labelled) graph G is the sequence $\boldsymbol{d} = (d(1), \dots, d(n))$ such that vertex i has degree $d(i)$. We call a sequence $\boldsymbol{d}$ *graphical* (or sometimes just a *degree sequence*) if there exists some (simple) graph with degree sequence $\boldsymbol{d}$.

Here we give some notation for several properties of a sequence $\boldsymbol{d}$. The value of the largest element in the sequence is its *maximum degree*, and is denoted by $\Delta := \Delta(\boldsymbol{d})$. The length of a sequence is denoted $n(\boldsymbol{d})$, to correspond with the notation for the number of vertices in a graph with that degree sequence. For some set $A \subset [n(\boldsymbol{d})]$, define the *total degree* of A, denoted by $d(A)$, as

$$d(A) = \sum_{i\in A} d(i).$$

For the case where $A = [n(\boldsymbol{d})]$, define

$$M(\boldsymbol{d}) := d([n(\boldsymbol{d})]) = \sum_{i\in[n(\boldsymbol{d})]} d(i).$$

We call $M(\boldsymbol{d})$ the *total degree* of $\boldsymbol{d}$. Equivalently, for a graph with degree sequence $\boldsymbol{d}$, we call $M(\boldsymbol{d})$ the total degree of G, and this is equal to $2|E(G)|$. When it is unambiguous, we simply refer to this as M. The *average degree* of a sequence is then $d(\boldsymbol{d}) := M(\boldsymbol{d})/n(\boldsymbol{d})$.

One major caveat to the $\mathcal{G}(n,p)$ and $\mathcal{G}(n,M)$ random graph models is that the distribution of many graph properties is far removed from that of many types of real-world networks [115]. For example, the degree of a vertex in $\mathcal{G}(n,p)$ is a binomially distributed random variable with $n-1$ trials (one for each other vertex in the graph) and probability of success p (the probability of each

edge being present). On the other hand, the degree sequences of many real-world networks have been conjectured to follow a power-law distribution, where the probability that a vertex has degree k is proportional to $k^{-\gamma}$, for some constant γ (typically $\gamma \in (2,3)$) [33]. As a result, Erdős–Rényi random graphs have entirely different structure to many families of graphs found in real-world applications. In order to study these types of graphs, other random graph models are needed. As mentioned in Chapter 1, one method for studying these real-world networks is to consider a null model with constraints derived from the properties of a particular network [90]. One natural constraint is the degree sequence, as it is a very "local" property of the network and thus easy to determine. In this case, the null model samples uniformly at random from every graph with the same degree sequence as the original network. This sets the stage for the model which we analyse for most of the thesis. We call this the *uniformly random graph with degree sequence* $\boldsymbol{d}$.

Definition 2.2.3. Let $\mathcal{G}(\boldsymbol{d})$ denote the set of labelled graphs with degree sequence $\boldsymbol{d}$ equipped with the uniform probability measure.

As with other random graph models, we also refer to "the graph $\mathcal{G}(\boldsymbol{d})$", meaning a graph sampled uniformly at random from the set of all graphs with degree sequence $\boldsymbol{d}$. We also sometimes abuse notation and refer to the set of all graphs with degree sequence $\boldsymbol{d}$ by $\mathcal{G}(\boldsymbol{d})$.

Uniform random graphs with given degree sequences are a broad class of commonly studied random graph models. Unlike binomial random graphs or Erdős-Rényi random graphs, these graph models can be used to study graphs and networks with extremely inhomogeneous degree sequences. For example, the probability that $\mathcal{G}(n,p)$ contains both a degree 1 and a degree $n-1$ vertex is immeasurably low, but in $\mathcal{G}(\boldsymbol{d})$ we can guarantee it by choosing such a sequence $\boldsymbol{d}$. This immediately fixes a notable limitation of $\mathcal{G}(n,p)$, as the degree of each vertex is no longer concentrated around np. However, the extra nuance also brings with it much more challenging analysis: events that were straightforward or even trivial to study in $\mathcal{G}(n,p)$ can be challenging in $\mathcal{G}(\boldsymbol{d})$. For example, the probability that two vertices are adjacent in $\mathcal{G}(n,p)$ is immediate from the definition, and is completely independent of all other edges in the graph. However, the probability of this event occurring in $\mathcal{G}(\boldsymbol{d})$ does not have an easy answer in general, and depends strongly on both the degree sequence $\boldsymbol{d}$ and the specific two vertices that are being considered. Furthermore, due to the fixed degrees of each vertex, the probability that some edge exists is dependent on every other edge in the graph. Thus, when trying to look for more complicated structures in these random graphs that use multiple edges, it is not enough to know information about the likelihood of each edge in isolation.

In the context of $\mathcal{G}(n,p)$, the notion of a sequence of probability spaces is fairly straightforward: each set Ω_n just contains all labelled graphs on n vertices. However, for $\mathcal{G}(\boldsymbol{d})$, the degree sequence $\boldsymbol{d}$ fixes the number of vertices in the graph, so it is not immediately clear what a "sequence" of probability spaces should look like. Various ideas exist in the literature to deal with this. Some authors define the notion of an *asymptotic degree sequence*, which is a sequence of sequences $(\boldsymbol{d}_n)_{n\geq 1}$ where $\boldsymbol{d}_n$ is an n-element, graphical sequence for each n satisfying some nice limiting properties. Molloy and Reed [113], Fountoulakis [46] and Janson [74] all assume that the asymptotic degree sequences are what they call smooth, where $(\boldsymbol{d}_n)_{n\geq 1}$ is smooth if

$$\frac{n_j(\boldsymbol{d}_k)}{k} \to \lambda_j$$

for some constant $\lambda_j \in [0,1]$ for each $j \geq 0$. They also assume that the asymptotic degree sequence is sparse, which means that there exists some $\lambda \in (0,\infty)$ such that

$$\frac{1}{k}\sum_{i\geq 1} in_i(\boldsymbol{d}_k) \to \lambda.$$

Paraphrasing, this condition says that the sum of every element in the degree sequence $\boldsymbol{d}_k$ (the total degree of $\boldsymbol{d}_k$) is asymptotically proportional to k, the number of vertices in a graph with degree sequence $\boldsymbol{d}_k$. Molloy and Reed [113] and Fountoulakis [46] also assume that

$$\frac{1}{k}\sum_{i\geq 1} i(i-2)n_i(\boldsymbol{d}_k) \to \sum_{i\geq 1} i(i-2)\lambda_i < \infty.$$

This last condition is assumed by Fountoulakis [46] because their result is proved by applying the result of Molloy and Reed, which itself assumes this condition. Other approaches, such as that of Janson [74], do not require this second assumption.

One thing to note is that many results about $\mathcal{G}(\boldsymbol{d})$ (and also many results about the configuration model $\mathcal{C}(\boldsymbol{d})$, defined in the next subsection) only consider sequences with minimum degree at least 1. If $\boldsymbol{d}$ has k terms equal to 0, then a uniformly random graph with degree sequence $\boldsymbol{d}$ has an identical distribution to a uniformly random graph with degree sequence $\boldsymbol{d}^*$ (where $\boldsymbol{d}^*$ is the sequence $\boldsymbol{d}$ with all zeroes removed) and k isolated vertices. Thus, allowing isolated vertices is not particularly interesting, and so a minimum degree of at least 1 is commonly assumed. With this in mind, in this thesis we also usually assume that $\boldsymbol{d}$ has a minimum degree of at least 1. In particular, we assume that $\boldsymbol{d}$ has a minimum degree of at least 1 when studying induced subgraphs of $\mathcal{G}(\boldsymbol{d})$ in Chapters 3 to 7. Note that since we study degree sequences of induced subgraphs, it is possible that the degree sequences of these induced subgraphs may contain zeroes. We discuss this in more detail in the relevant later chapters.

2.2.3 The configuration model

Due to the inherent difficulty of working with $\mathcal{G}(\boldsymbol{d})$, many other random graph models have been created to study graphs with inhomogeneous degree sequences. The *configuration (or pairing) model*, denoted $\mathcal{C}(\boldsymbol{d})$, was introduced by Bollobás [10] as a way of approximating and modelling random graphs with a given degree sequence. In this model, each vertex i is represented as a bin with label i which contains $d(i)$ labelled points. These points are called *half edges*. A uniformly random element of $\mathcal{C}(\boldsymbol{d})$ corresponds to a uniformly random perfect matching of the half edges. Such a matching is called a *configuration*, or a *pairing*. Two matched half edges $\{p_i, p_j\}$ are a *pair*, and we call p_j the *mate* of p_i. We enumerate the set of all half edges in a pairing by the set $[M] := \{1,\dots,M\}$, recalling that $M = M(\boldsymbol{d})$ is the total degree of a graph or (degree) sequence. Let $B(i) \subset [M]$ be the set of half edges in the bin corresponding to vertex i.

Pairings have a natural correspondence to what are called *multigraphs* with degree sequence $\boldsymbol{d}$: simply consider each pair as an edge in the multigraph and ignore the labels of the half edges. The result is an object similar to a graph that potentially has multiple edges between certain pairs of vertices, as well as edges which start and end at the same vertex (loops). In some works, these are also referred to as pseudographs, though some authors distinguish multigraphs as pseudographs with no loops. With this correspondence between pairings and multigraphs, many graph-theoretic

concepts, such as adjacency, connectivity, and subgraphs, extend to the configuration model. A *vertex* in a pairing is a bin of points, and the *degree* of the vertex is simply the number of points in the bin. An *edge* then corresponds to two half edges that have been paired, and we call these two half edges the *endpoints* of that edge. We differentiate between the set of half edges (all points in all bins) and the set of edges (the set of pairs of half edges in a given pairing). Two vertices u and v have an edge between them (are adjacent) if there exists a half edge u_i in $B(u)$, the bin corresponding to u, and a half edge v_j in $B(v)$ such that $\{u_i, v_j\}$ is a pair. The vertices u and v are *parent vertices* to the points u_i and v_j.

Similarly, concepts of connectedness carry over naturally following the definition of adjacency. The *order* and *size* of a component of a pairing can be respectively defined as the number of vertices and edges in the component. We can also define concepts analogous to subgraphs: from a pairing P we can define a partial pairing $Q \subset P$ where some of the vertices and half edges of P are removed. Similarly to induced subgraphs, define $P[S]$ to be the sub-pairing of P only containing half edges whose parent vertices are in $S \subset [n]$. Depending on the application, it can be beneficial to delete or to keep the unpaired half edges, that is, the half edges whose mate has been deleted.

In this thesis, a (uniformly) random configuration or random pairing refers to a uniformly random pairing chosen from $\mathcal{C}(\boldsymbol{d})$. The relevance of the configuration model to random graphs stems from the following well-known result. Suppose $G(P)$ is the (multi)graph corresponding to a given pairing P, and call a pairing *simple* if $G(P)$ is a simple graph (no multiple edges or loops).

Proposition 2.2.4. Let $\boldsymbol{d}$ be a graphical sequence. Let $G \in \mathcal{G}(\boldsymbol{d})$ be a simple graph with degree sequence $\boldsymbol{d}$. Then

$$\mathbb{P}\left(G(\mathcal{C}(\boldsymbol{d})) = G \middle| \, \mathcal{C}(\boldsymbol{d}) \text{ is simple}\right) = \mathbb{P}\left(\mathcal{G}(\boldsymbol{d}) = G\right).$$

Proof. Let P be a pairing that corresponds to the simple graph G, that is, $G(P) = G$. Each permutation σ which permutes the half edges within each bin of P creates a new pairing $\sigma(P)$ that corresponds to the same graph G. This holds for all $G \in \mathcal{G}(\boldsymbol{d})$. Since we assume that the pairing P is simple, $\sigma(P)$ is a unique pairing for each σ. Furthermore, for a given assignment of half edges to bins, all pairings $\mathcal{C}(\boldsymbol{d})$ that correspond to G can be represented as $\sigma(P)$ for some σ that fixes every bin $B(v)$. Thus, the number of pairings in $\mathcal{C}(\boldsymbol{d})$ that correspond to G is the number of ways of permuting the half edges in each bin, which is

$$\prod_{i=1}^{n} d(i)!.$$

Note that this property only depends on the degree sequence $\boldsymbol{d}$ and that G is simple. Therefore, the probability that a uniformly random pairing from $\mathcal{C}(\boldsymbol{d})$ corresponds to G is proportional to $\prod_{i=1}^{n} d(i)!$, for all $G \in \mathcal{G}(\boldsymbol{d})$. Since these graphs all have the same degree sequence, this probability is the same for all of them. Thus, if we condition that the corresponding graph is simple, all simple graphs are equally likely. □

On paper this sounds extremely useful, as random graphs with fixed degree sequences are hard to sample uniformly and random matchings of points are much easier to generate. Indeed this is the case, and there is an extensive body of literature dedicated to studying the configuration model and its implications for random graphs. However, the configuration model does have its

limitations. While Proposition 2.2.4 guarantees uniformity of the output if you condition on a simple pairing, this does not extend to uniformity over all multigraphs, as the argument does not hold if the pairing P contains multiple edges or loops. Furthermore, even if $\mathcal{C}(\boldsymbol{d})$ a.a.s. has some particular property, translating the result to the analogous claim for $\mathcal{G}(\boldsymbol{d})$ requires the probability that the pairing is simple to be bounded away from zero. Janson [75] gives necessary and sufficient conditions on the degree sequence of the pairing for this to be true.

Lemma 2.2.5. ([75, 77], Theorem 1.1) Let $G^*(\boldsymbol{d})$ be the multigraph corresponding to a uniform random pairing $P \in \mathcal{C}(\boldsymbol{d})$. Then

$$\liminf \mathbb{P}\left(G^*(\boldsymbol{d}) \text{ is simple}\right) > 0 \iff \sum_{i=1}^{n} d(i)^2 = O\left(\sum_{i=1}^{n} d(i)\right).$$

Vaguely speaking, this lemma means that if the degree sequence contains any large elements, or many elements that are somewhat large, then results about $\mathcal{C}(\boldsymbol{d})$ do not immediately imply analogous results for $\mathcal{G}(\boldsymbol{d})$. For degree sequences $\boldsymbol{d}$ that do satisfy this condition, this lemma implies that every event that happens asymptotically almost surely in $\mathcal{C}(\boldsymbol{d})$ also happens asymptotically almost surely in $\mathcal{G}(\boldsymbol{d})$.

Despite the limitations, the configuration model is a useful and widely utilised tool for studying graphs and networks [113, 24, 18, 71]. The configuration model has given insight into many graph problems which are difficult to solve in $\mathcal{G}(\boldsymbol{d})$. For example, the threshold for when a random $\mathcal{G}(\boldsymbol{d})$ graph is connected is currently unknown for many degree sequences (see Lemma 2.2.8, given in [56], for a contemporary result). However, Federico and Van Der Hofstad [45] characterised the threshold for when a random pairing is a.a.s. connected for all degree sequences. Beyond the study of graphs, the configuration model can be used to generate other objects randomly, and has also been used to study random hypergraphs [31] as well as random contingency tables [9].

Another possible correspondence between pairings and simple graphs is what is known as the erased configuration model [24], in which all loops (an edge where both endpoints are in the same bin) are deleted and all multiple edges are merged into a single edge to give a (non-uniformly distributed) simple graph. The notable downsides to this method are twofold: the resulting simple graph does not necessarily have the desired degree sequence, nor is it uniformly distributed. However, this method still has its uses. Firstly, under certain conditions on the degree sequence its distribution is asymptotically correct [24]. Furthermore, Janson [78] showed that, under certain conditions, the corresponding multigraph can be adjusted by using switchings (a concept we touch on later, albeit in a different context) to yield a random graph that is almost uniform, in the sense that the total variational distance between the distribution of the resulting graph and the uniform distribution is $o(1)$. Thus, the configuration model still gives valuable insights into random graphs.

2.2.4 Asymptotic notation

As discussed earlier, we are generally interested in studying properties of random graphs as the graph grows arbitrarily large. Thus, typically we consider a sequence of degree sequences indexed by n, the number of vertices in the corresponding graphs. The asymptotic notation we use is defined in terms of n, specifically as $n \to \infty$. Note that if $\boldsymbol{d}$ has a minimum degree of 1 and $M = M(\boldsymbol{d})$ is the total degree of $\boldsymbol{d}$, then $M \in [n, n(n-1)]$. This implies that $M \to \infty$ iff $n \to \infty$.

Thus, if we assume that $\boldsymbol{d}$ has a minimum degree of 1, it is equivalent to define the asymptotics with respect to $n \to \infty$ or $M \to \infty$.

We use the following definitions for Landau notation, as given by Wormald [63]. These definitions are equivalent to the standard definitions but allow us to more easily talk about bounds which apply asymptotically almost surely. Suppose that $|f| < \phi g$, where f, g, and ϕ are functions of n or M. Then if ϕ is bounded by a constant, we say that $f = O(g)$ (equivalently, $g = \Omega(f)$, using the definition of Knuth [91]). If $g = O(f)$ also, then we say that $f = \Theta(g)$. If $\phi \to 0$ as n goes to infinity, then we say that $f = o(g)$ (equivalently, $g = \omega(f)$). We say that $f \sim g$ if $f - g = o(f)$ (equivalently, $o(g)$).

We often write statements such as "a.a.s. $a = O(b)$" or "a.a.s. $a \sim b$" for two functions a and b (which are implicitly dependent on n or M and may be random variables). It is not immediately clear that "$a = O(b)$" is a well defined event on a probability space, and so we need to take some care. We write this using the conventions of Janson [76]: when we write "a.a.s. $a = O(b)$" we mean that there exists a constant $C > 0$ such that a.a.s. $|a| \leq Cb$, or equivalently $\mathbb{P}(|a| \leq Cb) \to 1$ as $n \to \infty$. Similarly, if we write "a.a.s. $a = o(b)$", we mean that there exists a sequence $\delta := \delta(n) \to 0$, such that $\mathbb{P}(|a| \leq \delta b) \to 1$ as $n \to \infty$. Thus, when we write "$a \sim b$ a.a.s.", we mean that $a - b = o(a)$ a.a.s. (equivalently, $o(b)$). Analogous definitions hold for $\Omega(\cdot)$ and $\omega(\cdot)$.

Some equalities and inequalities used throughout this thesis only hold asymptotically, that is, hold for n sufficiently large. In this case, there exists some N, not explicitly stated, such that for all $n > N$ (equivalently $M > N$) the inequality holds. When an inequality only holds for sufficiently large n or M, it is stated to hold asymptotically, or for n or M sufficiently large.

Another caveat worth mentioning is that in some places (particularly in examples) we want to define degree sequences in terms of functions of n or M. For example, one might want to consider a graph where every vertex has degree $\log M$, or a graph with $n^{1/2}$ vertices of degree $n^{1/2}$. In such cases, there are reasonable questions about whether such a sequence is graphical, or whether $\log M$ and $n^{1/2}$ are even integers. One option is to consider only such examples valid for values of n and M where everything is graphical. Otherwise, we can make this rigorous by adding a $O(1)$ term to each quantity to make it an integer and to make the sequence graphical. That is, in the first example we would really be considering a graph where every vertex has degree $\log M + O(1)$, and in the first example we would be considering a graph with $n^{1/2} + O(1)$ vertices of degree $n^{1/2} + O(1)$. For our purposes, this only comes up when discussing examples of sequences that inform our intuition. In these cases, we do not explicitly mention that we add $O(1)$ to these terms, but it is considered to be there implicitly.

2.2.5 Properties of random graphs

Here we give definitions for each graph property that we look at in our study of induced subgraphs of uniformly random graphs with a given degree sequence. In each case, we define the property, and give a brief overview and history of the literature about how this graph property behaves in various random graph models.

Degree sequence

Recall the definition of a graphical sequence from Definition 2.2.2. Erdős and Gallai [40] gave necessary and sufficient conditions for a sequence to be graphical.

Theorem 2.2.6. ([40]) A sequence $\boldsymbol{d}$ containing n elements ordered in non-decreasing order is graphical if and only if $\sum_{i=1}^{n} d(i)$ is even and

$$\sum_{i=0}^{k-1} d(n-i) \leq k(k-1) + \sum_{i=1}^{n-k} \min\{d(i), k\}$$

for all $k \leq n$.

The requirement that $\boldsymbol{d}$ is ordered is somewhat trivial: we can permute the elements of $\boldsymbol{d}$ to obtain a sequence $\boldsymbol{d'}$ (which corresponds to permuting the vertex labels of the underlying graph) ordered in non-decreasing order and apply the condition to $\boldsymbol{d'}$. On the more constructive side of things, the Havel-Hakimi algorithm [67, 64] either constructs a simple graph with degree sequence $\boldsymbol{d}$ if it is graphical, or proves that such a graph does not exist if $\boldsymbol{d}$ is not graphical, and does so in $O(n^2)$ time.

Erdős and Rényi [41, 42] began the study of the degree sequence of a random graph, a topic that has since received much attention. Bollobás [11] gave ranges $W(k)$ such that a.a.s. the degree of the k^{th} largest degree vertex in $\mathcal{G}(n,p)$ lies in $W(k)$, for all $k \leq \frac{1}{2}n$, as well as results about when $\mathcal{G}(n,p)$ has a unique maximum degree vertex and the distribution of the number of vertices of a given degree [14]. This work was later elaborated on by Palka [117, 118]. McKay and Wormald [111, 112] showed for various ranges of p (such as $p(1-p) = \omega(\log n/n^2)$ and $p = o(\sqrt{n})$) that the degree sequence of $\mathcal{G}(n,p)$ can be approximated by a sequence of n random variables $X_n \sim \text{Bin}\,(n-1, p')$ (meaning X_n is distributed as a binomial random variable with $n-1$ trials and success probability p') where p' is a carefully defined random variable. They also showed that the degree sequence of $\mathcal{G}(n,M)$ can be closely approximated by a sequence of binomial random variables, conditional on the event that their sum is equal to M. This result has recently been extended by Liebenau and Wormald [97] to much wider ranges, where $p(1-p) = \omega(\log^3 n/n^2)$ or $\min\{m, \binom{n}{2} - m\} = \omega(\log n)$ depending on whether $\mathcal{G}(n,p)$ or $\mathcal{G}(n,M)$ is being considered.

Connectivity

Another fundamental property of a graph is whether it is *connected*, that is, whether all vertices belong to the same component of the graph.

Definition 2.2.7. A graph G is connected if, for all $u, v \in V(G)$, there exists a path from u to v in G.

Gilbert [58], and independently Erdős and Rényi [42], showed that $\mathcal{G}(n,p)$ is a.a.s. connected if $p > (1+\varepsilon)\frac{\log n}{n}$ for some $\varepsilon > 0$, and a.a.s. disconnected if $p < (1-\varepsilon)\frac{\log n}{n}$. That is, $\frac{\log n}{n}$ is a sharp threshold for connectedness in $\mathcal{G}(n,p)$. This coincides with the threshold for the disappearance of isolated vertices in $\mathcal{G}(n,p)$. Around the threshold, $\mathcal{G}(n,p)$ is neither a.a.s. connected nor disconnected: if $p = (\log n + c)/n$ for some constant c, then the probability that $\mathcal{G}(n,p)$ is connected converges to $\exp(-\exp(-c))$ Erdős and Rényi [42]. More specifically, Ling [98] gave the exact

distribution of the number of edges that need to be added to $\mathcal{G}(n,p)$ such that the resulting graph is connected.

For graphs with given degree sequences, Wormald [130] showed that a random d-regular graph is a.a.s. connected if d is some constant greater than 3. More generally, he showed that if $\boldsymbol{d}$ is a degree sequence with minimum degree at least d and maximum degree at most D, for all constants $d \geq 3$ and D, then $\mathcal{G}(\boldsymbol{d})$ is a.a.s. d-connected, meaning that it has no vertex cutset containing less than d vertices. Łuczak [100] extended this to the case where the sequence has minimum degree at least d and the maximum degree is unbounded, and also gave the probability that a $\mathcal{G}(\boldsymbol{d})$ is connected in the case that $d = 2$. For regular graphs, Frieze [49] showed that a random d-regular graph is a.a.s. Hamiltonian (and thus connected) for $d = o(n^{1/5})$. This was then extended by Cooper et al. [35] to show that this holds for all $d \in [3, c_0 n]$ for some sufficiently small constant c_0. For nearly regular graphs (specifically $\boldsymbol{d}$ with average degree $d(\boldsymbol{d}) = \omega(\log n)$ and $d(i) \sim d(\boldsymbol{d})$ for all i), Gao et al. [55] proved that $\mathcal{G}(\boldsymbol{d})$ is Hamiltonian by coupling the generation of $\mathcal{G}(\boldsymbol{d})$ with two binomial random graphs $\mathcal{G}(n,p_1)$ and $\mathcal{G}(n,p_2)$ such that a.a.s. $\mathcal{G}(n,p_1) \subseteq \mathcal{G}(\boldsymbol{d}) \subseteq \mathcal{G}(n,p_2)$ for p_1 and p_2 that are close. The gaps in both time and generality between results for $\mathcal{G}(n,p)$ and $\mathcal{G}(\boldsymbol{d})$ random graph models highlight the extra difficulty that comes with studying random graphs with given degree sequences.

As mentioned earlier, Federico and Van Der Hofstad [45] recently gave the precise window for connectivity in the configuration model $\mathcal{C}(\boldsymbol{d})$. They showed that the probability that $\mathcal{C}(\boldsymbol{d})$ is connected is a function of the number of degree 1 vertices, the number of degree 2 vertices, and the average degree of $\boldsymbol{d}$; their result also shows that in the case that the average degree is $\omega(1)$ the number of degree 2 vertices is not relevant. This result carries over to the random graph model $\mathcal{G}(\boldsymbol{d})$ when $\boldsymbol{d}$ satisfies the conditions of Lemma 2.2.5. This was then expanded upon by Gao and Ohapkin [56] to carry over to the random graph model for a much larger range of degree sequences $\boldsymbol{d}$. They give the following result in their paper. Again, the important factors are the number of degree 1 and 2 vertices in the graph, as well as the total degree.

Lemma 2.2.8. ([56], Theorem 10 & Corollary 11) Let $\boldsymbol{d}$ be a graphical sequence with $\Delta^2 = o(M)$ and minimum degree at least 1, and let $c > 0$ be a fixed constant. Then

(a) if $n_1(\boldsymbol{d}) = o\left(\sqrt{M}\right)$ and $n_2(\boldsymbol{d}) = o(M)$, then $\mathcal{G}(\boldsymbol{d})$ is a.a.s. connected,

(b) if $n_1(\boldsymbol{d}) = \omega\left(\sqrt{M}\right)$ then a.a.s. $\mathcal{G}(\boldsymbol{d})$ is disconnected,

(c) if $n_1(\boldsymbol{d}) \geq c\sqrt{M}$ or $n_2(\boldsymbol{d}) \geq cM$, then there exists $\delta = \delta(c) > 0$ such that for all sufficiently large n, $\mathbb{P}\left(\mathcal{G}(\boldsymbol{d}) \text{ disconnected}\right) \geq \delta$.

This connectivity threshold also doubles as the threshold for the disappearance of isolated K_2 and K_3 subgraphs in $\mathcal{G}(\boldsymbol{d})$. This hints at an underlying similarity to the connectivity threshold for $\mathcal{G}(n,p)$: while $\mathcal{G}(\boldsymbol{d})$ cannot have isolated vertices (unless $\boldsymbol{d}$ contains zeroes), the connectivity threshold still corresponds to the a.a.s. disappearance of minuscule, isolated subgraphs.

Chromatic number

In 1852, Francis Guthrie conjectured that all maps could be coloured using at most four colours such that no two countries of matching colour shared a border. In graph theory circles this became known as the Four Colour Conjecture, and has a long history of work associated with it. Over a

century later, this claim was proved, first by Appel and Haken [5, 6], with a simpler proof given later by Robertson et al. [122]. This question inspired the more general study of graph colouring and the chromatic number of graphs.

Definition 2.2.9. The chromatic number of a graph, denoted $\chi(G)$, is the smallest number of colours needed to colour the vertex set of the graph G such that for all $\{u, v\} \in E(G)$, u and v receive different colours.

Far from being an esoteric property concerning maps, graph colouring has evolved into one of the most extensively studied fields in graph theory, with far-reaching applications both within graph theory research and in applications to other fields [29, 94, 95]. One of the earliest results on the chromatic number of a random graph was given by Grimmett and McDiarmid [62]. They showed that for constant p, the random graph $\mathcal{G}(n, p)$ a.a.s. has chromatic number

$$(1 + o(1))\frac{n}{2\log_b n} \leq \chi(G) \leq (1 + o(1))\frac{n}{\log_b n},$$

where $b = \frac{1}{1-p}$. They also conjectured that the true asymptotic value lies near the lower bound, which was later proved by Bollobás [17] and independently by Matula [105]. It was known at the time that the $o(1)$ function in the lower bound is actually quite small: for example, Shamir and Spencer [126] showed that $\chi(\mathcal{G}(n, p))$ is concentrated in a window of size $\sqrt{n}\omega$, for some function $\omega \to \infty$ growing arbitrarily slowly. Along with this, they showed that if $p < n^{-5/6-\varepsilon}$ for some fixed $\varepsilon > 0$, then the concentration is much stronger: in fact, $\chi(\mathcal{G}(n, p))$ is a.a.s. one of five consecutive integers (known as 5-point concentration). This was improved by Łuczak [102] to show that $\chi(\mathcal{G}(n, p))$ is a.a.s. one of two integers, and this idea was extended by Alon and Krivelevich [4] to prove 2-point concentration for $p = n^{-1/2-\varepsilon}$. Notably, the particular values of these integers were unknown. In the case that $p = d/n$, the search for these values culminated in Achlioptas and Naor [1] finding the specific two integers on which $\chi(\mathcal{G}(n, p))$ is concentrated: if k is the smallest integer such that $d < 2k \log k$, then a.a.s. the chromatic number is either k or $k + 1$. Similar x-point concentration results for small x have also been proved for wider ranges of p. For example, Coja-Oghlan et al. [34] show that if $p \leq n^{-3/4-\varepsilon}$ for some constant $\varepsilon > 0$, then the chromatic number of $\mathcal{G}(n, p)$ is a.a.s. concentrated on one of three values, as well as establishing 2-point concentration of $\chi(\mathcal{G}(n, p))$ for a subset of this range. Recently, Kargaltsev et al. [86] show that this 2-point concentration holds for $p \leq n^{-3/4-\delta}$ where $\delta \in \left(0, \frac{1}{4}\right]$, proving the highly sought 2-point concentration for almost all of the remaining values of p that Coja-Oghlan et al. examined.

In the case of fixed, constant p, the distribution of $\chi(\mathcal{G}(n, p))$ is quite different. Improvements on the bound of Bollobás were given by McDiarmid [106], followed by an improved upper bound by Fountoulakis et al. [48] and an improved lower bound by Panagiotou and Steger [119]. Most recently, Heckel [68] gave improved upper and lower bounds, in particular bounds that match up to a $o(1)$ term in the denominator. In stark contrast to the case where $p = o(1)$, the chromatic number of $\mathcal{G}(n, p)$ for fixed p is not concentrated on a finite set of values. Indeed, Heckel [69] showed that the chromatic number of $\mathcal{G}(n, \frac{1}{2})$ is not concentrated on fewer than $n^{1/4-\varepsilon}$ consecutive values. In a recent arXiv publication, Heckel and Riordan [70] show that for some values of n the width of the region of a.a.s. concentration is at least $n^{1/2-o(1)}$, which approximately matches the upper bound of $\sqrt{n}/\log n$ given by Noga Alon [116] (Section 7.9, Exercise 3).

For graphs with given degree sequences, Frieze and Łuczak [53] showed that a random d-regular graph a.a.s. has chromatic number $\frac{d}{2\log d}(1+o(1))$ when d is bigger than some constant d_0 and $d = o(n^{1/3-\varepsilon})$ for some fixed $\varepsilon > 0$. In the case where d is much larger, Krivelevich et al. [92] showed that if $d \in \left[n^{6/7+\varepsilon}, 0.9n\right]$ then a.a.s. a random d-regular graph has chromatic number $\frac{n}{2}\log_b d(1+o(1))$, where $b = n/(n-d)$. Notably, this coincides with the chromatic number of $\mathcal{G}(n,p)$ when $p = d/n$, that is, when the binomial random graph has average degree asymptotically equal to d. The intermediate gaps in the values of d were proved by Cooper et al. [36], who showed that a d-regular graph has chromatic number $d/2\log d$ for all $d \in \left[3, n^{1-\eta}\right]$ for an arbitrarily small constant $\eta > 0$.

A natural question that follows from all these results is whether there exists an analogous value to which $\chi(\mathcal{G}(\boldsymbol{d}))$ converges to for irregular degree sequences $\boldsymbol{d}$. If $\boldsymbol{d}$ is not regular, the first step is to consider what to replace the value d with - one logical choice would be the average degree, which we denote $d(\boldsymbol{d})$. One might then hope that it is true that $\mathcal{G}(\boldsymbol{d})$ a.a.s. has chromatic number $d(\boldsymbol{d})/2\ln d(\boldsymbol{d})$, or something similar. Frieze et al. [52] gave a range of degree sequences for which $\chi(\mathcal{G}(\boldsymbol{d})) = \Theta(d(\boldsymbol{d})/\ln d(\boldsymbol{d}))$ a.a.s., and Gao and Ohapkin [56] recently improved the range of degree sequences for which this result holds. We give the latter result below. For a sequence $\boldsymbol{d}$ ordered in non-increasing order, define

$$D_k(\boldsymbol{d}) = \sum_{i=1}^{k} d(i),$$

the sum of the k highest-degree elements in a sequence $\boldsymbol{d}$.

Lemma 2.2.10. ([56], Theorem 9) Let $\boldsymbol{d}$ be an n-element graphical sequence. Suppose that $\boldsymbol{d}$ satisfies the following conditions:

(a) $\Delta(\boldsymbol{d}) = o(n)$,

(b) $D_{\Delta(\boldsymbol{d})}(\boldsymbol{d}) = o(M)$, and

(c) there exist constants $\alpha \in \left(\frac{1}{2}, 1\right)$ and $\varepsilon, K_0 \in \mathbb{R}^+$ such that $D_k(\boldsymbol{d}) \leq K_0 dn\left(\frac{k}{n}\right)^{\alpha}$ for all $k \in \{1, \ldots, \varepsilon n\}$.

Then a.a.s. $\mathcal{G}(\boldsymbol{d})$ has chromatic number $\Theta\left(\frac{d(\boldsymbol{d})}{\ln d(\boldsymbol{d})}\right)$.

In their paper, Frieze and Łuczak [53] also give an example of a sequence for which it is definitely not true that a.a.s. $\chi(\mathcal{G}(\boldsymbol{d})) = \Theta(d(\boldsymbol{d})/\ln d(\boldsymbol{d}))$. Thus, for random graphs with given, irregular degree sequences, it is unclear if there is a single function that $\chi(\mathcal{G}(\boldsymbol{d}))$ should converge to for all $\boldsymbol{d}$, let alone what such a function would be.

Automorphism group

One thing to note is that in this thesis we only consider labelled graphs. That is, every vertex in $V(G)$ has a unique label, and thus the two graphs in Figure 2.2 are not the same labelled graph. However, these graphs are obviously similar - they have the same "shape", but the labels have been rearranged. By permuting the label set, we can transform the first graph into the second. More formally, this is called an isomorphism. An isomorphism that preserves the edge set of the graph is called an *automorphism*.

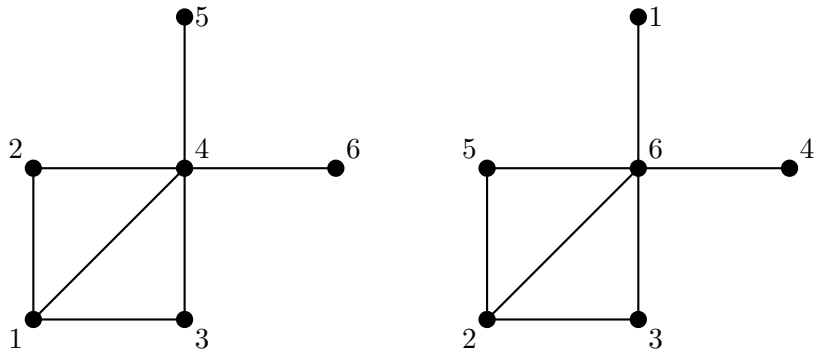

Figure 2.2: Two graphs that differ by a permutation of the vertex set, specifically the permutation $(1\ 2\ 5)\ (4\ 6) \in S_6$. These graphs are isomorphic, however this permutation is not an automorphism as the graphs have different edge sets (for example, the second graph does not contain the edge $\{1, 2\}$).

Definition 2.2.11. An automorphism of a graph $G = (V, E)$ is a permutation $\sigma : V \to V$ such that $\{\{\sigma(u), \sigma(v)\}\}_{\{u,v\}\in E} = E$. That is, σ is a permutation of the vertex set V that preserves the edge set E. The set of automorphisms for a given graph G is known as the automorphism group of G.

Every graph has at least one automorphism: the identity permutation, which maps each vertex to itself. This is often called the trivial automorphism, for fairly obvious reasons. The study of graph automorphisms (also sometimes called graph *symmetries*) is a topic of graph theory in and of itself, but automorphisms are also important in other areas of graph theory, such as enumerating graphs. Every labelled graph is uniquely determined by its edges, and thus the number of labelled graphs on n vertices is $2^{\binom{n}{2}}$. But what about unlabelled graphs? Two labelled graphs might correspond to the same unlabelled graph - in fact, up to $n!$ labelled graphs might correspond to the same unlabelled graph. By determining the automorphism group of each graph, we can fix this double counting.

The size of the automorphism group is another long-studied property of random graphs. Erdős and Rényi [43] showed that if $\min\{p, 1-p\} \geq (1+\varepsilon)\frac{\log n}{n}$, then $\mathcal{G}(n, p)$ is a.a.s. asymmetric. Keeping in mind that an automorphism of G is an automorphism of the complement graph $\overline{G}$, this aligns with the threshold for both $\mathcal{G}(n, p)$ and its complement to be a.a.s. connected. If p does not satisfy this condition, then a.a.s. either $\mathcal{G}(n, p)$ or its complement contain multiple isolated vertices, which gives rise to symmetries which permute the labels of these isolated vertices. This link between connectivity and non-trivial automorphisms is further seen in results of Bollobás [13], who showed that a random d-regular graph is a.a.s. asymmetric if $d \geq 3$ and d is a constant, which aligns with the connectivity results for random regular graphs discussed earlier. Note that if $d \in \{0, 1, 2\}$ then a random d-regular graph is symmetric with probability 1, and the similar results follow for $d \in \{n-3, n-2, n-1\}$ by considering the complement. McKay and Wormald [110] extended the result of Bollobás to $d = o(\sqrt{n})$, as well as extending to irregular degree sequences. The characterisation of the a.a.s. existence of non-trivial symmetries in random regular graphs was then completed by Kim et al. [89], who showed that if $d \in [3, n-4]$ then a random d-regular graph is a.a.s. asymmetric.

For irregular degree sequences, McKay and Wormald [110] showed that $\mathcal{G}(\boldsymbol{d})$ is a.a.s. asym-

metric if it contains sufficiently low numbers of degree 1 and 2 vertices. One easy-to-state result they give is as follows: if $\boldsymbol{d}$ is a bounded degree sequence (that is, $\Delta(\boldsymbol{d})$ is at most a constant), $n_1(\boldsymbol{d}) = O(n^{1/3})$, and $n_2(\boldsymbol{d}) = O(n^{2/3})$, then a.a.s. $\mathcal{G}(\boldsymbol{d})$ has no non-trivial automorphisms. Brick et al. [23] extend this result using similar methods. They precisely determine the threshold for non-trivial symmetries in the case where the maximum degree of $\boldsymbol{d}$ is bounded by a constant.

Lemma 2.2.12. ([23], Theorem 2) Fix a constant $\Delta > 0$ and assume that $\boldsymbol{d}$ is a graphical sequence where $1 \leq d(i) \leq \Delta$ for all $1 \leq i \leq n$.

(a) If $n_1 = o(n^{1/2})$ and $n_2 = o(n)$ then a.a.s. $\mathcal{G}(\boldsymbol{d})$ is asymmetric.

(b) If $n_1 = \omega(n^{1/2})$ then a.a.s. $\mathcal{G}(\boldsymbol{d})$ is symmetric.

(c) If there is a constant $c > 0$ such that

$$n_1 > cn^{1/2} \quad \text{or} \quad n_2 > cn$$

then there is $\delta = \delta(c) > 0$ such that for all sufficiently large n

$$\mathbb{P}\left(\mathcal{G}(\boldsymbol{d}) \text{ is symmetric}\right) > \delta.$$

It is noteworthy that this threshold is the same as the connectivity threshold for bounded degree sequences determined by Federico and Van Der Hofstad [45]. This echoes the overlap of the thresholds for connectivity and non-trivial automorphisms in $\mathcal{G}(n,p)$. In contrast to the connectivity threshold of Gao and Ohapkin given in Lemma 2.2.8, if the maximum degree is some growing function of n then the threshold for the existence of non-trivial automorphisms also depends on this maximum degree. As an extreme case, consider a degree sequence $\boldsymbol{d}$ which contains only two degree 1 terms but also a term of degree $n-1$. Then $\mathcal{G}(\boldsymbol{d})$ has a non-trivial automorphism with probability 1: a permutation of the vertex labels that swaps the two degree 1 vertices but fixes the rest of the graph is a non-trivial automorphism, since both degree 1 vertices are adjacent to the degree $n-1$ vertex with probability 1. Intuitively, the reason for this dependence is that the presence of vertices of high degree increases the chances that many degree 1 and 2 vertices are adjacent to these high-degree "hub" vertices. This in turn increases the chance of small, symmetric structures appearing in the random graph. In this sense, there is an intuitive synergy between high-degree vertices and degree 1 and 2 vertices when it comes to generating graphs with non-trivial automorphisms. The results of Brick et al. [23] that apply to degree sequences with larger maximum degree are given below. The first theorem gives sufficient conditions on $\boldsymbol{d}$ for $\mathcal{G}(\boldsymbol{d})$ to a.a.s. have no non-trivial automorphisms.

Lemma 2.2.13. ([23], Theorem 3) Suppose there are constants $R_1 > 0$, $R_2 > 0$ and $0 < \varepsilon < 1$ such that a graphical sequence $\boldsymbol{d}$ with minimum degree at least 1 satisfies the following conditions:

(A1) $\frac{\Delta^2}{d} = o(n^{\frac{1}{6} - \frac{1}{2R_1} - \frac{1}{R_2}})$,

(A2) $\frac{\Delta^2}{d} = o\left(\frac{n^{1/4}}{n_1^{1/2}}\right)$,

(A3) $\frac{\Delta^2}{d} = o\left(\frac{n^{\alpha_2/2}}{n_2^{\alpha_2/2}}\right)$,

(A4) $\left(\frac{n_i}{n}\right)^{\alpha_i(1-\varepsilon)}\left(\frac{\Delta^2}{d}\right)^{2-\varepsilon} = o(1)$, for $i \in \{1,2\}$,

where $\alpha_2 = 1/(R_2+4)$, and $\alpha_1 = (1-\alpha_2)/(R_1+4)$. Then $\mathcal{G}(\boldsymbol{d})$ is a.a.s. asymmetric.

The next lemma deals with the other case, giving sufficient conditions on $\boldsymbol{d}$ for $\mathcal{G}(\boldsymbol{d})$ to a.a.s. have a non-trivial automorphism. The proof of this lemma measures the likelihood of a few small, symmetric structures (such as isolated K_2 subgraphs or K_3 subgraphs with two degree 2 vertices) to appear in $\mathcal{G}(\boldsymbol{d})$. For the following theorem, recall that $M = M(\boldsymbol{d}) = \sum_{i=1}^{n} d(i)$ and define $M_2(\boldsymbol{d}) = \sum_{i=1}^{n} d(i)(d(i)-1)$.

Lemma 2.2.14. ([23], Theorem 4) Let $\boldsymbol{d}$ be an n-vertex graphical sequence with minimum degree at least 1 and assume $\Delta^2 = o(M)$.

(a) If $M_2 = o(M)$, $n_1 = \omega\left(M/\sqrt{M_2}\right)$, or $n_2 = \omega\left(\sqrt{M^3/M_2}\right)$ then a.a.s. $\mathcal{G}(\boldsymbol{d})$ is symmetric.

(b) If there is a constant $c > 0$ such that

$$n_1 > cM/\sqrt{M_2}, \quad \text{or} \quad n_2 > c\sqrt{M^3/M_2},$$

then there exists $\delta = \delta(c) > 0$ such that for all sufficiently large n,

$$\mathbb{P}\left(\mathcal{G}(\boldsymbol{d}) \text{ symmetric}\right) > \delta.$$

Existence of a giant component

The component structure of random graphs has been extensively studied in various random graph models since the early days of random graphs. For $\mathcal{G}_{n,M}$ (uniform random graphs with n vertices and M edges), Erdős and Rényi [42] showed the existence of what is classically called the "double jump" threshold: if $M = cn$ for some constant $c < \frac{1}{2}$, then a.a.s. all components have $O(\log n)$ vertices, whereas if $c > \frac{1}{2}$ then a.a.s. there exists a single component of order at least εn, for some constant $\varepsilon > 0$, with every other component of size $O(\log n)$. If $c = \frac{1}{2}$, then a.a.s. the largest component has size $\Theta(n^{\frac{2}{3}})$. Bollobás [15] and Łuczak [101] explored the growth of the largest component in the realm where $M \sim \frac{1}{2}n$, as well as giving estimates for the order of the k^{th} largest component. The component of order at least εn is referred to as the *giant component*. Since this result, there have been a plethora of results about thresholds for the existence of giant components in many random graph models [32, 127, 38, 128, 54].

The first result about giant components in graphs with given degree sequences was by Molloy and Reed [113], who gave a characterisation for sparse graphs. As is the case with many results on $\mathcal{G}(\boldsymbol{d})$, their results came about by translating results from the configuration model. In their paper, they analyse a random process which explores the random graph using a breadth-first search. Analogously, this can be viewed as generating the graph using a deferred decision process, revealing adjacencies stepwise as the graph is "explored" by the breadth-first search algorithm. They concluded for graphs with low maximum degree that a giant component exists a.a.s. when $\mathbb{E}\left[D(D-2)\right]$ is positive and bounded away from zero, where D is the degree of a uniformly randomly selected vertex. Conversely, if this expectation is negative, then almost surely a random graph with degree sequence $\boldsymbol{d}$ does not have a giant component. The intuition for this is straightforward: roughly speaking, in the early steps of the exploration process on $\mathcal{G}(\boldsymbol{d})$, the probability

that a given vertex is the next vertex discovered is proportional to its degree, up to some small error. Discovering a new vertex v consumes 1 half edge in the component currently being explored, but adds $d(v) - 1$ new ones. This is described in the diagram in Figure 2.3. Thus, $\mathbb{E}[D(D-2)]$ corresponds to the expected net gain of half edges at each step early in the exploration process. Thus, if the number of unpaired half edges in the component is expected to decrease at each step, it is much more likely that it reaches zero - meaning the component has been totally explored - before exploring εn vertices for any $\varepsilon > 0$. On the other hand, if the number of unpaired half edges is expected to increase at each step, this means the component explored by the breadth-first search is gaining half edges faster than it is losing them, and it is likely to grow to linear size.

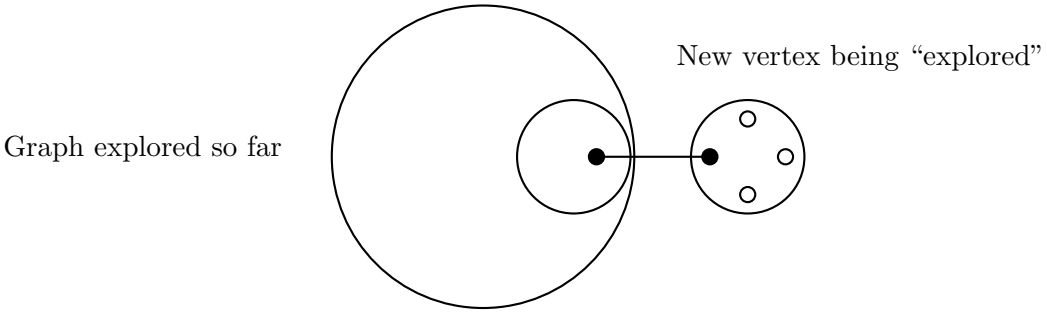

Figure 2.3: A diagram depicting a degree 4 vertex being added to the explored graph at a given time step. The component has a net gain of 2 unpaired half edges.

These methods were adapted by many people to prove results about giant components in both the configuration model and the random graph model for a wide variety of degree sequences [21, 66, 79, 85]. Recently, Joos, Perarnau, Rautenbach, and Reed [83] characterised the threshold for the a.a.s. existence of a giant component in $\mathcal{G}(\boldsymbol{d})$ for every sequence $\boldsymbol{d}$. Their method is also based on a breadth-first search, with extra preliminary steps to deal with the presence of vertices of high degree.

When considering $\mathcal{G}(\boldsymbol{d})$ models that contain high-degree vertices, a whole host of new problems arise. As mentioned earlier, the probabilities of many events are much more easily calculated in the configuration model. However, extra conditions (such as those given in Lemma 2.2.5) are needed to carry the analogous results from $\mathcal{C}(\boldsymbol{d})$ to $\mathcal{G}(\boldsymbol{d})$, which preclude using this method for the degree sequences in question. When working directly in the random graph model, the probabilities of these events are commonly computed via the switching method (which we describe in Section 2.3). This is much more involved, and is both more complicated and less precise in the presence of high-degree vertices.

The other problem is that the presence of high-degree vertices makes the exploration process less "stable" over time. Consider as an example the degree sequence $\boldsymbol{d} = (1, 1, \ldots, 1, n^{2/3})$ (example from Joos et al. [83]). No graph with this degree sequence has a giant component: they all contain a single component with $n^{2/3} + 1$ vertices, with every other component isomorphic to K_2. However, $\mathbb{E}[D(D-2)] = \Theta(n^{4/3})$ in this case, so extrapolating from the Molloy-Reed criterion one might expect a giant component. The issue is that $\mathbb{E}[D(D-2)]$ is only positive because of vertex n. As soon as this vertex is added to explored component, every unexplored vertex has degree 1, and $\mathbb{E}[D(D-2)]$ is no longer representative of the expected net gain in half edges at each step. Instead, the net gain at each subsequent step is -1, as every "undiscovered" vertex

has degree 1.

To combat this problem, the algorithm used in the proof of Joos et al. [83] starts by doing what they call *preprocessing*. This is a method of analysing the high-degree vertices before the breadth-first search process begins which excludes them from being discovered at a later step. The benefit of this is twofold. Firstly, it means that the expected number of half edges (or *open edges*, using the graph-centric terminology of Joos et al.) gained at each step does not change much throughout the early steps of the exploration process. Secondly, by removing high-degree vertices from consideration, the probabilities of various events throughout the process can be much more accurately determined, as the approximations arising from the switching method can be made much more precise.

The other concern that arises when considering arbitrary degree sequences is the curious role of vertices of degree 2. If a degree 2 vertex is discovered at some step in the exploration process, the number of open edges does not change. Because of this, random graphs with large numbers of degree 2 vertices behave differently: they correspond to the case where $\mathbb{E}[D(D-2)] \to 0$ as $n \to \infty$, which was not covered by the original results of Molloy and Reed [113]. The solution of Joos et al. is to create a new graph $H(\boldsymbol{d})$ from $G(\boldsymbol{d})$ where all paths of degree 2 vertices are removed and replaced with a single edge (potentially creating a multigraph). Then the exploration process is analysed to determine the component structure of $H(\boldsymbol{d})$, and a relationship between the giant components of $G(\boldsymbol{d})$ and $H(\boldsymbol{d})$ is established. Their result is given below. Notably, in this result the $o(1)$ terms are uniform across all n-vertex degree sequences. They also apply this theorem to give a result explicitly written in terms of sequences of sequences, which we state later in Theorem 4.4.1.

Theorem 2.2.15. ([83], Theorems 1 and 6) Let $\boldsymbol{d} = (d(1), \dots, d(n))$ with $d(1) \leq d(2) \leq \cdots \leq d(n)$. Define the following quantities:

$$j_{\boldsymbol{d}} = \min\left(\left\{j : j \in [n] \text{ and } \sum_{i=1}^{j} d(i)(d(i)-2) > 0\right\} \cup \{n\}\right),$$
$$R(\boldsymbol{d}) = \sum_{i=j_{\boldsymbol{d}}}^{n} d(i),$$
$$\widetilde{M}(\boldsymbol{d}) = \sum_{i \in [n], d(i) \neq 2} d(i).$$

Call a degree sequence *well-behaved* if $\widetilde{M}(\boldsymbol{d})$ is at least $\lambda(n)$ for any function $\lambda : \mathbb{N} \to \mathbb{N}$ where $\lambda \to \infty$ as $n \to \infty$. Then:

(a) For every function $\delta \to 0$ as $n \to \infty$, for every $\gamma > 0$, if $\boldsymbol{d}$ is a well-behaved graphical sequence with $R(\boldsymbol{d}) \leq \delta(n)\widetilde{M}(\boldsymbol{d})$, then the probability that $G(\boldsymbol{d})$ has a component of order at least γn is $o(1)$.

(b) For every positive constant ε, there is a $\gamma > 0$ such that if $\boldsymbol{d}$ is a well-behaved graphical sequence with $R(\boldsymbol{d}) \geq \varepsilon\widetilde{M}(\boldsymbol{d})$, then the probability that $G(\boldsymbol{d})$ has a component of order at least γn and a component of size at least $\alpha\widetilde{M}(\boldsymbol{d})$ is $1 - o(1)$.

(c) For every $b \geq 0$ and every $0 < \gamma < \frac{1}{8}$, there exist a positive integer $n_{b,\gamma}$ and a $0 < \delta < 1$ such that if $n > n_{b,\gamma}$ and $\boldsymbol{d}$ is a degree sequence with $\widetilde{M}(\boldsymbol{d}) \leq b$, then the probability that

there is a component of order at least γn in $G(\boldsymbol{d})$ lies between δ and $1-\delta$.

The intuition behind this criterion is that $R(\boldsymbol{d})$ roughly represents the size that a component reaches in the exploration process before the expected stepwise net gain in open edges becomes negative. This is because, intuitively, this roughly corresponds to when the vertices contributing to the $R(\boldsymbol{d})$ summation have been discovered. At this point, the component does not grow much more, and thus the largest component of the graph is, intuitively, of size roughly $R(\boldsymbol{d})$. It must be noted though that all of this is just intuition - the result of Joos et al. does not attempt to give any meaningful window for the size or order of the largest component beyond whether it is giant or not.

This theorem helps explain why previously found thresholds did not extend to random graphs with high-degree vertices. If large elements are present in the degree sequence, the expected increase in the number of open edges at each step can change drastically quickly. If the high-degree vertices are discovered early (which is, intuitively speaking, likely to be the case) then the expected net gain of open edges at each step can drop below 0 well before the component can grow to be a giant. This characterisation avoids this problem by splitting the sequence at $j_{\boldsymbol{d}}$, effectively separating the contribution of low-degree and high-degree terms in $\boldsymbol{d}$ and counting their contributions in different ways.

One consequence of Theorem 2.2.15 is the following proposition. We give this proposition early, as it is useful at many points throughout the thesis.

Proposition 2.2.16. If $\boldsymbol{d}$ is an n-element graphical sequence such that $M(\boldsymbol{d}) = \omega(n)$, then $R(\boldsymbol{d}) = M(\boldsymbol{d})(1-o(1))$.

This is an immediate consequence of the following claim of Joos et al. [83].

Claim 2.2.17. ([83], Claim 5) If $\boldsymbol{d}$ is an n-element graphical sequence, then $R(\boldsymbol{d}) \geq \widetilde{M}(\boldsymbol{d}) - 2(n - n_2(\boldsymbol{d}))$.

Note that $\widetilde{M}(\boldsymbol{d}) - 2(n-n_2(\boldsymbol{d})) = M(\boldsymbol{d}) - 2n$. So if $M(\boldsymbol{d}) = \omega(n)$, then $R(\boldsymbol{d}) \geq \widetilde{M}(\boldsymbol{d})(1-o(1))$. Thus, Theorem 2.2.15 implies that if $M(\boldsymbol{d}) = \omega(n)$, then $\mathcal{G}(\boldsymbol{d})$ a.a.s. contains a giant component. In their paper, Joos et al. [83] actually show the slightly stronger result that if $\widetilde{M}(\boldsymbol{d}) > 3n$, then $\mathcal{G}(\boldsymbol{d})$ a.a.s. contains a giant component. We do not need this stronger bound, but it gives some intuition about giant components in $\mathcal{G}(\boldsymbol{d})$. If $\widetilde{M}(\boldsymbol{d}) > 3n$, then the average degree of a vertex in G is strictly greater than 2. This intuitively suggests that, on average, the number of open edges gained at each step of the exploration process is positive, and thus the process is likely to explore a large component before running out of open edges.

2.3 The switching method

As alluded to in the previous sections, estimating and determining the probabilities of different events in $\mathcal{G}(\boldsymbol{d})$ is a non-trivial task at the best of times. The probability measure on $\mathcal{G}(\boldsymbol{d})$ is the uniform measure, which means that for some event $A \subset \mathcal{G}(\boldsymbol{d})$ we know that

$$\mathbb{P}(A) = \frac{|A|}{|\mathcal{G}(\boldsymbol{d})|}.$$

Thus, if we know the fraction of graphs with a certain property (for example, the fraction of graphs containing some particular edge uv), then we know the probability of this event. This can also be naturally extended to finding conditional probabilities: for two events A and B, it follows that

$$\mathbb{P}(A|\,B) = \frac{|A \cap B|}{|B|}.$$

Counting the number of graphs with degree sequence $\boldsymbol{d}$ is extremely difficult, and for most properties A (equivalently, events) the counting is similarly difficult. However, we do not need to know the exact values of $|A|$ and $|\mathcal{G}(\boldsymbol{d})|$, or even their approximate values - as long as we have good bounds on the ratio between the two quantities, then we have good bounds on the probability of the event A. This is the premise of the switching method.

For our purposes, a switching is an operation that maps a graph in $\mathcal{G}(\boldsymbol{d})$ to another graph in $\mathcal{G}(\boldsymbol{d})$ by deleting a subset of the edges and replacing them with other edges such that the degree sequence is maintained. For a given switching that maps a graph G_1 to another graph G_2, we call the inverse operation, mapping G_2 to G_1, a reverse switching. Let (A, B) be a partition of $\mathcal{G}(\boldsymbol{d})$. For example, for some $u, v \in V(G)$, let A be the set of graphs where uv is not an edge, and let B be the set of graphs where uv is an edge. We can define a switching that maps an element of A to an element of B by modifying the edges of the graph. There are many different possible switching operations that accomplish this, one such switching is given in Figure 2.4. On the left we have a (small part of a) graph in A, so it does not contain the edge uv. The switching then takes two edges ux and vy, deletes them, and replaces them with edges uv and xy. This creates a new graph with the same degree sequence which contains the edge uv, and is thus in B. A switching

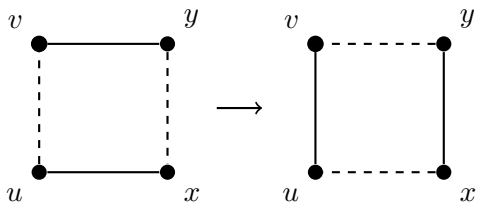

Figure 2.4: An example of a switching operation. Present edges are given as solid lines, forbidden edges are given as dashed. Other edges can be present or absent. The set of vertices and edges which the switching is defined on is sometimes called a *configuration* in the literature. We avoid that term where possible to minimise confusion with terminology for discussing the configuration model $\mathcal{C}(\boldsymbol{d})$.

operation defines a bipartite graph H with vertex set $V(H) = (A, B)$ and edges between graphs $G \in A$ and $G' \in B$ if and only if there exists a switching mapping G to G'. If every graph in A has degree a in H and every vertex in B has degree b, then counting the edges of H from both perspectives gives that

$$a|A| = b|B|. \tag{2.1}$$

This gives the relative of size of A with respect to B. Since (A, B) is a partition of $\mathcal{G}(\boldsymbol{d})$, it follows

that

$$\mathbb{P}(A) = \frac{|A|}{|\mathcal{G}(\boldsymbol{d})|} = \frac{|A|}{|A| + |B|} = \frac{\frac{|A|}{|B|}}{\frac{|A|}{|B|} + 1}.$$

The last form of this expression is useful because Equation (2.1) implies that $|A|/|B| = b/a$, which allows us to determine $\mathbb{P}(A)$ (in this example, the probability that u and v are not adjacent). With a little more effort we can define a similar idea for a partition with more than two parts, and determine their sizes relative to each other. A commonly used variant of this is the following: suppose $\{A_i\}_{i\in[N]}$, for $N \in \mathbb{N}$, is a partition of $\mathcal{G}(\boldsymbol{d})$. Suppose that we can define a switching that takes $G \in A_{i+1}$ to $G' \in A_i$ for all $i < N$ where $a_{i+1}|A_{i+1}| = a_i|A_i|$. Then we can determine the relative sizes of each A_i:

$$|A_i| = |A_{i-1}|\frac{a_{i-1}}{a_i} = \cdots = |A_1|\prod_{j=1}^{i-1}\frac{a_j}{a_{j+1}} = |A_1|\frac{a_1}{a_i}. \tag{2.2}$$

The catch of the switching method is that a and b in Equation (2.1) (and analogously $\{a_i\}_{i\in[N]}$ in (2.2)) are not often easy to determine exactly. In fact, in most applications the degrees (in H) of each element of A will not be equal, and similarly for graphs in B. We can instead think of a and b as functions of the graphs in A and B respectively, where $a(G)$ denotes the degree of G in H, for some $G \in A$. These functions are rarely constant and are usually difficult to determine exactly.

One option, often the simplest, is to find upper and lower bounds for a over all graphs in A, as well as upper and lower bounds for b over all graphs in B. This is often done in two steps. First, an upper bound is found for the number of ways to choose the vertices/edges involved in the switching for each graph G. Dependent on the particular graph G, some of these choices are *invalid* or *forbidden* in the sense that the output of the switching would not be an element of $\mathcal{G}(\boldsymbol{d})$. The second step is to then remove these invalid choices from the total count. For example, suppose we wanted to perform the switching given in Figure 2.4. For some $G \in A$ and choices of vertices x and y, it is possible that $xy \in E(G)$. Then the output of this switching would not be a simple graph. Thus, we must exclude this choice of $\{u, v, x, y\}$ from the lower bound on the number of ways to apply the switching to each $G \in A$.

Suppose that d_A and D_A are the lower and upper bounds on the degree of each $a \in A \subset V(H)$, and suppose that d_B and D_B are the corresponding bounds for B. Then these bounds can be used to obtain expressions similar to Equation (2.1), and to thus obtain upper and lower bounds for $\mathbb{P}(A)$:

$$d_A|A| \leq D_B|B| \quad \text{and} \quad d_B|B| \leq D_A|A|,$$

which implies that

$$\frac{d_B}{D_A} \leq \frac{|A|}{|B|} \leq \frac{D_B}{d_A}.$$

Provided that these bounds are sufficiently close (ideally asymptotically equal), this gives accurate bounds on the probability of the event A. This is the key idea used in Chapters 3 and 7 to

determine various probabilities in the random graph model.

In some circumstances, a small subset of graphs in A (or B) have many less valid switchings that can be applied to them in comparison to the rest of the graphs in that part. If we were to apply what has been discussed so far, the bounds obtained from the switchings may be very coarse, and thus might be insufficient or even trivial. The issue in this case is that $d_A|A|$ is not an accurate lower bound on the number of edges between A and B, since a few bad apples spoil the bunch. To fix this, we can partition A further, say, into A_{good} and A_{bad}, where A_{bad} contains the problematic graphs with a significantly worse lower bound. This way, we can replace $d_A|A|$ with the larger bound

$$d_{\text{Average}}|A| \geq d_{A_{\text{bad}}}|A_{\text{bad}}| + d_{A_{\text{good}}}|A_{\text{good}}|.$$

Then d_{Average} is a (lower bound on the) weighted average of the lower bound over all elements in A, and thus a lower bound on the number of edges between A and B in H. If one can get a good bound on the relative sizes of A_{good} and A_{bad}, this leads to a tighter bound on $|A|/|B|$ (and thus on $\mathbb{P}(A)$), since the small number of problematic graphs do not ruin the bound.

For even more nuanced switching ideas, we turn to Hasheminezhad and McKay [65]. They provide a much more general framework that encapsulates and extends what has been discussed so far. Let $\mathfrak{V}$ be an arbitrary set that indexes a partition $\{A(v)\}_{v\in\mathfrak{V}}$ of $\mathcal{G}(\boldsymbol{d})$. Let $\mathfrak{E}$ be a set of ordered pairs (u, v) of elements in $\mathfrak{V}$. Define a *structure graph* $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ for a given switching operation as follows. Each vertex $v \in \mathfrak{V}$ corresponds to the set $A(v)$. The edge (u, v) (also written uv for brevity, with implied $u \to v$ orientation) is in $\mathfrak{E}$ if and only if there exists an element $G \in A(u)$ and an element $G' \in A(v)$ such that there is a switching mapping G to G'. The structure graph permits loops, but has no multiple edges. Define $N(v) = |A(v)|$, and analogously define $N(V) = \sum_{v\in V} |A(v)|$ for some subset $V \subset \mathfrak{V}$.

Suppose that, on average, a graph in $A(v)$ can be subject to at least $a(v)$ switchings and be created by at most $b(v)$ switchings. Define $\alpha(vw) = b(w)/a(v)$; equivalently we write $\alpha(v \to w)$ to emphasise the direction. For a directed path P in $\mathfrak{G}$, let $\mathfrak{E}(P)$ be the edges of this path. Define

$$\alpha(P) = \prod_{uv\in\mathfrak{E}(P)} \alpha(uv),$$

and define

$$\mathfrak{Z} := \{v \in \mathfrak{V} \mid \alpha(vw) \geq 1 \text{ for some } vw \in \mathfrak{E}, \text{ or } v \text{ is a sink in } \mathfrak{G}\}.$$

Let $\mathfrak{X}$ be a superset of $\mathfrak{Z}$ (where $\mathfrak{Z}$ must be non-empty), and let $\mathfrak{Y}$ be a set that is disjoint from $\mathfrak{X}$. For $U, V \subset \mathfrak{V}$, define $\mathfrak{Q}_{\mathfrak{Y},\mathfrak{Z}}(U, V)$ to be the set of all non-trivial (containing at least two distinct vertices) directed paths in $\mathfrak{G}$ that start in U, end in V, and have no internal vertices in $\mathfrak{Y} \cup \mathfrak{Z}$. Finally, for subsets $A, B \subset \mathfrak{V}$, define

$$\alpha_{\mathfrak{Y},\mathfrak{Z}}(AB) := \max_{P\in\mathfrak{Q}_{\mathfrak{Y},\mathfrak{Z}}(A,B)} \alpha_P. \tag{2.3}$$

Lemma 2.3.1. ([65], Theorem 3) Let $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ be as defined above, along with N, α, $\mathfrak{X}$, $\mathfrak{Y}$,

and $\mathfrak{Z}$. Then

$$\frac{N(\mathfrak{Y})}{N(\mathfrak{X})} \leq \frac{\alpha_{\mathfrak{Y},\mathfrak{Z}}(\mathfrak{Y}\mathfrak{Z})}{1-\alpha_{\mathfrak{Y},\mathfrak{Z}}(\mathfrak{Y}\mathfrak{Y})},$$

where $\alpha_{\mathfrak{Y},\mathfrak{Z}}(AB) = 0$ if $\mathfrak{Q}_{\mathfrak{Y},\mathfrak{Z}}(A,B) = \emptyset$.

This is a large generalisation on the switching method described previously. In this language, the partition $\{A_i\}_{i\in[N]}$ forms the vertex set of the structure graph, which is simply the path $A_1 \dots A_N$. This method allows us to consider switchings to bound multiple related graph parameters simultaneously, and show that a.a.s. at least one of the parameters is within its desired range. As well as this, since loops are permissible in the structure graph, this method also allows us to consider switchings that might not work in the standard format where the structure graph is simply a path. This more generalised switching method will be useful in proving some results in Chapter 7.

2.4 Subgraphs of random graphs

The classic paper of Erdős and Rényi [42] provides a starting point for results on subgraphs of random graphs. They provide a threshold function for the existence of a small graph H as a subgraph of $\mathcal{G}(n,M)$ in the case that H is what they called "balanced", meaning that it contains no subgraphs denser than H itself. For example, trees, cycles, and complete graphs are all balanced in this sense. Matula [104] extended this result for complete graphs, analysing the distribution of complete subgraphs and maximal complete subgraphs of $\mathcal{G}(n,p)$. This was expanded upon by Schurger [125], who showed that the number of K_k subgraphs in $\mathcal{G}(n,p)$ converges to a Poisson distribution. The result of Erdős and Rényi was extended further by Bollobás [12] to give the threshold function for the probability that $\mathcal{G}(n,p)$ contains H for a wider class of graphs H. Specifically, he showed that if $m(H)$ is the density of the densest subgraph of H, then

$$\lim_{n\to\infty} \mathbb{P}\left(H \subset \mathcal{G}(n,p)\right) = \begin{cases} 0 & \text{if } p = o(n^{1-1/m(H)}), \\ 1 & \text{if } p = \omega(n^{1-1/m(H)}). \end{cases}$$

This result suggests that, roughly speaking, once a $\mathcal{G}(n,p)$ graph is likely to contain a copy of the densest subgraph of H, then it likely contains all of H - the sparser parts are "easier" to contain than the dense part. Ruciński [123] provided a general framework for determining the probability that a $\mathcal{G}(n,p)$ random graph has a given graph H as a subgraph, as well as the distribution of the number of copies of H. In particular, he gives conditions under which the number of copies of H is asymptotically distributed as a Poisson or a normal distribution. This result also extends beyond $\mathcal{G}(n,p)$ to arbitrary graphs G where the edges of G are deleted independently with probability p, that is, bond percolation on fixed graphs. In a later paper, Ruciński [124] gave necessary and sufficient conditions for the number of copies of a subgraph H in a $\mathcal{G}(n,p)$ random graph to asymptotically have a normal or Poisson distribution.

As is the case for many graph properties, the problem of subgraph containment is quite different for random graphs with a given degree sequence. McKay [109] gave an early result in this area, giving a series of general results about bounds on the probability that $\mathcal{G}(\boldsymbol{d})$ contains a particular subgraph, along with a particular application to the average number of spanning trees of $\mathcal{G}(\boldsymbol{d})$.

These results only apply in the case where the subgraph is small, and also require that the maximum degree of the whole graph is asymptotically negligible compared to the number of edges. These results were expanded upon by Greenhill and McKay [61], McKay [108], and then later by Greenhill, Isaev, and McKay [60]. In particular, these extensions consider when the graph G is dense, that is, has average degree λn for some constant $\lambda > 0$.

If we restrict to just the case where the degree sequence $\boldsymbol{d}$ is a constant sequence (that is, $\mathcal{G}(\boldsymbol{d})$ is a random regular graph), then Kim et al. [88] gave a threshold function for when $\mathcal{G}(\boldsymbol{d})$ contains a given subgraph H. They showed that $d = n^{1-1/m(H)}$ (where $m(H)$ is again the density of the densest subgraph of H) is the threshold — that is, if $d = o(n^{1-1/m(H)})$, then a.a.s. a random d-regular graph does not contain H as a subgraph, and a.a.s. does if $d = \omega(n^{1-1/m(H)})$. This mirrors the earlier analogous result for $\mathcal{G}(n,p)$ proved by Bollobás. The nearly thirty year gap between the result in $\mathcal{G}(n,p)$ and the result for random regular graphs is a testament to the extra difficulty in working with $\mathcal{G}(\boldsymbol{d})$ and the extra complexity brought about by the edge dependencies inherent in the model.

In the realm of induced subgraphs, Maehara [103] showed that, for many classes of subgraph H, the number of induced subgraphs of $\mathcal{G}(n,p)$ which are isomorphic to H is normally distributed. More recently, Krivelevich, Sudakov, and Wormald [93] studied the order of the largest induced regular subgraph of $\mathcal{G}(n,\frac{1}{2})$, showing that a.a.s. the largest such graph has roughly $n^{2/3}$ vertices. More recently again, Kamaldinov, Skorkin, and Zhukovskii [84] proved that a.a.s. the maximum size of an induced subtree of $\mathcal{G}(n,p)$ is concentrated on two consecutive values, specifically either $\lfloor 2\log_{1/(1-p)} np + 2.9 \rfloor$ or the next integer.

When it comes to graphs with a given degree sequence, Gao et al. [57] used switchings to estimate the probability that the subgraph of a uniformly random d-regular graph induced on $S \subset [n]$ is some given graph H, for $d = o(n^{1/3})$. This is also extended to the case where $\boldsymbol{d}$ is an arbitrary graphical sequence and $\Delta(\boldsymbol{d}) = o(M(\boldsymbol{d})^{1/4})$. On the denser side of things, McKay [108] asymptotically determined the analogous probability for the case where the average degree is very high (specifically $\Omega(n/\log n)$) using the multidimensional saddle point method. More recently, Gao and Ohapkin [56] gave an asymptotic formula for the probability that a given graph H is a subgraph (induced or otherwise) of $\mathcal{G}(\boldsymbol{d})$, with much less restrictive conditions on $\boldsymbol{d}$ and H (see Theorem 4 and Corollary 5 of their paper).

The results presented in this thesis approach this problem from a different perspective. Rather than asking for the probability that the induced subgraph is equal to some particular graph, we study the properties of induced subgraph without explicitly determining the exact graph itself. This allows us to get quite different results than the aforementioned papers, trading specific knowledge about the subgraph for more information about its overall structure. For example, it would be extremely non-trivial to determine the threshold for the existence of a giant component in the induced subgraph using the results of Gao et al. [57] or Gao and Ohapkin [56].

2.4.1 Percolation

Another well-studied method for generating random graphs is to take some graph (which can be fixed or random) and delete edges or vertices of the graph according to some probability distribution. Classically, edges (equivalently, vertices) are kept independently with probability p for some $p \in (0,1)$, and this is known in the literature as *bond percolation* (equivalently, *site*

percolation). As an extremely basic example, the random graph model $\mathcal{G}(n,p)$ is equivalent to the graph K_n after bond percolation where edges are kept with probability p. The study of percolation on graphs has received particular attention in the cases of large, infinite graphs, such as the infinite grid graph with vertex set $\mathbb{Z}^2$, the n-dimensional hypercube, or other lattice-like graphs. This area of study was started by Broadbent and Hammersley [26] using $\mathbb{Z}^2$ after bond percolation as a model for fluids moving through porous media. Since then, similar ideas have been applied to many problems such as traffic flow in cities [96], fragmentation of virus shells [27], and the deposition of metals on surfaces [81], as well as creating a robust mathematical field in its own right. The most commonly asked questions are about the connectivity structure of the remaining graph: after randomly deleting edges or vertices, what is the probability that the remaining graph is connected or contains a giant component?

There has been considerable study of percolation on the n-cube, the graph representation of the n-dimensional hypercube. Erdős and Spencer [44], and independently Burtin [28], studied the component structure of the n-dimensional hypercube subject to bond percolation, where edges are kept with probability p. They showed that the percolated graph is a.a.s. connected when $p > \frac{1}{2}$, a.a.s. disconnected when $p < \frac{1}{2}$, and connected with a probability that tends to e^{-1} for $p = \frac{1}{2}$. This is somewhat reminiscent of the connectivity threshold of $\mathcal{G}(n,p)$, where the probability that $\mathcal{G}(n,p)$ is connected converges to $\exp(\exp(-c))$ when $p = (\ln n + c)/n$ for constant c. Erdős and Spencer conjectured that the percolated hypercube a.a.s. had a giant component when $p = \frac{1+\varepsilon}{n}$ for some $\varepsilon > 0$, and a.a.s. had no giant component when $\varepsilon < 0$. This was proved by Ajtai et al. [2] in the case where ε is fixed, and Bollobás et al. [20] in the case where $\varepsilon \to 0$. This work has been elaborated on extensively and actively since then [22, 72, 73, 107].

One offshoot of this has been the study of percolated random graphs, or the study of random subgraphs of random graphs. Frieze et al. [51] studied the existence of giant components in bond-percolated quasi-random d-regular graphs (that is, d-regular graphs where all but the first eigenvalue are $o(d)$) when $d \to \infty$. They found that there exists a critical threshold of $p_c := 1/d$: if $p > (1+\varepsilon)p_c$ for some $\varepsilon > 0$, then the percolated graph a.a.s. contains a unique component of size $\Omega(n)$, but if $p < (1-\varepsilon)p_c$ then the graph has maximum component size $O(\log n)$. This idea was elaborated on by Bollobás et al. [19], who showed that for arbitrary dense graph sequences $\{G_n\}_{n\geq 1}$ subject to bond percolation, the threshold for the existence of a giant component is $\frac{1}{\lambda_n}$, where λ_n is the largest eigenvalue of G_n. This aligns with the result of Frieze et al. [51], as d is the largest eigenvalue of every d-regular graph. Alon et al. [3] studied bond percolation for finite d-regular expander graphs of high girth and bounded degree. They found that above a critical value $p_c = \frac{1}{d-1}$, a unique giant component appears a.a.s. in the bond-percolated expander graph. They also use this to show that for a random d-regular graph, the critical percolation threshold is almost surely $\frac{1}{d-1} + o(1)$. This agrees with a slightly earlier result of Goerdt [59], who proved that $\frac{1}{d-1}$ is the bond percolation threshold for almost all d-regular graphs. Pittel [120] then showed that the transition window around p_c is of order $n^{-1/3}$ in the case where d is constant. This work was expanded on by Nachmias and Peres [114], and then later by Joos and Perarnau [82]. The latter paper concluded that, at the critical percolation threshold, the largest component is a.a.s. of order $\Theta(n^{2/3})$ for all $d \in \{3, \ldots, n-1\}$. This is strongly reminiscent of the component structure around the threshold for the existence of giant components in $\mathcal{G}(n,M)$.

Focusing more on random graphs with given degree sequences, Fountoulakis [46] gave an early result on the the study of $\mathcal{G}(\boldsymbol{d})$ after bond and site percolation for irregular degree sequences. He

showed that, for an asymptotic sequence of degree sequences $(\boldsymbol{d}_n)_{n\geq 1}$ such that $\Delta(\boldsymbol{d}_n) \leq n^{1/9}$, the critical threshold for both site and bond percolation is given by

$$p_c := \frac{\sum_{v\in V(G)} d(v)}{\sum_{v\in V(G)} d(v)(d(v)-1)}. \tag{2.4}$$

That is, if $p < p_c$, then a uniformly random graph with degree sequence $\boldsymbol{d}$ after (site or bond) percolation has a giant component with probability $o(1)$, whereas if $p > p_c$ the percolated graph has a giant component with probability $1 - o(1)$. This turns out to be related to the study of branching processes. Consider a vertex $v \in V(G)$ and look at the distribution of the degrees of its neighbours. The maximum degree restrictions imply that the graph is locally tree-like, in the sense that the vertices at distance at most d from each vertex v a.a.s. form a tree. It also follows that the probability a given neighbour of some vertex has degree i is very close to $\frac{in_i(\boldsymbol{d})}{M}$ (recall that M is the sum of every element in $\boldsymbol{d}$). This means that the expected number of "children" of the vertex v in this tree is asymptotically p_c^{-1}. Thus, if $p < p_c$ and edges (or similarly vertices) are kept with probability p, then the expected number of children of each vertex in this tree is $p \cdot p_c^{-1} < 1$, whereas if $p > p_c$ this expectation is strictly greater than 1. Intuitively (and, as Fountoulakis shows, rigorously), this expectation determines whether the branching process that defines the structure of the component "dies out" quickly or whether the component containing v grows to linear size. This is very similar to how the expected gain/loss of half edges is an indicator for the a.a.s. existence of giant components in $\mathcal{G}(\boldsymbol{d})$ in the characterisation by Molloy and Reed [113] discussed earlier.

The proof of this is based around analysis of the configuration model, applying results analogous to Lemma 2.2.5 to carry the results into the random graph model. A key element of the proof is the following fact: if we condition on the resulting degree sequence after percolation — suppose we call this $\boldsymbol{k}$ — then the percolated pairing is distributed as a uniformly random element of $\mathcal{C}(\boldsymbol{k})$. Then by showing that a.a.s. $\boldsymbol{k}$ is "concentrated" in some sense, known results about giant components in $\mathcal{C}(\boldsymbol{d})$ are applied to prove results about the percolated case. This idea will be explored further in Chapters 3 and 4, where we employ a similar idea and study the degree sequence of induced subgraphs of random graphs. We then apply known results about $\mathcal{G}(\boldsymbol{d})$ to study a range of graph properties for these induced subgraphs, including the existence of a giant component.

The result of Fountoulakis was extended by Janson [74], who also studied the properties of a uniformly random pairing (and by Lemma 2.2.5, equivalently a uniformly random graph) after site and bond percolation. The treatment of the deleted vertices or edges in this method is slightly different to that of Fountoulakis, and Janson describes the deletion process in two steps. First, to-be-deleted vertices (or half edges) are "exploded": an exploded degree i vertex is replaced with i degree 1 vertices (each containing a unique half edge from that vertex), and an exploded half edge is moved into its own degree 1 vertex. Secondly, all these newly created degree 1 vertices (and the half edges contained therein) are deleted. However, since these newly created vertices are indistinguishable from normal degree 1 vertices, it is sufficient to delete the corresponding number of degree 1 vertices uniformly at random from all degree 1 vertices in the "exploded" graph. In this way, the problem again boils down to studying a modified degree sequence, albeit with the extra caveat of also deleting a specific number of degree 1 vertices. In particular, for

site percolation, Janson shows that, under sufficiently nice smoothness conditions for the degree sequence, the site percolated pairing a.a.s. has a giant component if and only if

$$\sum_{j=0}^{\infty} j(j-1)\pi_j p_j > \sum_{j=0}^{\infty} j p_j,$$

where p_j is the asymptotic fraction of vertices with degree j and π_j is the probability that each degree j vertex is deleted. Note that in the case where π_j is fixed and equal for all j, this aligns with earlier result of Fountoulakis [46] as the threshold aligns with (2.4). In the case where a giant component exists, he also determines its order up to an error of $o(n)$. This is done by analysing a Galton-Watson style branching process. These results are then related back to the random graph model by Lemma 2.2.5. This idea of "exploding" deleted vertices will be useful in Chapter 6, where a similar idea is used without the need for regularity conditions on the degree sequence $\boldsymbol{d}$. By combining this idea with a modified version of the exploration process used by Joos et al. [83], we determine the threshold for the existence of a giant component in the induced sub-pairing on a fixed set S for all possible degree sequences $\boldsymbol{d}$, provided that S is sufficiently large in some sense (specifically that $\sum_{i\in S} d(i) \sim cM(\boldsymbol{d})$ for some constant $c \in (0,1)$).

Recently, Fountoulakis et al. [47] studied the existence of giant components in edge percolated random graphs. This paper is notably different from the aforementioned results of Fountoulakis and Janson in that it does not use the configuration model. Instead, they use the switching method and apply the more recent result of Joos et al. [83] to give conditions that guarantee the a.a.s. existence of giant components after edge percolation in $\mathcal{G}(\boldsymbol{d})$ when $\boldsymbol{d}$ has bounded average degree. Their result shows that the critical condition depends on the tail of the degree sequence. If the tail is small, then there exists a percolation threshold, but if the tail is sufficiently large, then for all constants $p \in (0,1]$ the bond-percolated graph a.a.s. contains a giant component. In Chapter 5, we study site percolation on $\mathcal{G}(\boldsymbol{d})$. We do not require that our graphs have bounded average degree. Instead, our results require a condition on the maximum degree (the same condition as imposed in Chapters 3 and 4) which allows our results to apply in cases when the graph G is more dense than bounded average degree allows for. A consequence of our results, which is intuitively similar to results of Fountoulakis et al. [47], is that if $\boldsymbol{d}$ has total degree $\omega(n)$ then site-percolated graph a.a.s. has a giant component when p is constant.

Chapter 3

Degree sequence of the induced subgraph

In this chapter we analyse the distribution of the degree sequence of the induced graph $G[S]$, where $G \sim \mathcal{G}(\boldsymbol{d})$ and S is a fixed subset of the vertex set of G. We show that this degree sequence is a.a.s. "close" in some sense to an "average" sequence that can be defined deterministically in terms of $\boldsymbol{d}$ and S. We then use this to show that many basic quantities defined from the degree sequence of the induced subgraph are concentrated around their corresponding values for the average sequence. This forms the framework of what we call the reduction method in Chapter 1.

3.1 Overview

In this section we introduce the concepts on which we focus our study, as well as much of the notation. Let $\boldsymbol{d} = (d(1), \dots, d(n))$ be a graphical sequence, that is, the degree sequence of a graph with vertex set $[n]$. Let $\Omega_{n,\boldsymbol{d}}$ be the set of all graphs with degree sequence $\boldsymbol{d}$. Let $\mathcal{G}(\boldsymbol{d}) := (\Omega_{n,\boldsymbol{d}}, \mathcal{P}(\Omega_{n,\boldsymbol{d}}), \mathbb{P})$ be the probability space of all graphs with degree sequence $\boldsymbol{d}$ equipped with the uniform probability measure. Let G be a graph from $\mathcal{G}(\boldsymbol{d})$ sampled uniformly at random. Let $S := \{i_1, \dots, i_s\} \subset [n]$ be a set of vertices and let $\overline{S} := S \backslash [n]$. Recall that $G[S]$ is the induced subgraph of G on S. For each $v \in S$, define $d_S(v)$ to be the degree of vertex v in the induced subgraph $G[S]$, which we call its *induced degree*. Then in the probability space $\mathcal{G}(\boldsymbol{d})$, the function $d_S(v)$ is a random variable for each $v \in S$. This random variable takes integer values between 0 and $d(v)$ (including 0 and $d(v)$) counting how many edges incident to v are also incident to another vertex in S; more succinctly, $d_S(v)$ is the number of neighbours of v in S. Let $\boldsymbol{d}_S$ be the degree sequence of $G[S]$, where $d_S(j)$ is the degree of vertex i_j in $G[S]$ for $j \in [s]$. We refer to $G[S]$ as the induced (sub)graph, and the degree sequence $\boldsymbol{d}_S$ as the *induced degree sequence*.

For graphs G where the maximum degree is sufficiently low (roughly $o(\sqrt{M})$), we show that these random variables are binomially distributed up to some small error. This means that for vertices of moderate degree (in some soon-to-be-defined sense), their degrees in $G[S]$ are highly concentrated around their respective averages. Intuitively, for a given vertex $v \in S$, the probability that some edge incident to v is also incident to another vertex in S should be approximately $\frac{d(S)}{M}$, where M is the total degree of $\boldsymbol{d}$. This intuition follows from the configuration model: each half edge in the bin for vertex v is paired with a uniformly random half edge in G, and exactly $d(S)-1$ of the $M-1$ possible mates for this half edge belong to vertices in S. According to this intuition,

the expected induced degree of this vertex is approximately $\frac{d(S)}{M}d(v)$. In later lemmas we show that, under some assumptions on the maximum degree of G, this intuition is correct, and we nail down the error in these approximations more precisely.

However, one should not expect an induced subgraph of a regular graph to itself be a regular graph. Rather, the degrees of the vertices in $G[S]$ are roughly binomially distributed, where the induced degree of a vertex $v \in S$ can be approximated by the random variable $\mathrm{Bin}\,(d(v), p)$ where $p = \frac{d(S)}{M}$, where $d(S)$ is the total degree of S in G and $M := M(\boldsymbol{d})$ is the total degree of G. In this sense, the induced degree of a specific low-degree vertex is not precisely predictable, but the number of low-degree vertices with a given induced degree is concentrated. Thus, while an induced subgraph of (for example) a 5-regular graph is not regular, we can get precise estimates on how many vertices of each degree it is likely to contain.

The relevance of studying $\boldsymbol{d}_S$ comes from the following observation.

Proposition 3.1.1. Conditional on the event that $\boldsymbol{d}_S = \boldsymbol{k}$ for some arbitrary graphical sequence $\boldsymbol{k}$, the graph $G[S]$ is distributed as a uniformly random graph with degree sequence $\boldsymbol{k}$.

Proof sketch. Let H_1, H_2 be graphs with degree sequence $\boldsymbol{k}$, and let $\Omega_{n,\boldsymbol{d}}[H]$ be the subset of $\Omega_{n,\boldsymbol{d}}$ such that $G[S] = H$ for an arbitrary graph H. Then define a map $\Phi(H_1, H_2; \cdot) : \Omega_{n,\boldsymbol{d}}[H_1] \to \Omega_{n,\boldsymbol{d}}[H_2]$ that takes a graph $G \in \Omega_{n,\boldsymbol{d}}[H_1]$, deletes all the edges in $E(H_1)$, and adds in all the edges in $E(H_2)$. This map is a bijection since H_1 and H_2 have the same degree sequence, and thus $|\Omega_{n,\boldsymbol{d}}[H_1]| = |\Omega_{n,\boldsymbol{d}}[H_2]|$. Since $\mathcal{G}(\boldsymbol{d})$ is a uniform probability space, this means that

$$\mathbb{P}\left(G[S] = H_1\right) = \frac{|\Omega_{n,\boldsymbol{d}}[H_1]|}{|\Omega_{n,\boldsymbol{d}}|} = \frac{|\Omega_{n,\boldsymbol{d}}[H_2]|}{|\Omega_{n,\boldsymbol{d}}|} = \mathbb{P}\left(G[S] = H_2\right).$$

Since H_1, H_2 are arbitrary graphs with degree sequence $\boldsymbol{k}$, this means that $\mathbb{P}\left(G[S] = H\right)$ is the same for all graphs H with degree sequence $\boldsymbol{k}$. Since $\boldsymbol{k}$ is arbitrary, this holds for all possible sequences. □

This proposition means that we can study the properties of the induced graph $G[S]$ by applying known results about random graphs with given degree sequences to $\boldsymbol{d}_S$. Since $\boldsymbol{d}_S$ is a random variable, this reduces our problem to understanding the distribution of $\boldsymbol{d}_S$. This observation also underpins the study of percolated random graphs done by Fountoulakis [46].

3.1.1 Notation and definitions

Here we recall some notation from Chapter 2, as well as introduce some new notation. Let $\boldsymbol{d}$ be an arbitrary sequence of non-negative integers. If $\boldsymbol{d}$ is graphical, let $G := G(\boldsymbol{d})$ be a uniformly random graph with degree sequence $\boldsymbol{d}$. Define $\boldsymbol{d}'$ to be the sequence $\boldsymbol{d}$ ordered in non-decreasing order; equivalently one can think of this as $d(i) = d'(\sigma(i))$ for an appropriate permutation σ. For convenience, we define σ such that it maintains relative orderings of elements with the same value in $\boldsymbol{d}$. For example, if $\boldsymbol{d}$ is already ordered in non-decreasing order, $\boldsymbol{d}' = \boldsymbol{d}$ and σ is the identity permutation. Define $\boldsymbol{d}^*$ to be the maximal subsequence of $\boldsymbol{d}$ with all positive entries (that is, all elements equal to 0 removed). Naturally, $(\boldsymbol{d}')^* = (\boldsymbol{d}^*)'$.

Since we talk about many different degree sequences at once, often with different numbers of elements, we need to lay out some notation that accommodates this. Most of this notation is similar to standard notation for graphs and their degree sequences, but written as a function that

takes a sequence as an argument. Let $n(\boldsymbol{d})$ be the number of elements in the sequence $\boldsymbol{d}$. For each $k \in \mathbb{Z}_{\geq 0}$, let

$$n_k(\boldsymbol{d}) = |\{i \in [n(\boldsymbol{d})] \mid d(i) = k\}|,$$

Recall that $\Delta(\boldsymbol{d})$ is the largest term in $\boldsymbol{d}$, or the maximum degree of a vertex in a graph with that degree sequence (if it is graphical). For some set $A \subset [n(\boldsymbol{d})]$, recall the *total degree* of A, denoted by $d(A)$, as

$$d(A) = \sum_{i \in A} d(i).$$

Also recall that $d([n]) = M(\boldsymbol{d})$ is the total degree of $\boldsymbol{d}$. For convenience we write $M = M(\boldsymbol{d})$ (where $\boldsymbol{d}$ is the degree sequence of the graph G) and use $M(\cdot)$ when referring to other sequences. If $\boldsymbol{d}$ is a graphical sequence and G is a graph with degree sequence $\boldsymbol{d}$, then $n(\boldsymbol{d})$ the number of vertices in G, $n_k(\boldsymbol{d})$ is the number of degree k vertices in G, and $M(\boldsymbol{d}) = 2|E(G)|$.

3.1.2 The idealised degree sequence

The aim of this work is to predict properties of the induced graph $G[S]$ based on the degree sequence $\boldsymbol{d}_S$. However, the degree sequence $\boldsymbol{d}_S$ is a random variable. We define an "idealised" degree sequence $\boldsymbol{d}_I(\boldsymbol{d}, S)$, which we abbreviate to $\boldsymbol{d}_I$, which in some sense represents what an average degree sequence of $G[S]$ looks like. This sequence is based on the idea that the induced degree of a given vertex $v \in S$ can be approximated by a binomial random variable with $d(v)$ trials and success chance $\frac{d(S)}{M}$.

When defining the idealised degree sequence, we impose various restrictions on $\boldsymbol{d}$ and S to make this intuition more reflective of reality. Firstly, we require that both S and $\overline{S}$ are large, in the sense that they have large total degree. We also impose a condition on the maximum degree of $\boldsymbol{d}$. Specifically, we impose the conditions that

$$d(S) = \Theta(M),\ d(\overline{S}) = \Theta(M) \quad \text{and} \quad \Delta := \Delta(\boldsymbol{d}) \leq \frac{\sqrt{M}}{\log^7 M}.$$

These conditions apply throughout Chapters 3 and 4. The conditions on S are quite general. Since $M \in [n, \Delta n]$, this condition implies that $|S|,\ |\overline{S}| = \Omega(\sqrt{M}\log^7 M)$. In particular, this allows for the case where $|S| = o(n)$, as long as the set S has sufficiently high average degree in G. The condition on Δ implies that $\Delta^2 = o(M)$, which allows for tight error bounds on the switchings used to determine the induced degree of a given vertex. The extra factors of $\log M$ in the denominator are relevant for various union bounds used in Section 3.2.3. Notably, this condition on the maximum degree is much more general than the condition that $\Delta = o(\sqrt{n})$, and allows for much higher-degree vertices in dense graphs. For example, if $\boldsymbol{d} = (n^{1-\varepsilon}, \ldots, n^{1-\varepsilon})$, then the condition is satisfied, as $\sqrt{M}/\log^7 M = n^{1-\varepsilon/2}\log^7 M \geq \Delta$.

Here we lay the foundation to define $\boldsymbol{d}_I$ formally. Let $\boldsymbol{d} = (d(1), \ldots, d(n))$ be an n-element graphical sequence. Without loss of generality, we suppose that $\boldsymbol{d}$ is ordered in non-decreasing order. Recall that $S \subset [n]$ is defined to be the subsequence $\{i_k\}_{k \in [s]}$, so $|S| = s$. Define $J :=$

$\log M \log\log M$ for brevity and

$$\ell := \max_{k\in\{1,\dots,s\}} \{d(i_k) \le J\}.$$

Define $S_{\text{small}} = \{i_1, \dots, i_\ell\}$ and $S_{\text{big}} = \{i_{\ell+1}, \dots, i_s\}$, where $S_{\text{big}} = \emptyset$ if $\ell = s$. For this definition and throughout this chapter, let $Z_j \sim \text{Bin}\left(j, \frac{d(S)}{M}\right)$. For $k \in \{0, \dots, J\}$, define

$$\tilde{y}_k = \sum_{i\in S_{\text{small}}} \mathbb{P}\left(Z_{d(i)} = k\right). \tag{3.1}$$

Effectively, we want to define $\boldsymbol{d}_I$ (or at least the first ℓ entries of $\boldsymbol{d}_I$) to contain exactly $\tilde{y}_k$ elements equal to k in its first ℓ elements. However, since $\tilde{y}_k$ is not necessarily an integer, we employ what is often called cascade rounding to round them to integer values while maintaining their sum. Let $\lfloor x \rceil$ be the nearest integer to x (rounding up if $x - \lfloor x \rfloor = 1/2$) and let $(y_k)_{k=0}^J$ be the sequence $(\tilde{y}_k)_{k=0}^J$ after cascade rounding is applied. Then

$$y_0 = \lfloor \tilde{y}_0 \rceil \quad \text{and} \quad y_i = \left\lfloor \sum_{k=0}^{i} \tilde{y}_k \right\rceil - \sum_{k=0}^{i-1} \tilde{y}_k \quad \text{for } i \in \{1, \dots, J\}. \tag{3.2}$$

This ensures that $\sum_{k=0}^J y_k = \sum_{k=0}^J \tilde{y}_k = \ell$ and $|y_k - \tilde{y}_k| \le 1$ for all k. Now we define $\boldsymbol{d}_I$.

Definition 3.1.2. Let $(y_k)_{k=0}^J$ be the sequence $(\tilde{y}_k)_{k=0}^J$ after cascade rounding is applied, as defined in (3.2). For each $k \in \{0, \dots, J\}$, $\boldsymbol{d}_I$ contains y_k elements equal to k, ordered in non-decreasing order. In addition to this, $\boldsymbol{d}_I$ also contains the terms

$$\left\{ \left\lfloor d(i_{\ell+1}) \frac{d(S)}{M} \right\rfloor, \dots, \left\lfloor d(i_s) \frac{d(S)}{M} \right\rfloor \right\}.$$

These terms are concatenated to the end in the order listed.

The sequence $\boldsymbol{d}_I$ in some sense represents the "average" degree sequence of $G[S]$. If we approximate the induced degree of a vertex $v \in S$ by a binomial $\text{Bin}\left(d(v), \frac{d(S)}{M}\right)$, then $\frac{d(S)}{M} d(i)$ is simply the expected degree of vertex i in $G[S]$ under this approximation. Similarly, $\tilde{y}_k$ is the expected number of vertices in S_{small} with induced degree k under this approximation, and y_k is simply this value rounded to an integer in such a way that the sum of the sequence is still $|S_{\text{small}}|$. We note that the summation in the definition of $\tilde{y}_k$ can be re-expressed as

$$\sum_{v\in S_{\text{small}}} \mathbb{P}\left(Z_{d(v)} = k\right) = \sum_{j\le J} |S_j| \mathbb{P}\left(Z_j = k\right) \tag{3.3}$$

where $S_j := \{i_k \in S : d(i_k) = j\}$ for all $j \le J$. These sums are equal as every $v \in S_{\text{small}}$ contributes to exactly one S_j, precisely the set where $d(v) = j$. In this way, the form on the right hand side is simply grouping summation terms by their value in $\boldsymbol{d}$. These two forms are used interchangeably throughout many proofs. It is worth noting that the definition for $\boldsymbol{d}_I$ is not the only possible definition for a sequence that can be used to predict properties of $G[S]$. At the end of this section we give another definition for such a sequence, and discuss the differences and similarities to $\boldsymbol{d}_I$.

The choice of $\log M \log\log M$ as the cutoff between the vertices that are concentrated indi-

vidually or as a group is not the most aesthetically pleasing choice. The cutoff function must be $\omega(\log M)$, in order to perform certain union bounds in Corollary 3.2.3. However, the more obvious choice of $\log^2 M$ requires $\Delta(\boldsymbol{d})$ to be lowered by an extra factor or two of $\log M$. For this reason we use the slightly less pleasing function, and abbreviate it to J for notational convenience.

The reason for using cascade rounding rather than a more simple method (e.g. rounding each $\tilde{y}_k$ in the conventional way) is that we want to maintain that $n(\boldsymbol{d}_I) = |S|$. For a pathological example of a pair $(\boldsymbol{d}, S)$ for which more simple rounding causes problems, consider the case where the degree sequence $\boldsymbol{d}$ and set S are given by

$$\boldsymbol{d} = (1, 1, \dots, 1, \lfloor J \rfloor) \quad \text{and} \quad S = \left\{ \frac{1}{2}n + 1, \dots, n \right\}.$$

Due to the value of the n^{th} element, it is not hard to show (formally speaking, a consequence of Lemma 3.2.5) that $\tilde{y}_k \in \left(0, \frac{1}{2}\right)$ for all $k > 1$. If we define y_k using conventional nearest-integer rounding, the sequence $\boldsymbol{d}_I$ would contain no term greater than 1. This might be unsettling, as one might (accurately) expect that vertex n in $G[S]$ is likely to have induced degree significantly greater than 1. On the other hand, if we always rounded up, the sequence $\boldsymbol{d}_I$ would contain almost J terms greater than 1, despite $\boldsymbol{d}$ only containing a single term greater than 1. In either case, using these other methods of rounding makes it possible (if not highly likely) that $n(\boldsymbol{d}_I) \neq |S|$, and instead we might only be able to say that $|n(\boldsymbol{d}_I) - |S|| \leq J$. From a predictive perspective this might not be an issue — in the grand scheme of things, $\log M \log\log M$ is quite small, so the addition or removal of a small number of terms of degree at most $\log M \log\log M$ does not affect many of the predictive properties of $\boldsymbol{d}_I$ that we prove later. However, it significantly simplifies the notation and proof arguments to impose that $\boldsymbol{d}_S$ and $\boldsymbol{d}_I$ have the same number of terms, and in many ways is simply more natural.

It is worth noting that the definition of $\boldsymbol{d}_I$ does not explicitly define an induced degree for each individual vertex. We avoid referring to vertices in the definition of $\boldsymbol{d}_I$ as there is no clear notion of identifying terms in S_{small} with a particular "idealised degree" in $\boldsymbol{d}_I$. Instead, $\boldsymbol{d}_I$ gives a more holistic view of the average behaviour of vertices in $G[S]$. None of the theorems that we apply to study $G[S]$ (the theorems mentioned in Chapter 2) depend on specific vertices in $G[S]$ having specific degrees. Thus, we never need the degree of a fixed low-degree vertex to be well concentrated.

This sequence is also not necessarily ordered, for example it is possible that $d_I(\ell) > d_I(\ell+1)$. As a small example of this, consider the pair $(\boldsymbol{d}, S)$ where $\boldsymbol{d} = (J, \dots, J, J+1, J+1)$ and $S = \{\frac{1}{2}n, \dots, n-1\}$. Then $d_I(s) \leq \lfloor \frac{1}{2}J \rfloor + 1$, but it is straightforward to show from the definition of $\boldsymbol{d}_I$ and y_k that $d_I(s-1) > \frac{1}{2}J + 1$. For times when it is convenient to talk about $\boldsymbol{d}_I$ being ordered, recall from Section 3.1.1 that $\boldsymbol{d}'_I$ is the sequence $\boldsymbol{d}_I$ ordered in non-decreasing order. This can be written as $d'_I(i) = d_I(\sigma(i))$ for an appropriate permutation σ. As we discuss later, the sequence $\boldsymbol{d}_I$ is "almost" ordered in the sense that any terms that are out-of-order all have degree $\frac{d(S)}{M}J(1 + o(1))$.

One point that might cause alarm is that the sequence $\boldsymbol{d}_I$ is not necessarily a graphical sequence. For instance, it is entirely possible that $M(\boldsymbol{d}_I)$ is odd. For our methods it is actually entirely unnecessary that $\boldsymbol{d}_I$ be a graphical sequence in its own right. However, one can force that $\boldsymbol{d}_I$ is always graphical with minimal modification: if $M(\boldsymbol{d}_I)$ is odd, then lower the value of the highest-degree term in $\boldsymbol{d}_I$ by 1 to create a new sequence $\boldsymbol{d}_I^f$. Otherwise, if $M(\boldsymbol{d}_I)$ is even, define

$\boldsymbol{d}_I^f = \boldsymbol{d}_I$. The next lemma and subsequent corollary state that $\boldsymbol{d}_I^f$ is always a graphical sequence. This modification does not meaningfully affect the properties of $\boldsymbol{d}_I$ that we consider, and allows us to talk about random graphs with degree sequence $\boldsymbol{d}_I$ without having to add the caveat that $\boldsymbol{d}_I$ is graphical. Thus, for intuition purposes one may simply assume that $\boldsymbol{d}_I$ is graphical.

Lemma 3.1.3. Let $\boldsymbol{d}$ be an n-element sequence of non-negative integers, for n sufficiently large. If $M(\boldsymbol{d})$ is even and $\Delta(\boldsymbol{d}) \leq c\sqrt{M}$ for some $c \leq \frac{1}{2}$, then $\boldsymbol{d}$ is a graphical sequence.

Proof. Without loss of generality, we can assume that $\boldsymbol{d}$ is ordered in non-decreasing order. Recall that the Erdős-Gallai theorem (given in Theorem 2.2.6) states that $\boldsymbol{d}$ is a graphical sequence if and only if

$$\sum_{i=0}^{k-1} d(n-i) \leq k(k-1) + \sum_{i=1}^{n-k} \min\{d(i), k\} \tag{3.4}$$

for all $k \leq n$ (since we assume that $M(\boldsymbol{d})$ is even). Note that because of the maximum degree constraints on $\boldsymbol{d}$, the left hand side of this inequality is at most $kc\sqrt{M}$.

First suppose that $k \geq c\sqrt{M}+1$. Then $k(k-1) \geq kc\sqrt{M}$, and the claim immediately follows. Now suppose that $k < c\sqrt{M}+1$. Then define $x_k := \min\{x \in [n] : d(x) \geq k\}$. It follows from this definition that

$$\sum_{i=1}^{n-k} \min\{d(i), k\} = \sum_{i=1}^{x_k} d(i) + \sum_{i=x_k+1}^{n-k} k.$$

Now consider two cases. First, consider the case that $n-k-x_k \geq c\sqrt{M}$. Then Theorem 2.2.6 holds immediately, as

$$\sum_{i=x_k+1}^{n-k} k \geq kc\sqrt{M}.$$

On the other hand, if $n-k-x_k < c\sqrt{M}$, then it follows that

$$\sum_{i=x_k+1}^{n-k} d(i) < c^2 M,$$

since $\boldsymbol{d}$ has maximum degree at most $c\sqrt{M}$. Thus,

$$\sum_{i=1}^{x_k} d(i) = M(\boldsymbol{d}) - \sum_{i=x_k+1}^{n-k} d(i) - \sum_{i=n-k+1}^{n} d(i) > M - 2c^2M - c\sqrt{M} = M(1-2c-o(1)).$$

Since $k < c\sqrt{M}+1$, it follows that the left hand side of (3.4) is less than $c^2M + c\sqrt{M}$. Since $c \leq \frac{1}{2}$, it then follows that

$$\sum_{i=0}^{k-1} d(n-i) \leq M(c^2+o(1)) < M(1-2c^2-o(1)) < \sum_{i=1}^{n-k} \min\{d(i), k\}.$$

Thus, $\boldsymbol{d}$ satisfies (3.4) for all $k \leq n$. Therefore, since we assume $M(\boldsymbol{d})$ is even, it follows that $\boldsymbol{d}$ is a graphical sequence. □

With this in mind, it is easy to argue that $\boldsymbol{d}_I^f$ is always a graphical sequence. We defer the proof for now, since it relies on a result about $M(\boldsymbol{d}_I)$ (specifically Lemma 3.3.1) that we prove later.

Corollary 3.1.4. The sequence $\boldsymbol{d}_I^f$ is a graphical sequence.

To relate this idealised degree sequence $\boldsymbol{d}_I$ to a random degree sequence $\boldsymbol{d}_S$, we use switchings to show that the degrees of vertices in $G[S]$ a.a.s. behave as expected. We show that a.a.s. all vertices $v \in S$ such that $d(v) > J$ (that is, the vertices in S_{big}) have an induced degree asymptotically equal to $d(v)\frac{d(S)}{M}$. Furthermore, we show that the number of vertices $v \in S_{\text{small}}$ with induced degree k is concentrated around its average value for all k. These ideas are formalised in the following lemma. For ease of writing, we define $\gamma := d(S)/M$; it immediately follows that $\gamma = \Theta(1)$. This definition will be used in many places throughout this chapter. We also use the notation that $a = b \pm c$ means that $a \in [b - c, b + c]$.

Lemma 3.1.5. Suppose that $\boldsymbol{d}$ is an n-element graphical sequence with maximum degree at most $\sqrt{M}/\log^7 M$, and suppose $S \subset [n]$ such that $d(S) = \gamma M$ and $d(\overline{S}) = (1-\gamma)M = \Theta(M)$. Then the following statements hold with probability $1 - o(1)$.

(a) For all vertices $v \in S$ such that $d(v) > J$,

$$d_S(v) = d(v)\frac{d(S)}{M}\left(1 \pm \frac{10}{\sqrt{\gamma \log\log M}}\right).$$

(b) Let Y_i be the number of vertices in S_{small} with induced degree i. Then

$$\mathbb{E}[Y_i] = \tilde{y}_i\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right)$$

and

$$|Y_i - \mathbb{E}[Y_i]| \leq \frac{1}{\log^5 M}\mathbb{E}[Y_i] + \log^7 M$$

for all $i \leq J$.

This is the key lemma we use to analyse the properties of $G[S]$. Many results in the study of random graphs with given degree sequences state that $G(\boldsymbol{d})$, a uniformly random graph with degree sequence $\boldsymbol{d}$, a.a.s. has a particular property if $\boldsymbol{d}$ satisfies some conditions. This lemma implies that $\boldsymbol{d}_S$ and $\boldsymbol{d}_I$ are very similar degree sequences. Thus, if $\boldsymbol{d}_I$ satisfies a particular condition, then depending on the specifics we can apply Lemma 3.1.5 to show that a.a.s. $\boldsymbol{d}_S$ satisfies the same condition or a very similar one. This allows us to determine the properties of $G[S]$ simply by knowing $\boldsymbol{d}$ and S. Using this method, we analyse the connectivity and the chromatic number $G[S]$, as well as the existence of non-trivial automorphisms or giant components in $G[S]$. These results and their proofs are given in Chapter 4.

One thing that follows from the definition of $\boldsymbol{d}_I$ is that $n_k(\boldsymbol{d}_I)$ and $n_{k+1}(\boldsymbol{d}_I)$ are closely related for low values of k. That is, if $\boldsymbol{d}_I$ has a very large number of elements equal to k with indexes in $\{1, \ldots, \ell\}$ (that is, in S_{small}), then it also has a relatively large number of elements equal to $k-1$. This is formalised in the following remark. Combining the following remark with Lemma 3.1.5

means, for example, that if $G[S]$ a.a.s. contains many degree 2 vertices, then a.a.s. it also contains a large number of degree 1 vertices. This is particularly useful later when proving properties of $G[S]$ in Chapter 4.

Remark 3.1.6. Recall that S_j is the set of $i \in S$ such that $d(i) = j$. Then for $k \geq 1$,

$$\sum_{j \leq J} |S_j| \mathbb{P}\left(Z_j = k\right) = \sum_{j \leq J} \frac{d(S)}{d(\overline{S})} \frac{j-k+1}{k} |S_j| \mathbb{P}\left(Z_j = k-1\right).$$

With some naive bounds on the value of $\frac{j-k+1}{k}$, this gives useful bounds on the ratios between successive values of y_k. Since $j \leq J$, it follows that

$$\tilde{y}_k = \sum_{j \leq J} |S_j| \mathbb{P}\left(Z_j = k\right) \leq \frac{J}{k} \frac{d(S)}{d(\overline{S})} \sum_{j \leq J} |S_j| \mathbb{P}\left(Z_j = k-1\right) = \frac{J d(S)}{k d(\overline{S})} \tilde{y}_{k-1},$$

and thus $y_{k-1} = \Omega(k y_k / J)$ for all $k \geq 1$.

It is worth noting that y_k, defined in Definition 3.1.2, is not necessarily the number of terms equal to k in $\boldsymbol{d}_I$, as it is possible that $d_I(i) = k$ for some $k > \ell$. Similarly, Y_k, as defined in Lemma 3.1.5(b), is not necessarily the number of degree k vertices in $G[S]$. However, as noted in the following remark, this is only an issue for $k > \frac{d(S)}{M} \log M \log\log M(1 - o(1))$.

Remark 3.1.7. Lemma 3.1.5 implies that with probability $1 - o(1)$, no vertex in S_{big} has induced degree less than $\gamma \log M \log\log M(1 - 10/\sqrt{\gamma \log\log M})$. This means that for all k smaller than this value, we have that $n_k(\boldsymbol{d}_S) = Y_k$ with probability $1 - o(1)$. This means, for example, when considering the number of vertices of degree 0 in $G[S]$, we do not need to consider the case where a vertex in S_{big} is isolated in $G[S]$, as it happens with probability $o(1)$.

We also give a more relaxed (but more general) concentration result for the induced degree of vertices in S_{small} with larger degrees. Analogous asymptotic concentration results to Lemma 3.1.5(a) still hold with probability $1 - o(1)$ if the cutoff of $\log M \log\log M$ is replaced with $\log M \log\log\log M$, with analogous replacements in the definition of the error margin.

Remark 3.1.8. For all vertices $v \in S$ such that $d(v) \geq \log M \log\log\log M$,

$$d_S(v) \in \left[\frac{d(S)}{M} d(v) \left(1 - \frac{10}{\sqrt{\gamma \log\log\log M}}\right), \frac{d(S)}{M} d(v) \left(1 + \frac{10}{\sqrt{\gamma \log\log\log M}}\right)\right]$$

with probability $1 - o(M^{-10})$.

The proof of the claim is almost identical to the proof of Lemma 3.1.5(a), which is given in Lemma 3.2.2 and Corollary 3.2.3. Despite being even less aesthetically pleasing, this remark has a useful consequence. We know from Lemma 3.1.5(a) that a.a.s. all vertices in S_{big} have degree $\Omega(J)$ in $G[S]$, however it says nothing about the behaviour of vertices in S_{small} with degree $\Theta(J)$. This corollary says that a.a.s. every vertex in S_{small} with $d(v) = \Theta(J)$ must have an induced degree that asymptotically matches its expected induced degree.

3.1.3 An alternative predictive sequence

As mentioned earlier, $\boldsymbol{d}_I$ as defined in Definition 3.1.2 is not the only possible sequence that we can use to predict properties of the induced subgraph. Alternatively, we can define a sequence

$\boldsymbol{d}_H$, based on a cumulative distribution function.

Definition 3.1.9. Let $\boldsymbol{d}$ be a graphical sequence. Let $S \subset [n]$ define a subsequence of $\boldsymbol{d}$. To define $\boldsymbol{d}_H$, let $Z_j \sim \text{Bin}\left(j, \frac{d(S)}{M}\right)$. Then define

$$N(k) = \left\lfloor \sum_{i \in S} \mathbb{P}\left(Z_{d(i)} \leq k\right) + \frac{1}{2} \right\rceil$$

for $k \geq 0$, and $N(-1) = 0$, where $\lfloor x \rceil$ is the nearest integer to x. Then define $n_k(\boldsymbol{d}_H) = N(k) - N(k-1)$ to be the number of elements in $\boldsymbol{d}_H$ with value k.

This sequence $\boldsymbol{d}_H$ does not treat big and small vertices in S differently, which is in some sense less artificial. Furthermore, the same methods that we use to show that $\boldsymbol{d}_S$ is a.a.s. close to $\boldsymbol{d}_I$ can also be used to show that $\boldsymbol{d}_S$ and $\boldsymbol{d}_H$ are a.a.s. close. As a result, it is not hard to show that $\boldsymbol{d}_H$ can also be used to predict many properties of $G[S]$. However, while $\boldsymbol{d}_H$ might have a more natural definition, we choose to study $\boldsymbol{d}_I$ as there is less work required to show that there is an appropriate term in $\boldsymbol{d}_I$ corresponding to each high-degree vertex in S, since this is built into the definition. In this case, the slightly unnatural definition of $\boldsymbol{d}_I$ saves us some work in showing that these two sequences are similar. However, $\boldsymbol{d}_H$ also has some advantages, for example the elements are sorted in non-decreasing order by construction. Both definitions have their advantages and disadvantages, but they are so similar that any result that can be proved with one can be proved with the other, with possibly a little more work.

3.2 Concentration of the induced degree sequence

Here we prove the main workhorse lemmas that we need to prove Lemma 3.1.5. Recall that $d_S(v)$ is the degree of vertex v in $G[S]$. Recall that $M = \sum_{i \in [n]} d(i)$ and that $Z_j \sim \text{Bin}\,(j, \gamma)$ where $\gamma = d(S)/M$. We restrict ourselves to the case where $\Delta(\boldsymbol{d}) \leq \frac{\sqrt{M}}{\log^7 M}$, $\gamma = \Theta(1)$, and $1 - \gamma = \Theta(1)$. Let $\boldsymbol{d}$ be an arbitrary graphical sequence satisfying these conditions. We show that, under these conditions, the degree of a given vertex in $G[S]$ is roughly binomially distributed with some small error that shrinks to 0 as M grows. We then use this to show that a.a.s. the sequence $\boldsymbol{d}_S$ is sufficiently "close" to the sequence $\boldsymbol{d}_I$ to satisfy Lemma 3.1.5. The proof, much like the lemma, is split into two parts: one for the vertices in S above the cutoff J (that is, vertices v such that $d(v) > \log M \log\log M$) and one for vertices below the cutoff. The proofs for each part are given in Sections 3.2.2 and 3.2.3.

3.2.1 Distribution of the induced degree of a given vertex

The following lemma forms the crux of how we analyse the induced degree sequence $\boldsymbol{d}_S$. Since $\mathcal{G}(\boldsymbol{d})$ is a uniform probability space, the number of graphs with a given property is directly proportional to the probability that a graph has that property. The ratio between $\mathbb{P}\,(Z_j = i)$ and $\mathbb{P}\,(Z_j = i+1)$ is given by

$$\frac{\mathbb{P}\,(Z_j = i)}{\mathbb{P}\,(Z_j = i+1)} = \frac{i+1}{j-i}\frac{d(\overline{S})}{d(S)}.$$

With this in mind, the following lemma implies that $d_S(v)$ is, in some sense, asymptotically binomially distributed. The approximation error is a result of the switching bounds, as described in Section 2.3. However, we also would not expect that the induced degree of each vertex is exactly binomial: as mentioned earlier, $\mathcal{G}(\boldsymbol{d})$ is a complicated probability space with where the existence and non-existence of each possible edge depends on every other possible edge. Because of the error term we cannot say that $d_S(v)$ is exactly binomial, however since the distributions are similar we show in the following sections that $d_S(v)$ satisfies many concentration results similar to a binomial random variable.

Lemma 3.2.1. Let v be an arbitrary vertex in S. Let X_i be the number of graphs in $\mathcal{G}(\boldsymbol{d})$ where $d_S(v) = i$. Then

$$\frac{X_i}{X_{i+1}} = \frac{i+1}{d(v)-i} \cdot \frac{d(\overline{S})}{d(S)} \left(1 + O\left(\frac{\Delta^2}{M}\right)\right).$$

Proof. Let A_v^i be the set of $G \in \mathcal{G}(\boldsymbol{d})$ such that $d_S(v) = i$, which implies that $X_i = |A_v^i|$. We define a switching that takes a graph $G \in A_v^{i+1}$ to some $G' \in A_v^i$. Let $G \in A_v^{i+1}$. To perform a switching, choose a vertex y such that $vy \in E(G)$ and $y \in S$, as well as an ordered pair of vertices (u, x) such that $ux \in E(G)$ and $u \in \overline{S}$ (x can be in either S or $\overline{S}$). The switching deletes edges vy and ux, replacing these edges with uv and xy and hence creating a new graph G'. This switching is considered valid if and only if $G' \in A_v^i$, which occurs if and only if

(a) the vertices $\{u, v, x, y\}$ are distinct, and

(b) $xy \notin E(G)$ and $uv \notin E(G)$.

Now let $G' \in A_v^i$. Then observe that a reverse switching takes an edge uv where $u \in \overline{S}$ and an ordered pair of vertices (x, y) where $y \in S$ and $xy \in E(G)$, and replaces the edges uv and xy with edges xu and vy, creating a new graph G. Again to ensure that the result of the reverse switching is an element of A_v^{i+1}, a reverse switching is considered valid if and only if

(i) the vertices $\{u, v, x, y\}$ are distinct, and

(ii) $vy \notin E(G)$ and $ux \notin E(G)$.

A diagram for the switching is given in Figure 3.1.

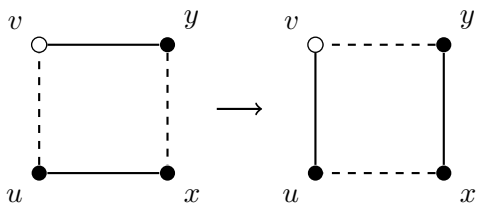

Figure 3.1: The switching used in this proof. Here $v, y \in S$ and $u \in \overline{S}$. Present edges are given as solid lines, forbidden edges are given as dashed. Other edges may be present or absent.

Now we find upper and lower bounds on the number of switchings that can be applied to some $G \in A_v^{i+1}$. The number of forward switchings is at most the number of ways of choosing an edge vy, where $y \in S$, and an ordered pair of vertices (u, x) such that $ux \in E(G)$ and $u \in \overline{S}$. Given

some $G \in A_v^{i+1}$, there are $i+1$ choices for vertex y such that $vy \in E(G)$ and $y \in S$. There are $d(\overline{S})$ choices for a vertex $u \in \overline{S}$ and neighbour x. Thus, there are at most $(i+1)d(\overline{S})$ switchings from $G \in A_v^{i+1}$ to some $G' \in A_v^i$.

To determine a lower bound on the number of switchings that can be applied to each $G \in A_v^{i+1}$, we bound from above the number of choices for $\{u, x, y\}$ as described above that do not correspond to a valid switching. This occurs exactly if at least one of (a) or (b) is not satisfied. Let W be the number of choices for $\{u, x, y\}$ such that at least one of (a) or (b) is not satisfied. Then the number of forward switchings from some graph $G \in A_v^{i+1}$ is $(i+1)d(\overline{S}) - W(G)$. We give an upper bound on W for all $G \in A_v^{i+1}$.

For case (a), we consider all the different ways that the vertices can be non-distinct. The vertices v and y must be different, as they are adjacent, as are the vertices u and x. The vertices $\{v, y\}$ must both be distinct from u, as $v, y \in S$ but $u \notin S$. Thus, the only two possibilities are if $x = v$ or $x = y$. In the first case, there are $i+1$ choices for the vertex y, and $d(v) - i - 1$ choices for the vertex u, since $ux \in E(G)$, $x = v$, and $u \in \overline{S}$. In the second case, there are $i+1$ choices for y, and at most $d(y)$ choices for u, since $ux \in E(G)$ and $x = y$. Since $d(v), d(y) \leq \Delta$, this means that there are at most $2(i+1)\Delta$ choices for $\{u, x, y\}$ that do not satisfy (a). For case (b), we first bound the number of choices such that $xy \in E(G)$. Again there are $i+1$ ways to choose y, and then given y there are at most $d(y)$ choices for x such that $xy \in E(G)$, and then at most $d(x)$ choices for u such that $ux \in E(G)$. Since $d(x), d(y) \leq \Delta$, there are at most $(i+1)\Delta^2$ choices such that $xy \in E(G)$. By similar reasoning, there are at most $(i+1)\Delta^2$ choices for $\{u, x, y\}$ such that $uv \in E(G)$. Since $\boldsymbol{d}$ has minimum degree at least 1, the number of valid switchings that can be applied to each $G \in A_v^{i+1}$ is at least $(i+1)(d(\overline{S}) - 4\Delta^2)$.

Now we use a very similar argument to count the number of reverse switchings that can be applied to each $G \in A_v^i$, or equivalently the number of switchings that create a particular $G \in A_v^i$. The number of reverse switchings is at most the number of ways of choosing an edge uv, where $u \in \overline{S}$, and an ordered pair of vertices (x, y), where $y \in S$ and $xy \in E(G)$. Given some $G \in A_v^i$, there are $(d(v) - i)$ choices for the vertex u. There are $d(S)$ choices for an ordered pair of vertices (x, y) such that $y \in S$ and $xy \in E(G)$. Thus, there are at most $(d(v) - i)d(S)$ choices for $\{u, x, y\}$ that correspond to a valid reverse switching. Now we bound from above the number of these choices that do not correspond to a valid reverse switching. A choice of $\{u, x, y\}$ does not correspond to a valid reverse switching if and only if one of (i) and (ii) are not satisfied. For case (i), we consider all the different ways that the vertices can be non-distinct. The vertices u and v are distinct, as they are adjacent, as are x and y. The vertices u and y are distinct, as $y \in S$ and $u \notin S$. Thus, the only possible clashes are if $x = v$, $x = u$, or $y = v$. Using similar reasoning to the analysis of the forward switching, there are at most $3(d(v) - i)\Delta$ choices for $\{u, x, y\}$ that do not satisfy (i). For case (ii), we consider the two possible "bad" edges separately. There are i choices for vy and $d(v) - i$ choices for u such that $y \in S$, $vy \in E(G)$, $u \in \overline{S}$, and $uv \in E(G)$. Given y, there are at most $d(y)$ choices for x such that $xy \in E(G)$. Thus, there are at most $(d(v) - i)\Delta^2$ choices for $\{u, x, y\}$ such that $vy \in E(G)$. Similarly, there are at most $(d(v) - i)\Delta^2$ choices such that $ux \in E(G)$. Thus, the number of valid backward switchings is at least $(d(v) - i)\left(d(S) - 5\Delta^2\right)$.

This implies the following inequalities.

$$\begin{aligned} X_{i+1}(i+1)(d(\overline{S}) - 4\Delta^2) &\leq X_i(d(v) - i)d(S), \\ X_{i+1}(i+1)d(\overline{S}) &\geq X_i(d(v) - i)\left(d(S) - 5\Delta^2\right). \end{aligned} \tag{3.5}$$

From these inequalities it follows that

$$\frac{X_i}{X_{i+1}} = \frac{i+1}{d(v)-i} \cdot \frac{d(\overline{S})}{d(S)} \left(1 + O\left(\frac{\Delta^2}{M}\right)\right),$$

since $d(S)$ and $d(\overline{S})$ are $\Theta(M)$. □

3.2.2 High-degree vertices: proof of Lemma 3.1.5(a)

Now we use Lemma 3.2.1 to show that vertices in S_{big} (that is, vertices $v \in S$ such that $d(v) > J = \log M \log\log M$) a.a.s. all have induced degree close to $d(v)\frac{d(S)}{M}$. This justifies the treatment of terms in S_{big} when defining $\boldsymbol{d}_I$ in Definition 3.1.2. Recall that $\gamma = d(S)/M$, and thus $\gamma = \Theta(1)$.

Lemma 3.2.2. Let $\varepsilon = 5(\sqrt{\gamma \log\log M})^{-1}$ and define $i_0 := i_0(v) = \gamma d(v)$ (not necessarily an integer). Then, for M sufficiently large,

$$\mathbb{P}\left(d_S(v) \in [i_0(1-2\varepsilon), i_0(1+2\varepsilon)]\right) < 2d(v) \exp\left(-\frac{1}{2}\varepsilon^2 i_0\right).$$

Proof. Recall that $d_S(v)$ is the degree of vertex v in $G[S]$. We first prove that the probability that $d_S(v) < i_0(1-2\varepsilon)$ is less than $d(v)\exp\left(-\frac{1}{2}\varepsilon^2 i_0\right)$. The argument for the upper bound is identical, and then the lemma follows from the union bound. Define $i_k = (1-k\varepsilon)i_0$ for all $k, \varepsilon > 0$. Recall that X_i is the number of graphs in $\mathcal{G}(\boldsymbol{d})$ such that $d_S(v) = i$. For all $i \le i_1 - 1$, Lemma 3.2.1 implies that

$$\begin{aligned}
\frac{X_i}{X_{i+1}} &= \frac{i+1}{d(v)-i}\frac{d(\overline{S})}{d(S)}\left(1 + O\left(\frac{\Delta^2}{M}\right)\right) \\
&\le \frac{i_1}{d(v)-i_1}\frac{d(\overline{S})}{d(S)}\left(1 + O\left(\frac{\Delta^2}{M}\right)\right) \\
&\le \frac{i_0 - \varepsilon i_0}{d(v)-i_0}\frac{d(\overline{S})}{d(S)}\left(1 + O\left(\frac{\Delta^2}{M}\right)\right) \\
&= (1-\varepsilon)\frac{i_0}{d(v)-i_0}\frac{d(\overline{S})}{d(S)}\left(1 + O\left(\frac{\Delta^2}{M}\right)\right).
\end{aligned}$$

By definition of i_0,

$$\frac{i_0}{d(v)-i_0} = \frac{\frac{d(S)}{M}}{1-\frac{d(S)}{M}} = \frac{d(S)}{d(\overline{S})}.$$

Thus, it follows from the above computation that, for all $i \le i_1 - 1$,

$$\frac{X_i}{X_{i+1}} \le (1-\varepsilon)\left(1 + O\left(\frac{\Delta^2}{M}\right)\right) < 1 - \frac{3}{4}\varepsilon, \tag{3.6}$$

where the second inequality holds for M sufficiently large, since $\Delta^2 = o(M)$. Thus for all $i < i_2$,

$$\begin{aligned}\frac{X_i}{X_{i_1}} &\leq \left(1 - \frac{3}{4}\varepsilon\right)^{\varepsilon i_0} \\ &= \exp\left(-\frac{3}{4}\varepsilon^2 i_0 + O\left(\varepsilon^3 i_0\right)\right) \\ &< \exp\left(-\frac{1}{2}\varepsilon^2 i_0\right),\end{aligned}$$

where the last inequality holds for M sufficiently large, since $\varepsilon \to 0$ as $M \to \infty$. Thus if $i < i_2$, it follows that $\mathbb{P}\left(d_S(v) = i\right) < \exp\left(-\frac{1}{2}\varepsilon^2 i_0\right)$. Performing a union bound over all possible induced degrees $i \leq i_2 \leq d(v)$ gives that

$$\mathbb{P}\left(d_S(v) \leq i_2\right) \leq d(v) \exp\left(-\frac{1}{2}\varepsilon^2 i_0\right).$$

An analogous calculation gives the corresponding bound for large i, which we give here for completeness. This time, let $i_k = (1 + k\varepsilon)$. For any $i \geq i_1$, it follows that

$$\begin{aligned}\frac{X_{i+1}}{X_i} &= \frac{d(v) - i}{i + 1}\frac{d(S)}{d(\overline{S})}\left(1 + O\left(\frac{\Delta^2}{M}\right)\right) \\ &\leq \frac{d(v) - i_1}{i_1}\frac{d(S)}{d(\overline{S})}\left(1 + O\left(\frac{\Delta^2}{M}\right)\right) \\ &\leq \frac{1}{1+\varepsilon}\frac{d(v) - i_0}{i_0}\frac{d(S)}{d(\overline{S})}\left(1 + O\left(\frac{\Delta^2}{M}\right)\right) \\ &\leq \left(1 - \varepsilon + O\left(\varepsilon^2\right)\right)\left(1 + O\left(\frac{\Delta^2}{M}\right)\right) \\ &< 1 - \frac{3}{4}\varepsilon.\end{aligned}$$

This statement is analogous to the inequality given in (3.6). From this point, the rest of the proof follows similarly. □

Corollary 3.2.3. Define $\varepsilon = 5(\sqrt{\gamma \log\log M})^{-1}$ as in Lemma 3.2.2. The probability that all vertices $v \in S$ such that $d(v) > J$ satisfy

$$d_S(v) \in \left[\frac{d(S)}{M}d(v)(1 - 2\varepsilon), \frac{d(S)}{M}d(v)(1 + 2\varepsilon)\right]$$

is $1 - o\left(M^{-10}\right)$.

Proof. Recall that S_{big} is the set of vertices in S such that $d(v) > J$. We apply Lemma 3.2.2 along with the union bound over all vertices in S_{big}. Lemma 3.2.2 implies that the probability that $d_S(v)$ is outside the specified range is at most $2n \exp\left(-\frac{1}{2}\varepsilon^2 i_0\right)$. Note that $i_0 \geq \gamma \log M \log\log M$ for all $v \in S_{\text{big}}$. Combining this with the union bound implies that the probability that there exists some vertex $v \in S_{\text{big}}$ such that $d_S(v)$ is outside its specified range is at most $2n^2 \exp\left(-\frac{1}{2}\varepsilon^2 \gamma \log M \log\log M\right)$. Due to the choice of ε, this is at most $2n^2 \exp\left(-\frac{25}{2}\log M\right)$, which is $n^2 \cdot o(M^{-12})$. Since $n \leq M$, the claim holds. □

3.2.3 Low-degree vertices: proof of Lemma 3.1.5(b)

As mentioned earlier, it is not expected that the induced degree of all low-degree vertices (that is, vertices in S_{small} with degree at most $J = \log M \log\log M$) is asymptotically $d(v)\frac{d(S)}{M}$. For example, if the original graph is a random d-regular graph for some constant d, it is highly, highly unlikely that $G[S]$ is regular. However, Lemma 3.2.1 implies that the distribution of $d_S(v)$ for each these vertices should be close to binomial. Here we formalise this idea. In the case where the expected number of vertices with degree k in $G[S]$ is small (that is, polylogarithmic in M), the concentration results obtained may be weak or even trivial. However, as we show in Chapter 4, these concentration windows are still sufficient to determine thresholds for various properties of $G[S]$. This justifies our treatment of low-degree vertices in the definition of $\boldsymbol{d}_I$. Recall that $\gamma = \frac{d(S)}{M}$ and that $Z_j \sim \text{Bin}\,(j,\gamma)$.

An important step to proving Lemma 3.1.5(b) is to show that the induced degrees of pairs of vertices in S_{small} are not strongly dependent. As a step towards showing this, we show that given two vertices in S_{small}, the probability that they are adjacent is very small. This is proved in the following lemma. We prove something slightly more general than this, where we also condition on the induced degree of a small (bounded) number of other vertices in S_{small}.

Lemma 3.2.4. Let $\{v_1,\dots,v_k\} \subset S_{\text{small}}$ be a set of k distinct vertices where $k = O(1)$, and let G be a uniformly random graph with degree sequence $\boldsymbol{d}$. Then

$$\mathbb{P}\left(v_1v_2 \in E(G) \middle| d_S(v_1) = i_1,\dots,d_S(v_k) = i_k\right) = O\left(\frac{J^2}{M}\right)$$

for all $i_j \le d(v_j)$ for $j \le k$, and $\mathbb{P}\left(v_1v_2 \in E(G)\right) = O\left(\frac{J^2}{M}\right)$.

Proof. First note that if one of i_1 or i_2 is equal to zero, then

$$\mathbb{P}\left(v_1v_2 \in E(G) \middle| d_S(v_1) = i_1,\dots,d_S(v_k) = i_k\right) = 0.$$

Thus we suppose that $i_1,\ i_2 > 0$.

Let A_{v_1,v_2} be the subset of $\mathcal{G}(\boldsymbol{d})$ where v_1 and v_2 are adjacent and each vertex v_j has induced degree i_j for $j \le k$. Similarly, let B_{v_1,v_2} be the subset of $\mathcal{G}(\boldsymbol{d})$ where v_1 and v_2 are not adjacent and each vertex v_j has induced degree i_j for $j \le k$. We define a switching between these subsets as follows. Suppose $G \in A_{v_1,v_2}$. To perform a switching, choose two ordered pairs of vertices in $V(G)$, (x,y) and (a,b), such that $ab,\ xy \in E(G)$, and $y,\ b \in S$. Then the switching deletes the edges $v_1v_2,\ xy,\ ab$ and replaces them with $v_1y,\ ax,\ v_2b$, creating a new graph G'. A diagram of this switching is given in Figure 3.2. This switching is considered valid if and only if $G' \in B_{v_1,v_2}$, which occurs if and only if

(a) the vertices $\{v_1,v_2,a,b,x,y\}$ are distinct, except $y = b$ is permissible,

(b) $v_1y,\ ax,\ v_2b \notin E(G)$, and

(c) the induced degrees of $v_1,\dots,v_k$ are unchanged by the switching.

Now let $G' \in B_{v_1,v_2}$. Then note that a reverse switching is equivalent to the following: take two edges v_1y and v_2b where $y,\ b \in S$, as well as an ordered pair of vertices (a,x) such that $ax \in E(G)$,

and replace $v_1y,\ ax,\ v_2b$ with $v_1v_2,\ xy,\ ab$, hence creating a new graph G. Such an operation corresponds to a valid reverse switching if and only if $G \in A_{v_1,v_2}$, which occurs if and only if

(i) the vertices $\{v_1, v_2, a, b, x, y\}$ are distinct, except $y = b$ is permissible,

(ii) $v_1v_2,\ ab,\ xy \notin E(G)$, and

(iii) the induced degrees of $v_1, \ldots, v_k$ are unchanged by the switching.

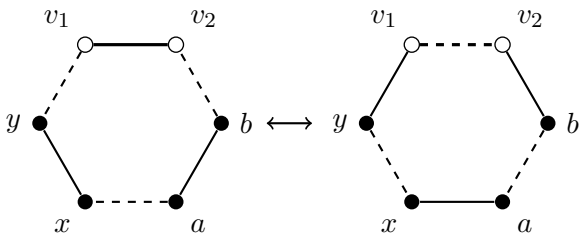

Figure 3.2: The switching used in this proof. Present edges are given as solid lines, forbidden edges are given as dashed. Other edges can be present or absent.

Now we determine upper and lower bounds on the number of switchings that can be applied to each $G \in A_{v_1,v_2}$. The number of possible forward switchings for any $G \in A_{v_1,v_2}$ is at most the number of ways of choosing (a, b) and (x, y) such that $ab,\ xy \in E(G)$, and $b, y \in S$. Since y and b are in S, there are $d(S)^2$ such choices for $\{a, b, x, y\}$ (including choices that do not correspond to valid switchings, for example where $ab = xy$). Thus, there are at most $d(S)^2$ switchings that can be applied to each $G \in A_{v_1,v_2}$.

Now let W be the number of choices for (a, b) and (x, y) as described above that do not correspond to a valid switching. This occurs exactly if at least one of (a) – (c) do not occur. We bound the number of choices such that one of the following does not occur:

(a′) the vertices $\{v_1, \ldots, v_k, a, b, x, y\}$ are distinct, except $y = b$ is permissible, and

(b′) $v_1y,\ ax,\ v_2b \notin E(G)$.

Since this implies $\{a, b, x, y\} \cap \{v_1, \ldots, v_k\} = \emptyset$, (a′) is a more strict condition than (a) and (c) combined. To bound W from above for all $G \in A_{v_1,v_2}$, we bound from above the number of choices Z such that at least one of (a′) or (b′) does not occur.

For case (a′), there are three types of ways in which the chosen vertices can be non-distinct:

(a1) $b = y$,

(a2) $a = x$, $a = y$, or $b = x$,

(a3) an element of $\{a, b, x, y\}$ is equal to an element of $\{v_1, \ldots, v_k\}$.

The definition of the switching permits $b = y$, and so we do not need to count this case. For case (a2), there are at most $d(S)$ choices for the pair (a, b), and at most $d(a)$ choices for a vertex y such that $ay \in E(G)$ (since $xy \in E(G)$ and $a = x$). Thus, the number of choices where $a = x$ at most $d(S)\Delta$. By similar arguments, the number of choices that satisfy (a2) is at most $3d(S)\Delta$. For case (a3), the number of choices for $\{a, b, x, y\}$ that intersect with $\{v_1, \ldots, v_k\}$ is at most $4d(S)(d(v_1) + \cdots + d(v_k))$.

For case (b′), there are at most $d(v_1)\Delta d(S)$ choices for $\{a, b, x, y\}$ such that $v_1 y \in E(G)$. Similarly, there are at most $d(v_2)\Delta d(S)$ choices for $\{a, b, x, y\}$ such that $v_2 b \in E(G)$. Finally, there are at most $d(S)\Delta^2$ choices for these vertices such that $ax \in E(G)$. Thus, at most $d(S)\Delta\left(i_1 + i_2 + \Delta\right)$ of the aforementioned $d(S)^2$ choices for $\{a, b, x, y\}$ do not satisfy (b). Therefore, it follows that for all $G \in A_{v_1,v_2}$,

$$W(G) \leq 4d(S)(d(v_1) + \cdots + d(v_k)) + d(S)\Delta\left(i_1 + i_2 + \Delta\right).$$

Therefore, the number of valid switchings that can be applied to each $G \in A_{v_1,v_2}$ is at least $d(S)^2 - 4d(S)(d(v_1) + \cdots + d(v_k)) - \Delta d(S)\left(i_1 + i_2 + \Delta\right)$. Since $i_j \leq d(v_j) \leq J$ for all $j \leq k$, and $k = O(1)$, this is $d(S)(d(S) - O(\Delta^2))$.

Now we determine similar bounds for the reverse switching operation. We abbreviate this argument as it is similar to previous switching bound arguments. Let $G' \in B_{v_1,v_2}$. The number of reverse switchings that can be applied to G' is at most the number of ways of choosing edges $v_1 y$, $v_2 b$, and an ordered pair of adjacent vertices (a, x) such that $b, y \in S$. By the definition of B_{v_1,v_2}, the number of choices for y is i_1, and similarly the number of choices for b is i_2. The number of choices for the adjacent pair (a, x) is at most M. Since we are only looking for an upper bound on the probability that $v_1 v_2 \in E(G)$, we do not need to consider the number of these choices that do not correspond to valid reverse switchings.

Therefore, noting that $d(S) = \Theta(M)$, it follows that

$$\frac{|A_{v_1,v_2}|}{|B_{v_1,v_2}|} \leq \frac{i_1 i_2 M}{d(S)^2}\left(1 + O\left(\frac{\Delta^2}{M}\right)\right).$$

Since $\Delta^2 = o(M)$ by assumption, the multiplicative error term is $1 + o(1)$. Thus, the probability that the vertices v_1 and v_2 are adjacent, conditional on the induced degrees of $\{v_1, \ldots, v_k\}$, is at most

$$\begin{aligned}
\mathbb{P}\left(v_1 v_2 \in E(G) \middle| d_S(v_1) = i_1, \ldots, d_S(v_k) = i_k\right) &= \frac{|A_{v_1,v_2}|}{|A_{v_1,v_2}| + |B_{v_1,v_2}|} \\
&\leq \frac{i_1 i_2 M}{d(S)^2}(1 + o(1)) \\
&\leq \frac{J^2 M}{d(S)^2}(1 + o(1)),
\end{aligned}$$

since both i_1 and i_2 are at most $d(v_1)$ and $d(v_2)$ respectively and $v_1, v_2 \in S_{\text{small}}$. Therefore, applying the law of total probability over all possible values for $(i_1, \ldots, i_k)$ implies that the probability that v_1 and v_2 are adjacent is at most

$$\sum_{(i_1,\ldots,i_k)} \mathbb{P}\left(d_S(v_1) = i_1, \ldots, d_S(v_k) = i_k\right) \frac{J^2 M}{d(S)^2}(1 + o(1)) = \frac{J^2 M}{d(S)^2}(1 + o(1)).$$

Since $d(S) = \Theta(M)$, this completes the proof. □

We now formally state and prove the lemma that the induced degrees of small sets (that is, sets of bounded size) of vertices in S_{small} are roughly independent.

Lemma 3.2.5. Suppose $(\boldsymbol{d}, S)$ is such that $d(S) = \Theta(M)$, $d(\overline{S}) = \Theta(M)$, and $\Delta(\boldsymbol{d}) \leq \frac{\sqrt{M}}{\log^7 M}$. Let

$\{v_1, \dots, v_k\} \subset S_{\text{small}}$ be a set of k distinct vertices for some $k = O(1)$. Let $Z_j \sim \text{Bin}\left(j, \frac{d(S)}{M}\right)$. Then

$$\begin{aligned}\mathbb{P}\left(d_S(v_1) = i_1, \dots, d_S(v_k) = i_k\right) &= \prod_{j=1}^{k} \mathbb{P}\left(d_S(v_j) = i_j\right)\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right) \\ &= \prod_{j=1}^{k} \mathbb{P}\left(Z_{d(v_j)} = i_j\right)\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right).\end{aligned}$$

Proof. Without loss of generality, we condition on the event that $d_S(v_j) = i_j$ for all $j \geq 2$, where $i_j \leq d(v_j)$. We show that

$$\mathbb{P}\left(d_S(v_1) = i_1 \middle| d_S(v_2) = i_2, \dots, d_S(v_k) = i_k\right) = \mathbb{P}\left(Z_{d(v_1)} = i_1\right)\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right),$$

as well as

$$\mathbb{P}\left(d_S(v_1) = i_1\right) = \mathbb{P}\left(Z_{d(v_1)} = i_1\right)\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right).$$

Then we prove the lemma quickly from these results, since v_1 is arbitrary. Let C_i be the set of graphs in $\mathcal{G}(\boldsymbol{d})$ such that $d_S(v_1) = i$ and $d_S(v_j) = i_j$ for all $j \geq 2$. That is, for all $G \in C_i$, $(d_S(v_1), d_S(v_2), \dots, d_S(v_k)) = (i, i_2, \dots, i_k)$. We apply a similar switching to the one used in Lemma 3.2.1 to switch between C_{i+1} and C_i. The diagram for this new switching is given in Figure 3.3. The important difference between this switching and the switching used in the proof of Lemma 3.2.1 is that the induced degrees of vertices $v_2, \dots, v_k$ are maintained. Other than this extra restriction, the edges are chosen in the same way as the switching used in Lemma 3.2.1.

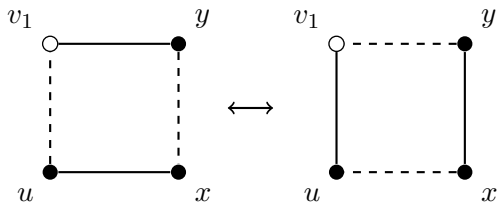

Figure 3.3: The switching used in this proof. Here $v_1, y \in S$ and $u \in \overline{S}$. Importantly, the switching does not alter $d_S(v_2), \dots, d_S(v_k)$.

Now we define the switching formally. Suppose $G \in C_{i+1}$. Choose a vertex y such that $v_1 y \in E(G)$ and $y \in S$, as well as an ordered pair of vertices (u, x) such that $ux \in E(G)$ and $u \in \overline{S}$. The switching deletes edges $v_1 y$ and ux, replacing these edges with uv_1 and xy and hence creating a new graph G'. This switching is considered valid if and only if $G' \in C_i$, which occurs if and only if

(a) the vertices $\{u, v_1, x, y\}$ are distinct,

(b) $xy \notin E(G)$ and $uv_1 \notin E(G)$,

(c) the induced degrees of $v_2, \dots, v_k$ are unchanged by the switching.

Similarly, a reverse switching applied to some $G' \in C_{i+1}$ is equivalent to taking an edge uv_1 where $u \in \overline{S}$ and an ordered pair of vertices (x, y) where $xy \in E(G)$ and $y \in S$, replacing the edges uv_1

and xy with edges xu and v_1y, thus creating a new graph G. Again to ensure that the result of the reverse switching is an element of C_{i+1}, this choice of $\{u, v_1, x, y\}$ is considered valid if and only if

(i) the vertices $\{u, v_1, x, y\}$ are distinct,

(ii) $v_1y \notin E(G')$ and $ux \notin E(G')$,

(iii) the induced degrees of $v_2, \dots, v_k$ are unchanged by the switching.

To count the number of switchings that can be applied to each $G \in C_{i+1}$, we carry out the analogous computation to what was done in the proof of Lemma 3.2.1. For each $G \in C_{i+1}$, there are $(i+1)d(\overline{S})$ choices for $\{u, v_1, x, y\}$ such that $v_1y, ux \in E(G)$, $y \in S$ and $u \notin S$. To obtain a lower bound on the number of switchings that satisfy conditions (a) – (c), we bound from above the number of these choices for $\{u, v_1, x, y\}$ such that one of the following is false:

(a) the vertices $\{u, v_1, x, y\}$ are distinct,

(b) $xy \notin E(G)$ and $uv_1 \notin E(G)$,

(c′) $\{v_2, \dots, v_k\} \cap \{u, v_1, x, y\} = \emptyset$.

Since (c′) is a more strict condition than (c), it follows that a choice for $\{u, v_1, x, y\}$ that satisfies (a), (b), and (c′) must also satisfy (a) – (c). Thus, if $W(G)$ is the number of choices for $\{u, v_1, x, y\}$ that do not satisfy at least one of (a), (b), or (c′), it follows that the number of switchings that can be applied to G is at least $(i+1)d(\overline{S}) - W(G)$. We now bound $W(G)$ from above for all $G \in C_{i+1}$. Identically to the analogous case in the proof of Lemma 3.2.1, there are at most $3(i+1)\Delta^2$ choices for $\{u, v_1, x, y\}$ that do not satisfy (a) or (b). All that remains is to give an upper bound on the number of choices that do not satisfy part (c′).

By assumption, $v_1 \neq v_j$ for $j \geq 2$, and $u \notin S$. Thus, the only possibilities for a non-empty intersection are if $x \in \{v_2, \dots, v_k\}$ or $y \in \{v_2, \dots, v_k\}$. Given $v_2, \dots, v_k$ are fixed, there are at most $k\Delta$ choices for a neighbour of one of these vertices. Thus, there are at most $(i+1)k\Delta$ choices for $\{u, v_1, x, y\}$ such that $x \in \{v_2, \dots, v_k\}$. For the case where $y \in \{v_2, \dots, v_k\}$, the number of such choices depends on which vertices in $\{v_2, \dots, v_k\}$ are adjacent to v_1 in G. If v_1 has no neighbours in the set $\{v_2, \dots, v_k\}$, then there are no choices for $\{u, v_1, x, y\}$ such that $y \in \{v_2, \dots, v_k\}$, since y is a neighbour of v_1. If $v_1v_j \in E(G)$ for some $j \geq 2$, then there are $d(\overline{S})$ choices for $\{u, v_1, x, y\}$ such that $y = v_j$, $u \in \overline{S}$, and $ux \in E(G)$. Lemma 3.2.4 implies that, for each $j \geq 2$,

$$\mathbb{P}\left(v_1v_j \in E(G) \middle| d_S(v_1) = i+1, \dots, d_S(v_k) = i_k\right) = O\left(\frac{J^2}{M}\right).$$

Since $k = O(1)$ it follows from the union bound that

$$\mathbb{P}\left(\bigcup_{j=2}^{k} \{v_1v_j \in E(G)\} \middle| d_S(v_1) = i+1, \dots, d_S(v_k) = i_k\right) = O(J^2/M).$$

Thus, the average number of choices for $\{u, v_1, x, y\}$ where $y \in \{v_2, \dots, v_k\}$, taken over all $G \in$

C_{i+1}, is

$$\sum_{i=0}^{\min\{d(v_1),k\}} id(\overline{S})\mathbb{P}\left(|N_G(v_1)\cap\{v_2,\dots,v_k\}|=i \,|\, d_S(v_1)=i+1,\dots,d_S(v_k)=i_k\right)$$
$$\leq k^2 d(\overline{S})O\left(\frac{J^2}{M}\right) = O(J^2).$$

Therefore, the average number of valid switchings that can be applied to each $G \in C_{i+1}$ is

$$(i+1)(d(\overline{S})-O(\Delta^2))-O(J^2) = (i+1)d(\overline{S})\left(1-O\left(\frac{\Delta^2+J^2}{M}\right)\right).$$

Without loss of generality, we can assume that $J \leq \Delta$ (otherwise, we could simply replace J with Δ in the previous analysis). Thus, the number of switchings that can be applied to each $G \in C_{i+1}$ is $(i+1)d(\overline{S})\left(1-O\left(\frac{\Delta^2}{M}\right)\right)$.

Now we determine upper and lower bounds for the number of reverse switchings that can be applied to each $G' \in C_i$. The enumeration of the number of choices for $\{u, v_1, x, y\}$ is very similar to enumeration done in the proof of Lemma 3.2.1: there are $(d(v_1)-i)$ choices for the vertex $u \notin S$ such that $v_1u \in E(G)$, and $d(S)$ choices for an ordered pair of adjacent vertices (x, y) such that $y \in S$. Now we bound the number of these choices that do not satisfy each of (i) – (iii). By the same reasoning as used in the proof of Lemma 3.2.1, the number of choices for these vertices that do not satisfy one of (i) or (ii) is $O((d(v_1)-i)\Delta^2)$. Now we bound the number of cases that do not satisfy (iii). Again, as an upper bound on the number of choices that do not satisfy (iii), we count the number of choices where $\{u, v_1, x, y\}$ and $\{v_2, \dots, v_k\}$ intersect.

Note that $v_1 \neq v_j$ for any $j \geq 2$ by assumption, and $u \neq v_j$ since $u \in \overline{S}$. Thus, it follows that the only choices for $\{u, v_1, x, y\}$ that do not satisfy (iii) are choices where either $x \in \{v_2, \dots, v_k\}$ or $y \in \{v_2, \dots, v_k\}$. In this case, there are at most $2(d(v_1)-i)\sum_{j=2}^{k} d(v_j)$ choices for $\{u, v_1, x, y\}$ such that either $x \in \{v_2, \dots, v_k\}$ or $y \in \{v_2, \dots, v_k\}$. Since $k = O(1)$ and $\{v_2, \dots, v_k\} \subset S_{\text{small}}$, the number of reverse switchings that can be applied to each $G' \in C_i$ is

$$(d(v_1)-i)d(S)\left(1-O\left(\frac{\Delta^2}{M}\right)\right).$$

Thus, it follows that

$$\frac{|C_i|}{|C_{i+1}|} = \frac{i+1}{d(v_1)-i}\frac{d(\overline{S})}{d(S)}\left(1+O\left(\frac{\Delta^2}{M}\right)\right). \tag{3.7}$$

Let $p_i = \mathbb{P}\left(d_S(v_1)=i \,|\, d_S(v_2)=i_2,\dots,d_S(v_k)=i_k\right)$ for $i \in \{0,\dots,d(v_1)\}$. Recall that $\gamma = d(S)/M$. Then Equation (3.7) implies that, for all $i \in \{0,\dots,d(v_1)\}$,

$$\frac{p_{i+1}}{p_i} = \frac{d(v_1)-i}{i+1}\frac{d(S)}{d(\overline{S})}\left(1+O\left(\frac{\Delta^2}{M}\right)\right) = \frac{d(v_1)-i}{i+1}\frac{\gamma}{1-\gamma}\left(1+O\left(\frac{\Delta^2}{M}\right)\right).$$

Thus, we can express p_i in terms of p_0:

$$\begin{aligned} p_i &= \binom{d(v_1)}{i}\left(\frac{1-\gamma}{\gamma}\right)^i p_0\left(1+O\left(\frac{\Delta^2}{M}\right)\right)^i \\ &= \binom{d(v_1)}{i}\left(\frac{1-\gamma}{\gamma}\right)^i p_0 \exp\left(O\left(\frac{\Delta^2 i}{M}\right)\right) \\ &= \binom{d(v_1)}{i}\left(\frac{1-\gamma}{\gamma}\right)^i p_0\left(1+O\left(\frac{\Delta^2 d(v_1)}{M}\right)\right), \end{aligned} \tag{3.8}$$

since $i \leq d(v_1)$ and $\frac{\Delta^2 d(v_1)}{M} \to 0$ when $d(v_1) \leq \log M \log\log M$. The sum of all p_i must be equal to 1, and thus

$$1 = \sum_{i=0}^{d(v_1)}\left[\binom{d(v_1)}{i}\left(\frac{\gamma}{1-\gamma}\right)^i\left(1+O\left(\frac{\Delta^2 d(v_1)}{M}\right)\right)p_0\right].$$

Since the error is uniformly bounded for all terms in the sum, and all terms are positive, the relative error of the whole sum is at most $\left(1+O\left(\frac{\Delta^2 d(v_1)}{M}\right)\right)$. Thus,

$$\sum_{i=0}^{d(v_1)}\left[\binom{d(v_1)}{i}\left(\frac{\gamma}{1-\gamma}\right)^i p_0\right] = 1+O\left(\frac{\Delta^2 d(v_1)}{M}\right).$$

It follows from the previous equation and the definition of the binomial distribution that

$$\begin{aligned} p_0 &= \left(\sum_{i=0}^{d(v_1)}\binom{d(v_1)}{i}\left(\frac{\gamma}{1-\gamma}\right)^i\right)^{-1}\left(1+O\left(\frac{\Delta^2 d(v_1)}{M}\right)\right) \\ &= (1-\gamma)^{d(v_1)}\left(1+O\left(\frac{\Delta^2 d(v_1)}{M}\right)\right). \end{aligned}$$

Applying Equation (3.8) for all $i \leq d(v_1)$, we obtain that

$$\mathbb{P}\left(d_S(v_1)=i_1 \middle| d_S(v_2)=i_2,\ldots,d_S(v_k)=i_k\right) = \mathbb{P}\left(Z_{d(v_1)}=i_1\right)\left(1+O\left(\frac{\Delta^2 d(v_1)}{M}\right)\right) \tag{3.9}$$

for each choice of $i_j \leq d(v_j)$ for all $j \leq k$. By the law of total probability, summing over all ordered tuples $(i_2,\ldots,i_k)$ such that $i_j \leq d(v_j)$ for all $j \geq 2$ gives that

$$\begin{aligned} \mathbb{P}\left(d_S(v_1)=i_1\right) &= \sum_{(i_2,\ldots,i_k)}\mathbb{P}\left(d_S(v_1)=i_1 \middle| \bigcup_{j=2}^{k}\{d_S(v_j)=i_j\}\right)\mathbb{P}\left(\bigcup_{j=2}^{k}\{d_S(v_j)=i_j\}\right) \\ &= \mathbb{P}\left(Z_{d(v_1)}=i_1\right)\left(1+O\left(\frac{\Delta^2 d(v_1)}{M}\right)\right). \end{aligned} \tag{3.10}$$

Now note that the choice of $v_1 \in \{v_1,\ldots,v_k\}$ was arbitrary, and the same argument holds for all v_j. Thus, analogous statements to Equations (3.9) and (3.10) apply for all vertices v_j. Thus,

applying these two equations for all $j \leq k$ and recalling that $k = O(1)$, it follows that

$$\begin{aligned}\mathbb{P}\left(d_S(v_1) = i_1,\ \ldots,\ d_S(v_k) = i_k\right) &= \prod_{j=1}^{k} \mathbb{P}\left(d_S(v_j) = i_j \middle| d_S(v_{j+1}) = i_{j+1}, \ldots, d_S(v_k) = i_k\right) \\ &= \prod_{j=1}^{k} \mathbb{P}\left(d_S(v_j) = i_j\right)\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right) \\ &= \prod_{j=1}^{k} \mathbb{P}\left(Z_{d(v_j)} = i_j\right)\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right).\end{aligned}$$

This completes the proof. □

Recall that Y_i is the number of vertices in S_{small} with induced degree i, for $i \in \{0, 1, \ldots, J\}$. Then the following corollary provides part of the proof to Lemma 3.1.5(b).

Corollary 3.2.6. For all $i \leq J$, $\mathbb{E}[Y_i] = \tilde{y}_i\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right)$.

Proof. By Lemma 3.2.5,

$$\mathbb{E}[Y_i] = \sum_{v \in S_{\text{small}}} \mathbb{P}\left(d_S(v) = i\right) = \sum_{v \in S_{\text{small}}} \mathbb{P}\left(Z_{d(v)} = i\right)\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right) = \tilde{y}_i\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right),$$

where the last inequality follows from the fact that the error term is uniformly bounded for all summands and that all summands are positive. □

Now we prove the rest of Lemma 3.1.5(b). We know from Lemma 3.2.5 and subsequently Corollary 3.2.6 that $\mathbb{E}[Y_i]$ is close to y_i for all $i \leq J$. The next step is to show that Y_i is concentrated around its expected value. We bound the variance of Y_i in order to later apply Chebyshev's inequality (given in Theorem A.2).

Lemma 3.2.7. For all $i \leq J$, $\text{Var}(Y_i) \leq \mathbb{E}[Y_i]\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\mathbb{E}[Y_i]\right)$.

Proof. Let V_k be the indicator variable for the event that vertex $k \in S_{\text{small}}$ has induced degree i. Then

$$Y_i = \sum_{k \in S_{\text{small}}} V_k.$$

Taking the variance of this sum gives that

$$\text{Var}(Y_i) = \text{Var}\left(\sum_{k \in S_{\text{small}}} V_k\right) = \sum_{k \in S_{\text{small}}} \text{Var}(V_k) + \sum_{j \neq k} \text{Cov}(V_j, V_k), \tag{3.11}$$

where the last sum is over all ordered pairs $(j, k) \in S_{\text{small}}^2$ where $j \neq k$. Each V_k is a Bernoulli

random variable, and thus

$$\begin{aligned}\operatorname{Var}(V_k) &= \mathbb{E}\left[V_k^2\right] - \mathbb{E}[V_k]^2 \\ &= \mathbb{E}[V_k] - \mathbb{E}[V_k]^2 \\ &= \mathbb{E}[V_k](1 - \mathbb{E}[V_k]) \\ &= \mathbb{P}(d_S(k) = i)(1 - \mathbb{P}(d_S(k) = i)).\end{aligned}$$

So the first summation is equal to

$$\sum_{k \in S_{\text{small}}} \operatorname{Var}(V_k) = \sum_{k \in S_{\text{small}}} \mathbb{P}(d_S(k) = i)(1 - \mathbb{P}(d_S(k) = i)) \leq \mathbb{E}[Y_i]. \tag{3.12}$$

An application of Lemma 3.2.5 provides a bound on the covariance terms in the second summation:

$$\begin{aligned}\operatorname{Cov}(V_j, V_k) &= \mathbb{P}(d_S(j) = i,\ d_S(k) = i) - \mathbb{P}(d_S(j) = i)\,\mathbb{P}(d_S(k) = i) \\ &= \mathbb{P}(d_S(j) = i)\,\mathbb{P}(d_S(k) = i)\, O\left(\frac{\Delta^2 J}{M}\right).\end{aligned}$$

Thus, the summation in Equation (3.11) of the covariances over all ordered pairs $(j,k) \in S_{\text{small}}^2$ where $j \neq k$ is equal to

$$\begin{aligned}&\left(\sum_{j \in S_{\text{small}}} \sum_{k \in S_{\text{small}}} \mathbb{P}(d_S(j) = i)\,\mathbb{P}(d_S(k) = i) - \sum_{j \in S_{\text{small}}} \mathbb{P}(d_S(j) = i)^2\right) O\left(\frac{\Delta^2 J}{M}\right) \\ &= \left(\left(\sum_{k \in S_{\text{small}}} \mathbb{P}(d_S(k) = i)\right)^2 - \sum_{j \in S_{\text{small}}} \mathbb{P}(d_S(j) = i)^2\right) O\left(\frac{\Delta^2 J}{M}\right).\end{aligned}$$

Therefore, noting that $\sum_{j \in S_{\text{small}}} \mathbb{P}(d_S(j) = i)^2 \in [0, \mathbb{E}[Y_i]^2]$, we obtain

$$\sum_{j \neq k} \operatorname{Cov}(V_j, V_k) \leq \mathbb{E}[Y_i]^2 \cdot O\left(\frac{\Delta^2 J}{M}\right). \tag{3.13}$$

Combining (3.12) – (3.13), we obtain an upper bound on the variance of Y_i:

$$\operatorname{Var}(Y_i) \leq \mathbb{E}[Y_i] + O\left(\frac{\Delta^2 J}{M}\right)\mathbb{E}[Y_i]^2.$$

The claim of the lemma immediately follows. □

Some straightforward applications of Chebyshev's inequality give the concentration bounds in Lemma 3.2.8. When $\mathbb{E}[Y_i]$ is not too small (in a sense defined shortly), we show that Y_i is a.a.s. within a $1 + o(1)$ window around its expected value. If Y_i is expected to be small, then it follows that a.a.s. Y_i is also small. We choose our bounds in such a way that these inequalities a.a.s. all hold simultaneously.

Lemma 3.2.8. If $\mathbb{E}[Y_i] \geq \log^{12} M$, then

$$\mathbb{P}\left(|Y_i - \mathbb{E}[Y_i]| \geq \frac{\mathbb{E}[Y_i]}{\log^5 M}\right) = O\left(\frac{1}{\log^2 M}\right).$$

If $\mathbb{E}[Y_i] \leq \log^{12} M$, then

$$\mathbb{P}\left(|Y_i - \mathbb{E}[Y_i]| \geq \log^k M\right) = O\left(\log^{12-2k} M\right).$$

Proof. First suppose that $\mathbb{E}[Y_i] \geq \log^{12} M$. Applying Chebyshev's inequality with $t := \alpha\mathbb{E}[Y_i]$ gives that

$$\begin{aligned}\mathbb{P}(|Y_i - \mathbb{E}[Y_i]| \geq \alpha\mathbb{E}[Y_i]) &\leq \frac{\operatorname{Var}(Y_i)}{\alpha^2\mathbb{E}[Y_i]^2} \\ &\leq \frac{1}{\alpha^2\mathbb{E}[Y_i]} + O\left(\frac{\Delta^2 J}{\alpha^2 M}\right).\end{aligned}$$

Recall that $\Delta = \Delta(\boldsymbol{d}) \leq \sqrt{M}/\log^7 M$. Now let $\alpha = \log^{-5} M$, then

$$\begin{aligned}\frac{\operatorname{Var}(Y_i)}{\alpha^2\mathbb{E}[Y_i]^2} &= \frac{1}{\alpha^2\mathbb{E}[Y_i]} + O\left(\frac{\Delta^2 J}{\alpha^2 M}\right) \\ &\leq \frac{1}{\log^2 M} + O\left(\frac{J}{\log^4 M}\right).\end{aligned}$$

Since $J = o(\log^2 M)$, this gives the first part of the lemma. Now suppose that $\mathbb{E}[Y_i] \leq \log^{12} M$. Then it follows that

$$\begin{aligned}\operatorname{Var}(Y_i) &= \mathbb{E}[Y_i] + \mathbb{E}[Y_i]^2 O\left(\frac{\Delta^2 J}{M}\right) \\ &\leq \log^{12} M + o\left(\log^{12} M\right) \\ &= \log^{12} M(1 + o(1)).\end{aligned}$$

Then applying Chebyshev's inequality with $t := \log^k M$ gives the desired result. □

The combination of these two concentration bounds and the union bound for all $i \leq J = \log M \log\log M$ gives the following corollary.

Corollary 3.2.9. With probability $1 - o(1)$,

$$|Y_i - \mathbb{E}[Y_i]| \leq \frac{1}{\log^5 M}\mathbb{E}[Y_i] + \log^7 M$$

for all $i \in \{0, 1, \ldots, J\}$.

Proof. Set $k = 7$ in Lemma 3.2.8. Since $J = o(\log^2 M)$, with probability $1 - o(1)$ the inequalities mentioned in the previous lemma a.a.s. simultaneously hold for all $i \leq J$. □

With this, we have proved all the claims in Lemma 3.1.5. We pull them all together here for completeness.

Proof of Lemma 3.1.5. Corollary 3.2.3 proves Lemma 3.1.5(a), and Corollaries 3.2.6 and 3.2.9 prove part (b). Applying the union bound, the desired result immediately follows. □

Remark 3.2.10. Recall that $|\tilde{y}_i - y_i| \leq 1$ for each $i \leq J$. It follows from Corollary 3.2.6 that $y_i = \mathbb{E}[Y_i](1 + o(1)) \pm 1$ for all $i \leq J$. With this in mind, it is an immediate consequence of Lemma 3.2.8 (say, taking $k = 6.75$ instead of $k = 7$ and $\alpha = \frac{1}{2}\log^{-5} M$ instead of $\alpha = \log^{-5} M$ in the proof) that a.a.s.

$$|Y_i - y_i| \leq \frac{y_i}{\log^5 M} + \log^7 M$$

for all $i \leq J$. This allows us to more directly relate $\boldsymbol{d}_I$ and $\boldsymbol{d}_S$.

3.2.4 A note on isolated vertices in $G[S]$

Many results about random graphs with given degree sequences only consider sequences with minimum degree at least 1. We also make this assumption on $\boldsymbol{d}$ throughout the entire thesis. If $\boldsymbol{d}$ has k terms equal to 0, then a uniformly random graph with degree sequence $\boldsymbol{d}$ has an identical distribution to a uniformly random graph with degree sequence $\boldsymbol{d}^*$ (recall that $\boldsymbol{d}^*$ is the sequence $\boldsymbol{d}$ with all zeroes removed) and k isolated vertices. Thus, allowing isolated vertices is not particularly interesting, and so a minimum degree of 1 is commonly assumed. This means that when applying such results to study $G[S]$, we technically need to apply them to $\boldsymbol{d}_I^*$ and $\boldsymbol{d}_S^*$ instead of $\boldsymbol{d}_I$ and $\boldsymbol{d}_S$. However, if the number of elements in $\boldsymbol{d}_I^*$ and $\boldsymbol{d}_S^*$ deviated wildly from each other, we might not be able to infer our desired properties of $G[S]$ from properties of $\boldsymbol{d}_I^*$. To address this, we show that the numbers of non-zero terms in both sequences are a.a.s. close to each other. This is straightforward to show just using Definition 3.1.2 and Lemma 3.1.5. The applications discussed in this work only require us to know the order of the number of vertices in $G[S]$, rather than its exact value. Thus, we can effectively ignore this issue altogether. The following lemma formalises this.

Lemma 3.2.11. With probability $1 - o(1)$, $n(\boldsymbol{d}_I^*) \sim n(\boldsymbol{d}_S^*)$. Furthermore, $n(\boldsymbol{d}_I^*) = \Theta(|S|)$ always.

Proof. We first show that $n(\boldsymbol{d}_I^*) = \Theta(|S|)$. If $\tilde{y}_0 = o(|S|)$ then the claim holds trivially; thus we assume that $\tilde{y}_0 = \Theta(|S|)$. Note that

$$\frac{\mathbb{P}(Z_k = 1)}{\mathbb{P}(Z_k = 0)} = k\frac{d(S)}{d(\overline{S})}.$$

Thus, by Remark 3.1.6 and the definition of $\boldsymbol{d}_I$ it follows that

$$\begin{aligned} n_1(\boldsymbol{d}_I) &= \sum_{i \in S_{\text{small}}} \mathbb{P}\left(Z_{d(i)} = 1\right) \pm 1 \\ &= \frac{d(S)}{d(\overline{S})} \sum_{i \in S_{\text{small}}} d(i)\mathbb{P}\left(Z_{d(i)} = 0\right) \pm 1. \end{aligned}$$

By assumption, $d(i) \geq 1$ for all $i \in [n]$ and $d(S) = \Theta(M)$. Thus, it follows that $n_1(\boldsymbol{d}_I) = \Omega(n_0(\boldsymbol{d}_I)) \pm 1$. Since we assume that $\tilde{y}_0 = \Theta(|S|)$, this implies that $n_1(\boldsymbol{d}_I) = \Theta(|S|)$. Since $n(\boldsymbol{d}_I^*) \geq n_1(\boldsymbol{d}_I)$, this proves the claim.

Now we prove that a.a.s. $n(\boldsymbol{d}_I^*) \sim n(\boldsymbol{d}_S^*)$. Note that $|S| = \Omega(\sqrt{M}\log^7 M)$, since $\Delta(\boldsymbol{d}) \leq \sqrt{M}/\log^7 M$ and $d(S) = \Theta(M)$. Also note by Lemma 3.1.5 that a.a.s. $|Y_0 - \mathbb{E}[Y_0]| \leq \frac{1}{\log^5 M}\mathbb{E}[Y_0] +$

$\log^7 M$ and $n_0(\boldsymbol{d}_S) = Y_0$. Using Remark 3.2.10 to relate this to y_0, this implies that a.a.s.

$$n(\boldsymbol{d}_S^*) = |S| - Y_0 = |S| - y_0 \pm \left(\frac{1}{\log^5 M} \mathbb{E}[Y_0] + \log^7 M \right).$$

Since $|S| - y_0 = n(\boldsymbol{d}_I^*)$, and $|S| = \omega(\sqrt{M} \log^7 M)$, and $Y_0 \leq |S|$, it follows that a.a.s. $n(\boldsymbol{d}_S^*) = n(\boldsymbol{d}_I^*) + o(|S|)$. From the earlier part of this lemma, $n(\boldsymbol{d}_I^*) = \Theta(|S|)$. Thus, a.a.s. $n(\boldsymbol{d}_I^*) \sim n(\boldsymbol{d}_S^*)$, which concludes the proof. □

As a note on asymptotics, this also means that if one of n, M, $|S|$, or $n(\boldsymbol{d}_I^*)$ goes to infinity, then they all do. Thus, when we defined asymptotics earlier in terms of n and M, we could equivalently define it in terms of $|S|$ or $n(\boldsymbol{d}_I^*)$. We do not explicitly use this anywhere, but it does appear implicitly - some results in the following sections talk about asymptotics in relation to the length or total degree of a degree sequence while assuming a minimum degree of 1. One can rest easy knowing that there are no peculiar cases where $n(\boldsymbol{d}_I)$ is misbehaving and not growing as a function of n or M.

3.3 Basic properties of the induced degree sequence

In this section we apply Lemma 3.1.5 to prove concentration results about some basic properties of $\boldsymbol{d}_S$ and then relate them to the analogous properties for $\boldsymbol{d}_I$. These form a foundation for proving more complicated properties in Chapter 4, such as thresholds for connectivity or the existence of a giant component in $G[S]$.

The following lemma shows that the total degrees of both $\boldsymbol{d}_I$ and $\boldsymbol{d}_S$ are highly concentrated around $\gamma^2 M$, where $\gamma = d(S)/M$. This makes sense intuitively: the total degree of S in G is γM, and the induced degree of each vertex in $v \in S$ is roughly distributed as $\text{Bin}(d(v), \gamma)$. Recall that $J = \log M \log\log M$, S_{big} is the set of vertices in S with degree greater than J in G, and $S_{\text{small}} = S \backslash S_{\text{big}}$. We continue to use the notation that $a = b \pm c$ implies that $a \in [b - c, b + c]$. For brevity, define $M_I := M(\boldsymbol{d}_I)$ and $M_S := M(\boldsymbol{d}_S)$.

Lemma 3.3.1. $M_I \sim \gamma^2 M$ always. Furthermore, a.a.s. $M_S \sim M_I$.

Proof. Recall that $M_I = \sum_{i \in S} d_I(i)$. First split the summation for M_I into

$$M_I = \sum_{i \in S_{\text{big}}} \left\lfloor d(i) \frac{d(S)}{M} \right\rfloor + \sum_{k \leq J} k y_k. \tag{3.14}$$

Now we look at each summation individually. The floor functions in the definition of $\boldsymbol{d}_I$ lower the total value of M_I by at most 1 per element of S_{big}. That is,

$$\left(d(i) \frac{d(S)}{M} - 1 \right) \leq \left\lfloor d(i) \frac{d(S)}{M} \right\rfloor \leq d(i) \frac{d(S)}{M}.$$

Thus it immediately follows that

$$\sum_{i \in S_{\text{big}}} d(i) \frac{d(S)}{M} - |S_{\text{big}}| \leq \sum_{i \in S_{\text{big}}} \left\lfloor d(i) \frac{d(S)}{M} \right\rfloor \leq \sum_{i \in S_{\text{big}}} d(i) \frac{d(S)}{M}. \tag{3.15}$$

Now we focus on the second summation in Equation (3.14), which represents the total degree in $\boldsymbol{d}_I$ of terms in S_{small}. Recall that S_j is the subset of S with degree j in $\boldsymbol{d}$. First note that by Definition 3.1.2 and Equation (3.3),

$$\begin{aligned}\sum_{k\le J} k\tilde{y}_k &= \sum_{k\le J} k \sum_{j\le J} |S_j|\mathbb{P}\left(Z_j = k\right)\\ &= \sum_{j\le J} |S_j| \sum_{k\le J} k\mathbb{P}\left(Z_j = k\right)\\ &= \sum_{j\le J} |S_j|\mathbb{E}\left[Z_j\right]\\ &= \sum_{j\le J} |S_j| j\frac{d(S)}{M}\\ &= \frac{d(S)}{M} d(S_{\text{small}}).\end{aligned}$$

Since y_k is simply the sequence $\tilde{y}_k$ after cascade rounding, it follows that $y_k = \tilde{y}_k \pm 1$ for all $k \le J$. Thus,

$$\sum_{k\le J} k y_k = \frac{d(S)}{M} d(S_{\text{small}}) \pm \sum_{k=1}^{J} k. \tag{3.16}$$

Combining together (3.14) to (3.16), we obtain

$$M_I \le \frac{d(S)^2}{M} + \sum_{k=1}^{J} k \quad \text{and} \quad M_I \ge \frac{d(S)^2}{M} - \sum_{k=1}^{J} k - |S_{\text{big}}|. \tag{3.17}$$

Note that by definition, $|S_{\text{big}}| = o(M)$, since $M > J|S_{\text{big}}|$. Since $\sum_{k=1}^{J} k \le J^2$, which is also $o(M)$, this proves that $M_I \sim d(S)^2/M$. Recalling that $\gamma = d(S)/M$, this proves the first claim of the lemma.

Now we consider the second claim. Recall that

$$M_S = \sum_{v\in S} d_S(v),$$

that is, M_S is the total degree of $G[S]$. We apply Lemma 3.1.5(a) to bound the induced degree of all vertices in S_{big}. Since $\gamma = \Theta(1)$, it follows that a.a.s.

$$\sum_{v\in S_{\text{big}}} d_S(v) = \sum_{v\in S_{\text{big}}} \lfloor \gamma d(v)\rfloor \left(1 + O\left(\frac{1}{\sqrt{\log\log M}}\right)\right).$$

Now we focus on S_{small}. Recall that Y_k is the number of vertices in S_{small} with induced degree k. We apply Lemma 3.1.5(b) and Remark 3.2.10 to the vertices in S_{small}. This implies that a.a.s. Y_k

differs from y_k by at most $\frac{1}{\log^5 M}\mathbb{E}[Y_k] + \log^7 M$ for all $k \le J$. Thus, a.a.s.

$$\begin{aligned}\left|\sum_{k\le J} k\,(Y_k - y_k)\right| &\le \sum_{k\le J} k\left(\frac{1}{\log^5 M}\mathbb{E}[Y_k] + \log^7 M\right)\\ &\le \frac{1}{\log^5 M}\sum_{k\le J} k\mathbb{E}[Y_k] + \log^{10} M\\ &\le \frac{1}{\log^5 M}\sum_{k\le J} k\tilde{y}_k(1+o(1)) + \log^{10} M\\ &\le \frac{M_I}{\log^5 M}(1+o(1)) + \log^{10} M\\ &\le \frac{M}{\log^5 M}.\end{aligned}$$

Thus, combining these results it follows that with probability $1 - o(1)$

$$|M_S - M_I| = O\left(\frac{d_S(S_{\text{big}})}{\sqrt{\log\log M}}\right) + \frac{M}{\log^5 M}.$$

Since $M_I = \Theta(M)$ and $d_S(S_{\text{big}}) \le d(S_{\text{big}}) \le M$, this completes the proof. □

With this result, we can give a short proof of Corollary 3.1.4, which states that $\boldsymbol{d}_I^f$ is always a graphical sequence.

Proof of Corollary 3.1.4. By definition, $M(\boldsymbol{d}_I^f)$ is even. It follows from the inequalities given in (3.17) that $M(\boldsymbol{d}_I) \ge \frac{d(S)^2}{M} - J^2 - |S_{\text{big}}|$, and thus $M(\boldsymbol{d}_I) \ge cM$ for some constant $c > 0$. Finally, it follows that $\Delta(\boldsymbol{d}_I) \le \Delta(\boldsymbol{d}) = o(\sqrt{M})$. Thus, Lemma 3.1.3 implies that $\boldsymbol{d}_I^f$ is a graphical sequence. □

Recall that $d(\boldsymbol{d})$ is the average degree of $\boldsymbol{d}$, that is,

$$d(\boldsymbol{d}) = \frac{M(\boldsymbol{d})}{n(\boldsymbol{d})}.$$

Another useful corollary of Lemma 3.3.1 is that the average degrees of $\boldsymbol{d}_I$ and $\boldsymbol{d}_S$ are a.a.s. equal.

Corollary 3.3.2. With probability $1 - o(1)$, $d(\boldsymbol{d}_I) \sim d(\boldsymbol{d}_S)$.

Proof. By Lemma 3.3.1, it follows that a.a.s. $M(\boldsymbol{d}_S) \sim M(\boldsymbol{d}_I)$, and thus it immediately follows that a.a.s. $M(\boldsymbol{d}_S)/|S| \sim M(\boldsymbol{d}_I)/|S|$. □

As mentioned earlier, we cannot concentrate the induced degree of every individual vertex in $G[S]$ when there exist vertices of low (that is, $O(\log M)$) degree. However, we can argue that the maximum degree of the induced graph is within a constant factor of what is expected.

Lemma 3.3.3. $\Delta(\boldsymbol{d}_I) = \Theta(d(i_s))$ always, and a.a.s. $\Delta(\boldsymbol{d}_S) = \Theta(\Delta(\boldsymbol{d}_I))$.

Proof. Naturally if $S_{\text{big}} \neq \emptyset$, then by Definition 3.1.2 and Lemma 3.1.5 it follows that $\Delta(\boldsymbol{d}_I) \sim \gamma d(i_s)$ and a.a.s. $\Delta(\boldsymbol{d}_I) \sim \Delta(\boldsymbol{d}_S)$. For the remainder of the proof we suppose that $S_{\text{big}} = \emptyset$. If $d(i_s) = O(1)$, then it follows immediately that both $\boldsymbol{d}_S$ and $\boldsymbol{d}_I$ are bounded and the claim holds.

To complete the proof, we show that if $d(i_s) = \omega(1)$ and $S_{\text{big}} = \emptyset$, then $\Delta(\boldsymbol{d}_I) = \Theta(d(i_s))$ and a.a.s. $\Delta(\boldsymbol{d}_S) = \Theta(d(i_s))$.

It is immediate that $\Delta(\boldsymbol{d}_I) \le d(i_s)$ and $\Delta(\boldsymbol{d}_S) \le d(i_s)$. Now we show that $d_I(s) = \Omega(d(i_s))$. Recall that $\gamma = d(S)/M$. Chebyshev's inequality (given in Theorem A.2) implies that

$$\mathbb{P}\left(|Z_{d(i_s)} - \gamma d(i_s)| \ge d(i_s)^{2/3}\right) \le \frac{\operatorname{Var}\left(Z_{d(i_s)}\right)}{d(i_s)^{4/3}} = \frac{d(i_s)\gamma(1-\gamma)}{d(i_s)^{4/3}} = o(1). \tag{3.18}$$

Therefore, by definition of $\tilde{y}_k$,

$$\sum_{k \ge \gamma d(i_s) - d(i_s)^{2/3}} \tilde{y}_k \ge 1 - o(1).$$

Thus, it follows that $y_k \ge 1$ for some $k \ge \gamma d(i_s) - d(i_s)^{2/3}$. This implies that $\Delta(\boldsymbol{d}_I) \ge \gamma d(i_s) - d(i_s)^{2/3}$. By a similar argument, we can show that $\boldsymbol{d}_S$ a.a.s. contains a term of degree at least $\gamma d(i_s) - d(i_s)^{2/3}$:

$$\begin{aligned}
\mathbb{P}\left(d_S(s) \ge \gamma d(i_s) - d(i_s)^{2/3}\right) &= \sum_{i \ge \gamma d(i_s) - d(i_s)^{2/3}} \mathbb{P}\left(d_S(s) = i\right) \\
&= \sum_{i \ge \gamma d(i_s) - d(i_s)^{2/3}} \mathbb{P}\left(Z_{d(i_s)} = i\right)\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right) \\
&= 1 - o(1),
\end{aligned}$$

where the last two equalities follow from Lemma 3.2.5 and Equation (3.18) respectively. Thus, a.a.s. $\boldsymbol{d}_S$ has maximum degree $\Theta(d(i_s))$. This completes the proof. □

Next we prove a concentration result about the sum of the k largest terms in each sequence, for all $k \in [s]$. We in fact prove a slightly more general statement about the following function: for $\boldsymbol{d}$ ordered in non-decreasing order, define

$$D_k(\boldsymbol{d}, t) = \sum_{i=0}^{k-1} (d(n(\boldsymbol{d}) - i))^t.$$

When $t = 1$, this is equal to the sum of the k largest terms in the sequence $\boldsymbol{d}$. By considering larger t, we can consider "moments" of the sequence, which are used in some results about random graphs with given degree sequences. First recall that $\boldsymbol{d}'$ is the sequence $\boldsymbol{d}$ ordered in non-decreasing order, for an arbitrary sequence $\boldsymbol{d}$. The following lemma shows that $D_k(\boldsymbol{d}'_I, t)$ and $D_k(\boldsymbol{d}'_S, t)$ are close for all fixed t and all $k \le |S|$. This lemma may be of independent interest, but in particular allows us to prove results about the chromatic number of $G[S]$, which are given in the next chapter.

Lemma 3.3.4. With probability $1 - o(1)$,

$$|D_k(\boldsymbol{d}'_S, t) - D_k(\boldsymbol{d}'_I, t)| = o(D_k(\boldsymbol{d}'_I, t) + \log^{9+t} M)$$

for all $k \le |S|$ and fixed t.

We defer the proof for a moment to first state two helpful claims. The first claim follows by summing the bound in Lemma 3.1.5(b) over all $i \ge k$, and is presented without proof.

Claim 3.3.5. Let $B_k(\boldsymbol{d}_S)$ be the number of vertices in S_{small} with degree at least k in $G[S]$. Then a.a.s.

$$\left| B_k(\boldsymbol{d}_S) - \sum_{i=k}^{J} y_k \right| = O\left(\sum_{i=k}^{J} \left[\frac{y_k}{\log^5 M} + \log^7 M \right] \right).$$

The second preliminary result is a technical lemma on the effects of reordering the sequences, in particular its effect on the values of the sequence that are at least $\gamma J(1-o(1))$. Recall that $S_{\text{big}} = \{i_{\ell+1}, \dots, i_s\}$ and $S_{\text{small}} = S \backslash S_{\text{big}}$. Define $\hat{\boldsymbol{d}}_S = (d_S(\pi(i)))_{i\in[s]}$ to be a reordering of $\boldsymbol{d}_S$ such that $\hat{\boldsymbol{d}}_S(1) \le \cdots \le \hat{\boldsymbol{d}}_S(\ell)$ and $\pi(i) = i$ for all $i \ge \ell+1$. That is, $\hat{\boldsymbol{d}}_S$ is the degree sequence $\boldsymbol{d}_S$ with terms in S_{small} ordered in non-decreasing order and the terms in S_{big} unchanged. The following claim shows that this new sequence is extremely similar to $\boldsymbol{d}'_S$.

Claim 3.3.6. For all $i \le |S|$, $d_I(i) \sim d'_I(i)$. Furthermore, a.a.s. $\hat{d}_S(i) \sim d'_S(i)$ for all $i \le |S|$.

Proof. Recall that by definition of $\boldsymbol{d}_I$, $d_I(1) \le \cdots \le d_I(\ell)$, where $|S_{\text{small}}| = \ell$, and also that $d_I(\ell+1) \le \cdots \le d_I(s)$. From the definition of $\tilde{y}_k$ and McDiarmid's inequality (given in Theorem A.5), we get that

$$\sum_{k > \gamma J(1+10/\gamma\sqrt{\log\log M})} \tilde{y}_k = \sum_{v \in S_{\text{small}}} \mathbb{P}\left(Z_{d(v)} > \gamma J(1 + 10/\gamma\sqrt{\log\log M}) \right) = o(M^{-10}).$$

Thus, it follows that $y_k = 0$ for all $k > \gamma J(1 + 10/\gamma\sqrt{\log\log M})$; this implies that $d_I(\ell) \le \gamma J(1+o(1))$. By definition of $\boldsymbol{d}_I$, we know that $d_I(\ell+1) \ge \lfloor \gamma J \rfloor$.

Let k_1 be the smallest index such that $d_I(k_1) \ge \lfloor \gamma J \rfloor$, and let k_2 be the largest index such that $d_I(k_2) \le \gamma J(1 + 10/\gamma\sqrt{\log\log M})$. It follows that $d_I(i) = d'_I(i)$ for all $i \notin [k_1, k_2]$. Furthermore, $d_I(i) \sim \gamma J$ for all $i \in [k_1, k_2]$. Thus, $d_I(j) \sim d_I(k)$ for all $j, k \in [k_1, k_2]$. Therefore, $d'_I(i) \sim d_I(i)$ for all $i \in [k_1, k_2]$. This completes the proof of the first claim.

The proof of the second claim is very similar, but complicated slightly by the randomness inherent in $\boldsymbol{d}_S$. By Lemma 3.1.5(a), it follows that a.a.s. $d_S(i) \ge \gamma J(1 - 10/\sqrt{\gamma \log\log M})$ for all $i \in S_{\text{big}}$. Let k_3 be the largest index such that $\hat{d}_S(k_3) < \gamma J(1 - 10/\sqrt{\gamma \log\log M})$. Then a.a.s. $\hat{d}_S(i) = d'_S(i)$ for all $i \le k_3$, since a.a.s. no vertices in S_{big} have induced degree smaller than $\gamma J(1 - 10/\sqrt{\gamma \log\log M})$.

Now choose j such that $\hat{d}_S(j) \ge \gamma J(1 - 10/\sqrt{\gamma \log\log M})$, that is, $j > k_3$. It follows that $d(i_j) \ge 2\gamma^{-1} \log M \log\log\log M$, since $\gamma = \Theta(1)$. Then Remark 3.1.8 implies that a.a.s. $\hat{d}_S(j) = \gamma d(i_j)(1 \pm o(1))$ for all $j > k_3$. Suppose that $d'_S(j) \nsim \hat{d}_S(j)$ for some $j > k_3$. There are then two possible cases:

(a) $d'_S(j) < \hat{d}_S(j)(1 - f(M))$ for all $f(M) = o(1)$, or

(b) $d'_S(j) > \hat{d}_S(j)(1 + f(M))$ for all $f(M) = o(1)$.

However, for a sufficiently slowly-shrinking choice of $f(M)$, the number of terms in $\hat{\boldsymbol{d}}_S$ with degree at least $\hat{d}_S(j)(1-f(M))$ (equivalently, $\gamma d(i_{\pi^{-1}(j)})(1-f(M))$) is a.a.s. at least $s - j + 1$ by Lemma 3.1.5 and Remark 3.1.8. This means that case (a) a.a.s. does not occur. Similarly, there are a.a.s. at least j terms in $\hat{\boldsymbol{d}}_S$ with degree at most $\hat{d}_S(j)(1+f(M))$; this means that case (b) also a.a.s. does not occur. Thus, it follows that a.a.s. $d'_S(j) \sim \hat{d}_S(j)$ a.a.s. for all $j > k_3$. This completes the proof. □

Now we return to the proof of Lemma 3.3.4. For the purposes of the proof we extend the definition of $D_k(\boldsymbol{d}, t)$ to degree sequences that are not ordered in non-decreasing order. In this case, if $\boldsymbol{d}$ is some unordered degree sequence, then $D_k(\boldsymbol{d}, 1)$ is not necessarily the sum of the k largest terms in $\boldsymbol{d}$.

Proof of Lemma 3.3.4. It follows from Claim 3.3.6 that

$$D_k(\hat{\boldsymbol{d}}_S, t) \sim D_k(\boldsymbol{d}'_S, t) \text{ a.a.s. and } D_k(\boldsymbol{d}'_I, t) \sim D_k(\boldsymbol{d}_I, t) \tag{3.19}$$

for all fixed t and all $k \leq s$ (recall $s := |S|$). Thus, to prove Lemma 3.3.4 it is sufficient to prove the claim that a.a.s. $|D_k(\hat{\boldsymbol{d}}_S, t) - D_k(\boldsymbol{d}_I, t)| = o(D_k(\boldsymbol{d}_I, t) + \log^{9+t} M)$ for all $k \leq s$ and all fixed t.

First suppose that $k \leq |S_{\text{big}}|$. Then every term contributing to either sum is an element of S_{big}. Then Lemma 3.1.5(a) implies that a.a.s. $d_S(i) \sim d_I(i)$ for all $i \geq s-k+1$. Since $\hat{d}_S(i) = d_S(i)$ for all $i > \ell$ (since $S_{\text{big}} = \{\ell+1, \dots, s\}$), the claim follows immediately. Now suppose that $k > |S_{\text{big}}|$. Since the terms in S_{big} a.a.s. do not differ much between the two sequences, we now only need to bound the difference over the $k - |S_{\text{big}}|$ largest terms in S_{small} in each sequence. That is, a.a.s.

$$\left|D_k(\hat{\boldsymbol{d}}_S, t) - D_k(\boldsymbol{d}_I, t)\right| = o(D_k(\boldsymbol{d}_I, t)) + \left|\sum_{i=s-\ell}^{k-1} \left(\hat{d}_S(s-i)^t - d_I(s-i)^t\right)\right|. \tag{3.20}$$

Define

$$a = \min\{d_I(s-k+1), \hat{d}_S(s-k+1)\},$$
$$b = \max\{d_I(s-k+1), \hat{d}_S(s-k+1)\}.$$

Recall the definition of y_j from Definition 3.1.2, the number of elements in $\boldsymbol{d}_I$ equal to j with index in S_{small}. The definition of y_j implies that $y_j = 0$ for all $j \geq (1+\varepsilon)\gamma J$ for every constant $\varepsilon > 0$. Thus, it follows that $b < J$. Suppose $a = b$. Applying Claim 3.3.5, the numbers of terms with degree b in S_{small} contributing to $D_k(\boldsymbol{d}_I, t)$ and $D_k(\hat{\boldsymbol{d}}_S, t)$ differ by

$$O\left(\sum_{i=b+1}^{J} \left(\frac{y_k}{\log^5 M} + \log^7 M\right)\right).$$

Combining this with Equation (3.20) implies that

$$\begin{aligned}\left|D_k(\hat{\boldsymbol{d}}_S, t) - D_k(\boldsymbol{d}_I, t)\right| &= o(D_k(\boldsymbol{d}_I, t)) + O\left(\sum_{j=b+1}^{J} j^t \frac{y_j}{\log^5 M} + \log^7 M\right) \\ &\quad + O\left(b^t \sum_{j=b+1}^{J} \left[\frac{y_j}{\log^5 M} + \log^7 M\right]\right) \\ &= o(D_k(\boldsymbol{d}_I, t)) + O\left(\frac{D_k(\boldsymbol{d}_I, t)}{\log^5 M}\right) + O\left(J^{t+1} \log^7 M\right).\end{aligned}$$

Thus, this difference is $o(D_k(\boldsymbol{d}_I, t) + \log^{9+t} M)$ for all fixed t. Now suppose that $a < b$. Without loss of generality, suppose that $d_I(s-k+1) = b$, and thus $\hat{d}_S(s-k+1) = a$. Then all degree b terms in $\hat{\boldsymbol{d}}_S$ in S_{small} contribute to $D_k(\hat{\boldsymbol{d}}_S, t)$. From the definition of b and our assumptions about

$d_I(s-k+1)$ and $d_S(s-k+1)$, we have that a.a.s.

$$\sum_{j=b}^{J} Y_j + |S_{\text{big}}| < k \text{ and } \sum_{j=b}^{J} y_j + |S_{\text{big}}| \geq k.$$

By Claim 3.3.5, this implies that a.a.s.

$$\sum_{j=b}^{J} Y_j + |S_{\text{big}}| \geq k - O\left(\sum_{j=b}^{J}\left[\frac{y_j}{\log^5 M} + \log^7 M\right]\right).$$

Thus, there are $O\left(\sum_{j=b}^{J}\left[\frac{y_j}{\log^5 M} + \log^7 M\right]\right)$ terms in S_{small} with degree less than b contributing to $D_k(\hat{\boldsymbol{d}}_S, t)$. Therefore, in this case a.a.s.

$$\begin{aligned}\left|D_k(\hat{\boldsymbol{d}}_S, t) - D_k(\boldsymbol{d}_I, t)\right| &= o(D_k(\boldsymbol{d}_I, t)) + O\left(\sum_{j=b}^{J} j^t \frac{y_j}{\log^5 M} + \log^7 M\right) \\ &\quad + O\left((b^t - a^t)\sum_{j=b}^{J}\left[\frac{y_j}{\log^5 M} + \log^7 M\right]\right).\end{aligned}$$

By identical reasoning to the case where $a = b$, this is also $o(D_k(\boldsymbol{d}_I, t) + \log^{9+t} M)$ for all fixed t. Thus, a.a.s.

$$|D_k(\hat{\boldsymbol{d}}_S, t) - D_k(\boldsymbol{d}_I, t)| = o(D_k(\boldsymbol{d}_I, t) + \log^{9+t} M).$$

Therefore, it follows from (3.19) that a.a.s.

$$|D_k(\boldsymbol{d}'_S, t) - D_k(\boldsymbol{d}'_I, t)| = o(D_k(\boldsymbol{d}'_I, t) + \log^{9+t} M),$$

which completes the proof. □

Chapter 4

Properties of the induced subgraph

In this chapter we apply the results given in Chapter 3 to determine more complex properties of the induced subgraph $G[S]$. This forms the second part of what we call the reduction approach to studying $G[S]$. Many of the results in the literature about $\mathcal{G}(\boldsymbol{d})$ are of the form "if $\boldsymbol{d}$ has a particular property, then a uniformly random graph with degree sequence $\boldsymbol{d}$ a.a.s. has some property". Proposition 3.1.1 states that, conditional on a fixed degree sequence $\boldsymbol{k}$, the graph $G[S]$ is a uniformly random graph with degree sequence $\boldsymbol{k}$. Furthermore, Lemma 3.1.5 implies that $\boldsymbol{d}_S$ is a.a.s. "close" to $\boldsymbol{d}_I$ in some sense.

With this in mind, we have a framework for predicting a broad range of properties of $G[S]$ by applying known results about random graphs with given degree sequences and Lemma 3.1.5. The aim is to show that if $\boldsymbol{d}_I$ satisfies some conditions, then $\boldsymbol{d}_S$ a.a.s. satisfies the same (or very similar) conditions, and use this to show that a.a.s. $G[S]$ has a particular property. In this chapter we focus on four graph properties in particular: whether $G[S]$ is connected, its chromatic number, whether it has a non-trivial automorphism group, and whether it contains a giant component. For each of these applications, the framework is similar: we introduce (or recall from Chapter 2) a theorem from the literature about the property in question in the random graph model $\mathcal{G}(\boldsymbol{d})$, and then show that $\boldsymbol{d}_S$ a.a.s. satisfies the conditions of the theorem if and only if $\boldsymbol{d}_I$ does, and vice versa.

4.1 Connectivity of $G[S]$

The first property of $G[S]$ that we study using $\boldsymbol{d}_I$, given in Definition 3.1.2, is whether $G[S]$ is likely to be connected or not. Recall Lemma 2.2.8 (restated below for convenience), which characterises when $\mathcal{G}(\boldsymbol{d})$ is a.a.s. connected or not for all graphical sequences $\boldsymbol{d}$ such that $\Delta(\boldsymbol{d})^2 = o(M)$. Since our assumptions on $\boldsymbol{d}$ are more strict than this, we can use this result to determine a threshold for when $G[S]$ is likely to be connected. Recall that $n_k(\boldsymbol{d})$ is the number of terms equal to k in a given sequence $\boldsymbol{d}$.

Lemma 2.2.8. *([56], Theorem 10 & Corollary 11) Let $\boldsymbol{d}$ be a graphical sequence with $\Delta^2 = o(M)$ and minimum degree at least 1, and let $c > 0$ be a fixed constant. Then*

(a) if $n_1(\boldsymbol{d}) = o\left(\sqrt{M}\right)$ and $n_2(\boldsymbol{d}) = o(M)$, then $\mathcal{G}(\boldsymbol{d})$ is a.a.s. connected,

(b) if $n_1(\boldsymbol{d}) = \omega\left(\sqrt{M}\right)$ then a.a.s. $\mathcal{G}(\boldsymbol{d})$ is disconnected,

(c) if $n_1(\boldsymbol{d}) \geq c\sqrt{M}$ *or* $n_2(\boldsymbol{d}) \geq cM$*, then there exists* $\delta = \delta(c) > 0$ *such that for all sufficiently large* n*,* $\mathbb{P}(\mathcal{G}(\boldsymbol{d})$ *disconnected*$) \geq \delta$*.*

To apply this result to $G[S]$, the number of degree 1 and degree 2 vertices in $\boldsymbol{d}_S$ need to be sufficiently concentrated. However, due to the distribution of degrees in $\boldsymbol{d}_I$ (and, by extension, $\boldsymbol{d}_S$), the conditions for connectedness of $G[S]$ look somewhat different to those given in Lemma 2.2.8. For example, Remark 3.1.6 implies that if the number of degree 2 vertices in $G(\boldsymbol{d}_I)$ is $\Omega(J\sqrt{M})$, then number of degree 1 vertices is $\Omega(\sqrt{M})$. Lemma 3.1.5 then implies that similar bounds a.a.s. apply to $\boldsymbol{d}_S$. This means that the conditions in Lemma 2.2.8 (a) and (c) involving degree 2 vertices are superfluous - that is, if $G[S]$ contains $\Omega(M)$ degree 2 vertices, then a.a.s. it also contains $\omega(\sqrt{M})$ degree 1 vertices, and thus the graph is a.a.s. disconnected by applying part (a) of Lemma 2.2.8. Furthermore, the graph $G[S]$ is trivially disconnected if it contains a degree 0 vertex. With this in mind, it is natural to ask not just whether $G[S]$ is connected, but whether $G[S]$ is connected if all degree 0 vertices are also deleted. Recall that $\overline{S} = [n]\backslash S$, and recall that $\tilde{y}_k$, defined in Equation (3.1), is the sum of $\mathbb{P}\left(Z_{d(i)} = k\right)$ over all $i \in S_{\text{small}}$.

Lemma 4.1.1. Let $\boldsymbol{d}$ be an n-element graphical sequence with maximum degree at most $\sqrt{M}/\log^7 M$, and let $S \subset [n]$ be a set such that $d(S) = \Theta(M)$ and $d(\overline{S}) = \Theta(M)$. Then:

(a) If $\tilde{y}_0 = o(1)$, then $G[S]$ is a.a.s. connected.

(b) If $\tilde{y}_0 = \Theta(1)$, then there exists some constant $c > 0$ such that

$$c \leq \mathbb{P}(G[S] \text{ is disconnected}) \leq 1 - c.$$

(c) If $\tilde{y}_0 = \omega(1)$ and $\tilde{y}_1 = o(\sqrt{M})$, then $G[S]$ is a.a.s. disconnected, but is a.a.s. connected if the degree 0 vertices are deleted.

(d) If $\tilde{y}_1 = \Theta(\sqrt{M})$, then $G[S]$ is a.a.s. disconnected, and there exists some constant $c > 0$ such that the probability that $G[S]$ is connected after deleting degree 0 vertices is at least c.

(e) If $\tilde{y}_1 = \omega(\sqrt{M})$, then $G[S]$ is a.a.s. disconnected even after deleting degree 0 vertices.

Lemma 3.1.5(b) implies that $\mathbb{E}[Y_i] \sim \tilde{y}_i$ for all $i \leq J$. This means that if $\tilde{y}_0 = o(1)$, then so is $\mathbb{E}[Y_0]$; similar results follow for $\Theta(1)$ and $\omega(1)$, as well as for $\tilde{y}_1$ and $\mathbb{E}[Y_1]$. The values of these $\tilde{y}_i$ are purely dependent on the degree sequence $\boldsymbol{d}$ and choice of subset S. Thus, even though this lemma does not directly refer to $\boldsymbol{d}_I$, for some pair $(\boldsymbol{d}, S)$ it is still straightforward to check which of (a) – (e) is satisfied.

This lemma is an example of where the rounding in the definition of $\boldsymbol{d}_I$ can hide details about the threshold for some graph properties. For example, if $\mathbb{E}[Y_0] = \frac{1}{3}$, then it follows from the definition of $\tilde{y}_k$ and Lemma 3.1.5(b) that $y_0 = 0$. Thus, $\boldsymbol{d}_I$ has minimum degree 1. However, with probability $c > 0$ there is at least one degree 0 vertex in $G[S]$.

Proof of Lemma 4.1.1. Recall that S_{small} is the set of vertices $v \in S$ such that $d(v) \leq J$ (recall that $J = \log M \log\log M$) and that $S_{\text{big}} = S\backslash S_{\text{small}}$. Corollary 3.2.3 implies that the probability that some vertex in S_{big} has degree 0, 1, or 2 in $G[S]$ is less than M^{-10}. So a.a.s. $G[S]$ contains no

vertices in S_{big} with induced degree 0, 1, or 2. Thus, we focus our attention on S_{small}. Markov's inequality implies that

$$\mathbb{P}(Y_0 \neq 0) = \mathbb{P}(Y_0 \geq 1) \leq \mathbb{E}[Y_0]. \tag{4.1}$$

Let V_k be the indicator variable for the event that $d_S(k) = 0$ for each $k \in S_{\text{small}}$. Lemma 3.1.5(b) states that

$$\mathbb{E}[Y_0] = \tilde{y}_0 \left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right) \sim \tilde{y}_0, \tag{4.2}$$

since $\Delta^2 J = o(M)$. We also use a second moment inequality (given in Lemma A.3) which states that

$$\mathbb{P}(Y_0 > 0) \geq \frac{\mathbb{E}[Y_0]^2}{\mathbb{E}\left[Y_0^2\right]}. \tag{4.3}$$

To evaluate $\mathbb{E}\left[Y_0^2\right]$, we use Lemma 3.2.5. Let $p_i := \mathbb{P}(Z_i = 0)$ for each $i \in S_{\text{small}}$. In line with our previously established convention, the summation range "$j \neq k$" means that we sum over all ordered pairs $(j,k) \in S_{\text{small}}^2$ such that $j \neq k$. Then

$$\begin{aligned}
\mathbb{E}\left[Y_0^2\right] &= \mathbb{E}\left[\sum_{k \in S_{\text{small}}} V_k^2 + \sum_{j \neq k} V_j V_k\right] \\
&= \sum_{k \in S_{\text{small}}} \mathbb{E}[V_k] + \sum_{j \neq k} \mathbb{E}[V_j V_k] \\
&= \mathbb{E}[Y_0] + \sum_{j \neq k} p_{d(j)} p_{d(k)} \left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right) \\
&\leq \left(\mathbb{E}[Y_0] + \mathbb{E}[Y_0]^2\right) \left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right),
\end{aligned}$$

where the last two steps follow from an application of Lemma 3.2.5 and the following naive upper bound:

$$\sum_{j \neq k} p_{d(j)} p_{d(k)} = \sum_{j \in S_{\text{small}}} \sum_{k \in S_{\text{small}}} p_{d(j)} p_{d(k)} - \sum_{i \in S_{\text{small}}} p_{d(j)}^2 \leq \left(\sum_{j \in S_{\text{small}}} p_{d(j)}\right)^2.$$

Applying the inequality given in (4.3) then gives that

$$\mathbb{P}(Y_0 > 0) \geq \frac{\mathbb{E}[Y_0]^2}{\mathbb{E}\left[Y_0^2\right]} \geq \frac{\mathbb{E}[Y_0]^2}{\mathbb{E}[Y_0]\left(1 + \mathbb{E}[Y_0]\right)} \left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right),$$

which simplifies to

$$\mathbb{P}(Y_0 > 0) \geq \frac{1}{\frac{1}{\mathbb{E}[Y_0]} + 1} \left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right). \tag{4.4}$$

Altogether, we have a threshold for the existence of degree 0 vertices in terms of $\mathbb{E}[Y_0]$:

(i) If $\mathbb{E}[Y_0] = o(1)$, then the inequality given in (4.1) implies that there exists no degree 0

vertices with probability $1 - o(1)$.

(ii) If $\mathbb{E}[Y_0] = \Theta(1)$, then together the inequalities given in (4.1) and (4.4) imply that there exists some $c > 0$ such that

$$c \leq \mathbb{P}(Y_0 > 0) \leq 1 - c.$$

(iii) If $\mathbb{E}[Y_0] = \omega(1)$, then the inequality given in (4.4) implies that there exists a degree 0 vertex with probability $1 - o(1)$.

Applying (4.2), this proves claims (a) and (b) of the lemma, as well as half of the claim given in part (c). To prove the second half of the lemma, we need to consider the size of $\mathbb{E}[Y_1]$, the expected number of degree 1 terms in $\boldsymbol{d}_S$, with respect to the value of $M(\boldsymbol{d}_S)$. For $Z_j \sim \text{Bin}\left(j, \frac{d(S)}{M}\right)$, Lemma 3.2.5 implies that the expected numbers of vertices in S_{small} with induced degree 0 and 1 satisfy the following relation:

$$\begin{aligned}
\mathbb{E}[Y_1] &= \sum_{j \leq J} |S_j| \mathbb{P}(Z_j = 1) \left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right) \\
&= \frac{d(S)}{d(\overline{S})} \sum_{j \leq J} j |S_j| \mathbb{P}(Z_j = 0) \left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right) \\
&< J \frac{d(S)}{d(\overline{S})} \mathbb{E}[Y_0] \left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right).
\end{aligned}$$

So if $\tilde{y}_1 = \omega(J)$, then (4.2) and (4.4) imply that a.a.s. there exists at least one vertex with induced degree 0 and thus $G[S]$ is a.a.s. disconnected. To prove the remaining claims of the lemma, recall from Lemma 3.1.5(b) that a.a.s.

$$|Y_i - \mathbb{E}[Y_i]| \leq \frac{1}{\log^5 M} \mathbb{E}[Y_i] + \log^7 M$$

for all $i \leq J$. Specifically for $i = 1$, this implies that a.a.s.

$$Y_1 \leq \mathbb{E}[Y_1] \left(1 + \frac{1}{\log^5 M}\right) + \log^7 M$$

and

$$Y_1 \geq \mathbb{E}[Y_1] \left(1 - \frac{1}{\log^5 M}\right) - \log^7 M.$$

Since $\mathbb{E}[Y_1] \sim \tilde{y}_1$, this implies that if $\tilde{y}_1 = o(\sqrt{M})$, then a.a.s. $Y_1 = o(\sqrt{M})$. Simialrly, if $\tilde{y}_1 = \omega(\sqrt{M})$, then a.a.s. $Y_1 = \omega(\sqrt{M})$. Finally, if $\tilde{y}_1 \geq c\sqrt{M}$, then a.a.s. $Y_1 \geq c'\sqrt{M}$ for every $c' < c$.

Now we apply Lemma 2.2.8 to $\boldsymbol{d}_S$. Lemma 3.3.1 implies that $M(\boldsymbol{d}_I) \sim \gamma^2 M$, and a.a.s. $M(\boldsymbol{d}_S) \sim M(\boldsymbol{d}_I)$. Let $\varepsilon > 0$ be some small constant such that $\varepsilon < \gamma^2$. Suppose that $\tilde{y}_1 = o(\sqrt{M})$ (and thus $\mathbb{E}[Y_1] = o(\sqrt{M})$), and let A be the event that $G[S]$ is connected after deleting degree 0

vertices. Then Lemma 2.2.8(a) and Lemma 3.3.1 together imply that

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(A|\, M(\boldsymbol{d}_S) \geq \varepsilon M)\,\mathbb{P}(M(\boldsymbol{d}_S) \geq \varepsilon M) + \mathbb{P}(A|\, M(\boldsymbol{d}_S) < \varepsilon M)\,\mathbb{P}(M(\boldsymbol{d}_S) < \varepsilon M)\\ &= (1-o(1))(1-o(1)) + o(1)\\ &= 1 - o(1).\end{aligned}$$

This proves the second half of claim (c) in the lemma statement. Claims (d) and (e) follow similarly by applying Lemma 2.2.8(b) and (c). □

4.2 Chromatic number of $G[S]$

In this section, we apply Lemma 2.2.10 to give results on the chromatic number of $G[S]$. We restate the lemma here for a sequence $\boldsymbol{d}$ ordered in non-decreasing order. Recall that for a sequence $\boldsymbol{d}$ ordered in non-decreasing order,

$$D_k(\boldsymbol{d}) := D_k(\boldsymbol{d}, 1) = \sum_{i=0}^{k-1} d(n(\boldsymbol{d}) - i), \tag{4.5}$$

the total degree of the k largest elements in $\boldsymbol{d}$. Also recall that $d(\boldsymbol{d}) = M(\boldsymbol{d})/n(\boldsymbol{d})$ is the average degree of $\boldsymbol{d}$ (which is naturally the same for all graphs in $\mathcal{G}(\boldsymbol{d})$).

Lemma 2.2.10. *([56], Theorem 9) Let $\boldsymbol{d}$ be an n-element graphical sequence ordered in non-decreasing order. Suppose that $\boldsymbol{d}$ satisfies the following conditions:*

(a) $\Delta(\boldsymbol{d}) = o(n)$,

(b) $D_{\Delta(\boldsymbol{d})}(\boldsymbol{d}) = \sum_{i=0}^{\Delta-1} d(n-i) = o(M)$, *and*

(c) there exist constants $\alpha \in \left(\frac{1}{2}, 1\right)$ *and* $\varepsilon, K_0 \in \mathbb{R}^+$ *such that* $D_k(\boldsymbol{d}) \leq K_0 dn \left(\frac{k}{n}\right)^\alpha$ *for all* $k \in \{1, \ldots, \varepsilon n\}$.

Then a.a.s. $\mathcal{G}(\boldsymbol{d})$ *has chromatic number* $\Theta\left(\frac{d(\boldsymbol{d})}{\ln d(\boldsymbol{d})}\right)$.

For the sequences we are considering, the two conditions (a) and (b) are immediately satisfied, since we assume that $\Delta^2 = o(M)$. Thus, condition (c) is the main focus of our attention. Intuitively, this condition is satisfied when the tail of a degree sequence (the sum of its largest elements) is sufficiently small compared to the total degree of the sequence. On the other hand, sequences that do not satisfy this have elements that are much larger than the average degree, and thus have dense tails. For example, if $\boldsymbol{d} = (3, \ldots, 3)$, then condition (c) is satisfied; however, if $\boldsymbol{d} = (3, \ldots, 3, \frac{\sqrt{M}}{\log^7 M})$, then there does not exist a valid choice of α or K_0 for the case $k = 1$ and thus condition (c) is not satisfied.

We show that $\boldsymbol{d}'_S$ a.a.s. satisfies this condition if and only if $\boldsymbol{d}'_I$ satisfies this condition, potentially with slightly different values of K_0 or α. Due to the restrictions on α, we know that $K_0 dn \left(\frac{k}{n}\right)^\alpha = \Omega(n^\varepsilon)$ for some $\varepsilon > 0$. Thus, the required result is a direct consequence of Lemma 3.3.4 applied with $t = 1$. It follows that a.a.s. $D_k(\boldsymbol{d}'_S)$ satisfies part (c) of Lemma 2.2.10 if and only if $D_k(\boldsymbol{d}'_I)$ does too. This is the content of the following corollary. Note that we actually compare $\boldsymbol{d}'_I$ to $\boldsymbol{d}'_S$, since we want to talk about the sequences when they are both ordered.

Theorem 4.2.1. Let $\boldsymbol{d}$ be an n-element graphical sequence with maximum degree at most $\frac{\sqrt{M}}{\log^7 M}$, and let $S \subset [n]$ be a set such that $d(S) = \Theta(M)$ and $d(\overline{S}) = \Theta(M)$. Then if $\boldsymbol{d}'_I$ satisfies condition (c) of Lemma 2.2.10, then $G[S]$ a.a.s. has chromatic number $\Theta\left(\frac{\gamma^2 M/s}{\ln(\gamma^2 M/s)}\right)$.

Proof. We show that the sequence $\boldsymbol{d}'_S$ a.a.s. satisfies the conditions of Lemma 2.2.10 if and only if $\boldsymbol{d}'_I$ satisfies condition (c). The maximum degree restriction on $\boldsymbol{d}$ implies that both sequences $\boldsymbol{d}'_I$ and $\boldsymbol{d}'_S$ immediately satisfy conditions (a) and (b). Corollary 3.3.2 states that a.a.s. $d(\boldsymbol{d}_S) \sim \boldsymbol{d}(\boldsymbol{d}_I)$. Thus, a.a.s.

$$d(\boldsymbol{d}_S)s\left(\frac{k}{s}\right)^{\alpha} = d(\boldsymbol{d}_I)s\left(\frac{k}{s}\right)^{\alpha}(1+o(1)) \tag{4.6}$$

for all $k \leq s$ and all $\alpha \in \left(\frac{1}{2}, 1\right)$. Suppose that $\boldsymbol{d}'_I$ satisfies part (c) of Lemma 2.2.10. Then there exists a triple of constants $(\alpha, \varepsilon, K_0)$ such that

$$D_k(\boldsymbol{d}'_I) \leq K_0 d(\boldsymbol{d}_I)s\left(\frac{k}{s}\right)^{\alpha}$$

for all $k \leq \varepsilon s$. Lemma 3.3.4 then implies that a.a.s. $|D_k(\boldsymbol{d}'_S) - D_k(\boldsymbol{d}'_I)| = o(D_k(\boldsymbol{d}'_I) + \log^{10} M)$ for all such $k \leq s$. Then Equation (4.6) implies that a.a.s. $\boldsymbol{d}'_S$ satisfies part (c) of Lemma 2.2.10 with constants $(\alpha, \varepsilon, 2K_0)$.

Conversely, suppose $\boldsymbol{d}'_I$ does not satisfy Lemma 2.2.10(c). Then for all possible sets of constants $(\alpha, \varepsilon, K_0)$, there exists some $k \leq \varepsilon s$ such that

$$D_k(\boldsymbol{d}'_I) > K_0 d(\boldsymbol{d}_I)s\left(\frac{k}{s}\right)^{\alpha}.$$

Then by Equation (4.6) it similarly follows that a.a.s.

$$D_k(\boldsymbol{d}'_S) > \frac{1}{2}K_0 d(\boldsymbol{d}_I)s\left(\frac{k}{s}\right)^{\alpha}.$$

Since K_0 is arbitrary, this implies that $\boldsymbol{d}'_S$ a.a.s. does not satisfy Lemma 2.2.10(c). Therefore, if $\boldsymbol{d}'_I$ satisfies condition (c) of Lemma 2.2.10, then a.a.s. $\boldsymbol{d}'_S$ satisfies all the conditions of the lemma and $d(\boldsymbol{d}_S) \sim d(\boldsymbol{d}_I)$. This completes the proof. □

4.3 Symmetry of $G[S]$

In this section, we apply the results of Brick et al. [23] to give conditions under which $G[S]$ a.a.s. has (or does not have) non-trivial automorphisms. Recall that we call a graph G symmetric if it has a non-trivial automorphism, and asymmetric if it does not. Specifically, we show that Lemmas 2.2.12 to 2.2.14 can be applied to the sequence $\boldsymbol{d}^*_I$ (since they assume a minimum degree of 1) to prove results about the symmetry of $G[S]$. First we recall Lemma 2.2.12.

Lemma 2.2.12. *([23], Theorem 2) Fix a constant $\Delta > 0$ and assume that $\boldsymbol{d}$ is a graphical sequence where $1 \leq d(i) \leq \Delta$ for all $1 \leq i \leq n$.*

(a) If $n_1 = o(n^{1/2})$ and $n_2 = o(n)$ then a.a.s. $\mathcal{G}(\boldsymbol{d})$ is asymmetric.

(b) If $n_1 = \omega(n^{1/2})$ then a.a.s. $\mathcal{G}(\boldsymbol{d})$ is symmetric.

(c) If there is a constant $c > 0$ such that

$$n_1 > cn^{1/2} \quad \text{or} \quad n_2 > cn$$

then there is $\delta = \delta(c) > 0$ such that for all sufficiently large n

$$\mathbb{P}(\mathcal{G}(\boldsymbol{d}) \text{ is symmetric}) > \delta.$$

Lemma 3.3.3 implies that $\boldsymbol{d}_S$ is a.a.s. bounded if and only if $\boldsymbol{d}_I$ is bounded, and $\boldsymbol{d}_I$ is bounded if and only if $\boldsymbol{d}$ is bounded. Thus, we can assume that $\boldsymbol{d}$ is bounded, otherwise Lemma 2.2.12 does not apply to $\boldsymbol{d}_I$ and a.a.s. does not apply to $\boldsymbol{d}_S$. This simplifies the proof considerably. Recall that Y_i is the number of vertices in S_{small} with induced degree i, and recall from Lemma 3.1.5(b) that $\mathbb{E}[Y_i] \sim \tilde{y}_i$ for all $i \leq J$.

Theorem 4.3.1. Let $\boldsymbol{d}$ be a bounded n-element graphical sequence, and let $S \subset [n]$ be a set such that $d(S),\ d(\overline{S}) = \Theta(M)$. Then $G[S]$ a.a.s. has non-trivial automorphisms, permuting both degree 0 and degree 1 vertices.

Proof. Since $\boldsymbol{d}$ is bounded, we immediately know that $M = \Theta(n)$. Since $d(S) = \Theta(M)$, this also implies that $|S| = \Theta(n)$. Lemma 3.3.1 then implies that $M_I = \Theta(n)$ and a.a.s. $M_S = \Theta(n)$. Since $\boldsymbol{d}$ is bounded, it contains no terms that are greater than $\log M$, and thus $S_{\text{small}} = S$. Lemma 3.2.5 implies that, for all $v \in S$,

$$\mathbb{P}(d_S(v) = 0) = (1-\gamma)^{d(v)}\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right)$$

and

$$\mathbb{P}(d_S(v) = 1) = d(v)\gamma(1-\gamma)^{d(v)-1}\left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right).$$

Since γ and $1-\gamma$ are both bounded away from 0 and $d(v) = O(1)$ by assumption, there exist some constants $c, c' > 0$ such that $\mathbb{P}(d_S(v) = 0) \geq c$ and $\mathbb{P}(d_S(v) = 1) \geq c'$ for all $v \in S$. Thus, by linearity of expectation,

$$\mathbb{E}[Y_0] = \sum_{v \in S_{\text{small}}} \mathbb{P}(d_S(v) = 0) \geq c|S| \quad \text{and} \quad \mathbb{E}[Y_1] = \sum_{v \in S_{\text{small}}} \mathbb{P}(d_S(v) = 1) \geq c'|S|.$$

Lemma 3.1.5 then implies that a.a.s. $|Y_0 - \mathbb{E}[Y_0]| = o(|S|)$ and $|Y_1 - \mathbb{E}[Y_1]| = o(|S|)$, since $|S| = \Theta(M)$ under these assumptions. This implies that a.a.s. $n_0(\boldsymbol{d}_S) = \Theta(|S|)$ and $n_1(\boldsymbol{d}_S) = \Theta(|S|)$. This means that a.a.s. $n_0(\boldsymbol{d}_S) \geq 2$, and thus a.a.s. $G[S]$ has a non-trivial automorphism permuting the degree 0 vertices. Since $n(\boldsymbol{d}_S) = |S| = \Theta(n)$, this also implies that a.a.s. $n_1(\boldsymbol{d}_S) = \Theta(M(\boldsymbol{d}_S))$. Thus, Lemma 2.2.12 implies that a.a.s. $G[S]$ has a non-trivial automorphism that permutes degree 1 vertices. □

Now we focus on the case that $\boldsymbol{d}$ is not bounded. The aim is to apply Lemmas 2.2.13 and 2.2.14 to give results about whether $G[S]$ a.a.s. has non-trivial automorphisms or not. We restate both lemmas below for convenience.

Lemma 2.2.13. *([23], Theorem 3) Suppose there are constants $R_1 > 0$, $R_2 > 0$ and $0 < \varepsilon < 1$ such that a graphical sequence $\boldsymbol{d}$ with minimum degree at least 1 satisfies the following conditions:*

(A1) $\frac{\Delta^2}{d} = o(n^{\frac{1}{6}-\frac{1}{2R_1}-\frac{1}{R_2}})$,

(A2) $\frac{\Delta^2}{d} = o\left(\frac{n^{1/4}}{n_1^{1/2}}\right)$,

(A3) $\frac{\Delta^2}{d} = o\left(\frac{n^{\alpha_2/2}}{n_2^{\alpha_2/2}}\right)$,

(A4) $\left(\frac{n_i}{n}\right)^{\alpha_i(1-\varepsilon)}\left(\frac{\Delta^2}{d}\right)^{2-\varepsilon} = o(1)$, *for* $i \in \{1,2\}$,

where $\alpha_2 = 1/(R_2+4)$, *and* $\alpha_1 = (1-\alpha_2)/(R_1+4)$. *Then* $\mathcal{G}(\boldsymbol{d})$ *is a.a.s. asymmetric.*

Lemma 2.2.14. *([23], Theorem 4) Let* $\boldsymbol{d}$ *be an* n*-vertex graphical sequence with minimum degree at least 1 and assume* $\Delta^2 = o(M)$.

(a) If $M_2 = o(M)$, $n_1 = \omega\left(M/\sqrt{M_2}\right)$, *or* $n_2 = \omega\left(\sqrt{M^3/M_2}\right)$ *then a.a.s.* $\mathcal{G}(\boldsymbol{d})$ *is symmetric.*

(b) If there is a constant $c > 0$ *such that*

$$n_1 > cM/\sqrt{M_2}, \quad \text{or} \quad n_2 > c\sqrt{M^3/M_2},$$

then there exists $\delta = \delta(c) > 0$ *such that for all sufficiently large* n,

$$\mathbb{P}\left(\mathcal{G}(\boldsymbol{d}) \text{ symmetric}\right) > \delta.$$

Showing an analogous result to Lemma 2.2.14 for $G[S]$ is relatively straightforward. If $M_2(\boldsymbol{d}_S) \sim M_2(\boldsymbol{d}_I)$ a.a.s., then all conditions of Lemma 2.2.14 that are satisfied by $\boldsymbol{d}_I$ are a.a.s. also satisfied by $\boldsymbol{d}_S$. The conditions of Lemma 2.2.13 are more nuanced, and thus a little more work is required to show $\boldsymbol{d}_I$ and $\boldsymbol{d}_S$ a.a.s. satisfy conditions (A1) to (A4) at the same time. The issue arises from condition (A4): if $n_1(\boldsymbol{d}_I)$ or $n_2(\boldsymbol{d}_I)$ are $O(\log^7 M)$, then Lemma 3.1.5 implies that a.a.s. $n_i(\boldsymbol{d}_S) = O(\log^7 M)$ too — however, the actual value could theoretically could be anywhere between $\Theta(\log^7 M)$ and 0. For certain values of R_1, R_2, ε, and Δ, it is plausible that this difference could be enough to alter whether the condition (A4) is satisfied.

Thankfully, we are saved by the quantifiers in Lemma 2.2.13: we do not necessarily need $\boldsymbol{d}_S$ and $\boldsymbol{d}_I$ to satisfy conditions (A1) to (A4) for the same constants (R_1, R_2, ε). This flexibility allows us to get past this pathological case by changing ε slightly for one or both of the sequences. This gives the following result about automorphisms of $G[S]$. Here we ignore degree 0 terms in both $\boldsymbol{d}_S$ and $\boldsymbol{d}_I$, since Lemma 2.2.13 assumes a minimum degree of at least 1. Thus, we compare $\boldsymbol{d}_S^*$ and $\boldsymbol{d}_I^*$, where the degree 0 terms have been removed from each sequence.

Theorem 4.3.2. Let $\boldsymbol{d}$ be an n-element graphical sequence with maximum degree at most $\frac{\sqrt{M}}{\log^7 M}$, and let $S \subset [n]$ be a set such that $d(S),\ d(\overline{S}) = \Theta(M)$.

(a) If $\boldsymbol{d}_I^*$ satisfies conditions (A1) to (A4) of Lemma 2.2.13, then $G[S]$ a.a.s. has no non-trivial automorphisms other than automorphisms of isolated vertices.

(b) If $\boldsymbol{d}_I^*$ satisfies part (a) of Lemma 2.2.14, then $G[S]$ is a.a.s. symmetric. If $\boldsymbol{d}_I^*$ satisfies part (b) of Lemma 2.2.14, then there exists some $\delta > 0$ such that for all sufficiently large n, $\mathbb{P}\left(\mathcal{G}(\boldsymbol{d}) \text{ symmetric}\right) \geq \delta$.

Similar to our previous methods, we prove this theorem by proving that $\boldsymbol{d}_S^*$ a.a.s. satisfies the conditions of each lemma if and only if $\boldsymbol{d}_I^*$ satisfies the conditions. Notably, we do not explicitly show that each of (A1) to (A4) are a.a.s. satisfied by $\boldsymbol{d}_S^*$ if and only if they are satisfied by $\boldsymbol{d}_I^*$. Instead we show that $\boldsymbol{d}_S^*$ a.a.s. satisfies all four conditions at once if and only if $\boldsymbol{d}_I^*$ satisfies all four conditions.

Lemma 4.3.3. Let $\boldsymbol{d}$ be an n-element graphical sequence with maximum degree at most $\frac{\sqrt{M}}{\log^7 M}$, and let $S \subset [n]$ be a set such that $d(S),\ d(\overline{S}) = \Theta(M)$.

(a) The degree sequence $\boldsymbol{d}_S^*$ a.a.s. satisfies conditions (A1) to (A4) of Lemma 2.2.13 if and only if $\boldsymbol{d}_I^*$ satisfies the conditions.

(b) The degree sequence $\boldsymbol{d}_S^*$ a.a.s. satisfies part (a) of Lemma 2.2.14 if and only if $\boldsymbol{d}_I^*$ satisfies part (a). An identical statement holds for part (b).

We start with Lemma 4.3.3(b), as it is more straightforward than the proof of Lemma 4.3.3(a). We know from Lemma 3.3.1 that a.a.s. $M(\boldsymbol{d}_S) \sim M(\boldsymbol{d}_I)$, and since $M(\boldsymbol{d}) = M(\boldsymbol{d}^*)$ this immediately implies the same result for $\boldsymbol{d}_S^*$ and $\boldsymbol{d}_I^*$. The only other necessary ingredient is to show that the same holds for $M_2(\boldsymbol{d}_S^*)$, where $M_2(\boldsymbol{d}) = \sum_{i\in[n]} d(i)\,(d(i)-1)$. This is proved in the following lemma. This proof is very similar to the corresponding proof for the concentration of $M(\boldsymbol{d}_S)$ around $M(\boldsymbol{d}_I)$. Recall the notation that $a = b \pm c$ means that $a \in [b-c, b+c]$.

Lemma 4.3.4. With probability $1 - o(1)$, $|M_2(\boldsymbol{d}_S^*) - M_2(\boldsymbol{d}_I^*)| = o(M_2(\boldsymbol{d}_I^*) + |S|)$.

Proof. Recall that Y_i is the number of vertices in S_{small} with induced degree i. We can split the sum for $M_2(\boldsymbol{d}_S^*)$ into one sum over S_{big} and another sum over S_{small}:

$$\begin{aligned} M_2(\boldsymbol{d}_S^*) &= \sum_{k\in S_{\text{big}}} d_S(k)\,(d_S(k)-1) + \sum_{k\in S_{\text{small}}} d_S(k)\,(d_S(k)-1) \\ &= \sum_{k\in S_{\text{big}}} d_S(k)^2 - \sum_{k\in S_{\text{big}}} d_S(k) + \sum_{k=1}^{J} k(k-1)Y_k. \end{aligned}$$

By Lemma 3.1.5, we can relate each of these terms to their corresponding values for $\boldsymbol{d}_I^*$. Applying this lemma it follows that, with probability $1 - o(1)$,

$$\begin{aligned} M_2(\boldsymbol{d}_S^*) = &\sum_{k\in S_{\text{big}}} \left(d(i_k)\frac{d(S)}{M}(1+o(1))\right)^2 - \sum_{k\in S_{\text{big}}} d(i_k)\frac{d(S)}{M}(1+o(1)) \\ &+ \sum_{k=1}^{J} k(k-1)\mathbb{E}\left[Y_k\right] + \sum_{k=1}^{J} k(k-1)O\left(\frac{\mathbb{E}\left[Y_k\right]}{\log^5 M} + \log^7 M\right). \end{aligned} \tag{4.7}$$

Since $d(i_k)\frac{d(S)}{M} \to \infty$, it follows immediately that

$$\begin{aligned} &\sum_{k\in S_{\text{big}}} \left(d(i_k)\frac{d(S)}{M}(1+o(1))\right)^2 - \sum_{k\in S_{\text{big}}} d(i_k)\frac{d(S)}{M}(1+o(1)) \\ &= \sum_{k\in S_{\text{big}}} d(i_k)\frac{d(S)}{M}\left(d(i_k)\frac{d(S)}{M} - 1\right)(1+o(1)), \end{aligned}$$

since the $1 + o(1)$ factor is uniformly bounded for all $k \in S_{\text{big}}$. For the latter two sums in Equation (4.7), note that the sum $\sum_{k=1}^{J} k(k-1)$ is less than J^3. This means that

$$\sum_{k=1}^{J} k(k-1)\left(\frac{\mathbb{E}[Y_k]}{\log^5 M} + \log^7 M\right) \leq \frac{J^2}{\log^5 M}\sum_{k=1}^{J}\mathbb{E}[Y_k] + J^3\log^7 M$$
$$\leq J^2\left(\frac{|S_{\text{small}}|}{\log^5 M} + J\log^7 M\right).$$

Recall from Lemma 3.1.5(b), Equation (3.1), and the definition of y_k given in (3.2) that $\mathbb{E}[Y_k] \sim \tilde{y}_k$ and $y_k = \tilde{y}_k \pm 1$ for all $k \leq J$. Altogether, this implies that a.a.s.

$$\begin{aligned} M_2(\boldsymbol{d}_S^*) &= \sum_{k\in S_{\text{big}}} d(i_k)\frac{d(S)}{M}\left(d(i_k)\frac{d(S)}{M} - 1\right)(1+o(1)) + \sum_{k=1}^{J} k(k-1)\tilde{y}_k(1+o(1)) \\ &\quad + O\left(J^2\left(\frac{|S_{\text{small}}|}{\log^5 M} + J\log^7 M\right)\right) \\ &= M_2(\boldsymbol{d}_I)(1+o(1)) + O\left(J^3\left(\frac{|S_{\text{small}}|}{\log^5 M} + \log^7 M\right)\right). \end{aligned}$$

We know that $J^3 = o(\log^5 M)$, $|S_{\text{small}}| \leq |S|$, and $|S| \geq \sqrt{M}\log^7 M$. Thus, it immediately follows that

$$J^3\left(\frac{|S_{\text{small}}|}{\log^5 M} + \log^7 M\right) = o(|S|).$$

This proves the lemma. □

With this result, the proof of Lemma 4.3.3(b) is straightforward. Recall that for all $k > 0$, $n_k(\boldsymbol{d}) = n_k(\boldsymbol{d}^*)$. Thus, $M(\boldsymbol{d}) = M(\boldsymbol{d}^*)$ and $M_2(\boldsymbol{d}) = M_2(\boldsymbol{d}^*)$. With this in mind, we sometimes omit the asterisk to avoid clutter.

Proof of Lemma 4.3.3(b). By Lemma 3.3.1, we know that a.a.s. $M(\boldsymbol{d}_S) \sim M(\boldsymbol{d}_I) \sim \frac{d(S)^2}{M}$. Since $M \geq |S|$, it follows from Lemma 4.3.4 that $M_2(\boldsymbol{d}_I) = o(M)$ if and only if a.a.s. $M_2(\boldsymbol{d}_S) = o(M)$ as well. Thus, if $M_2(\boldsymbol{d}_I) = o(M)$, both sequences a.a.s. satisfy part (a) of Lemma 2.2.14. For the remainder of the proof, suppose that $M_2(\boldsymbol{d}_I) = \Omega(M)$. This implies that $M_2(\boldsymbol{d}_I) = \Omega(|S|)$, and also implies that a.a.s. $M_2(\boldsymbol{d}_I) \sim M_2(\boldsymbol{d}_S)$. Thus, in this case it follows that a.a.s.

$$M(\boldsymbol{d}_I^*)/\sqrt{M_2(\boldsymbol{d}_I^*)} \sim M(\boldsymbol{d}_S^*)/\sqrt{M_2(\boldsymbol{d}_S^*)} \tag{4.8}$$

and

$$\sqrt{M(\boldsymbol{d}_I^*)^3/M_2(\boldsymbol{d}_I^*)} \sim \sqrt{M(\boldsymbol{d}_S^*)^3/M_2(\boldsymbol{d}_S^*)}. \tag{4.9}$$

From Remark 3.1.7 we know that a.a.s. $n_i(\boldsymbol{d}_S) = Y_i$ for $i \in \{1, 2\}$. Also, recall from Lemma 3.1.5 and Remark 3.2.10 that a.a.s.

$$|Y_i - y_i| \leq \frac{y_i}{\log^5 M} + \log^7 M$$

for all $i \leq J$. Since $n_i(\boldsymbol{d}_I) = y_i$ for $i \in \{1, 2\}$, it follows that a.a.s.

$$n_i(\boldsymbol{d}_S) = n_i(\boldsymbol{d}_I) + o(n_i(\boldsymbol{d}_I)) \pm \log^7 M \tag{4.10}$$

for $i \in \{1, 2\}$. Recall that $\Delta = o(\sqrt{M})$ and $M(\boldsymbol{d}_I) = \Theta(M)$. This implies that

$$\frac{M(\boldsymbol{d}_I^*)}{\sqrt{M_2(\boldsymbol{d}_I^*)}} \geq \frac{M(\boldsymbol{d}_I^*)}{\sqrt{\Delta(\boldsymbol{d}_I^*)M(\boldsymbol{d}_I^*)}} \geq M(\boldsymbol{d}_I^*)^{1/4} = \omega(\log^7 M)$$

and

$$\sqrt{\frac{M(\boldsymbol{d}_I^*)^3}{M_2(\boldsymbol{d}_I^*)}} \geq \sqrt{\frac{M(\boldsymbol{d}_I^*)^3}{\Delta(\boldsymbol{d}_I^*)M(\boldsymbol{d}_I^*)}} \geq M(\boldsymbol{d}_I^*)^{3/4} = \omega(\log^7 M).$$

As a result of (4.8) and (4.9), the same inequalities hold a.a.s. when $\boldsymbol{d}_I^*$ is replaced with $\boldsymbol{d}_S^*$. From here we split the proof into two cases depending on the size of $\mathbb{E}[Y_i]$. Firstly, if $n_i(\boldsymbol{d}_I) \leq 3\log^7 M$ for either $i = 1$ or $i = 2$, then a.a.s. $n_i(\boldsymbol{d}_S) \leq 5\log^7 M$ by Equation (4.10). Otherwise, if $n_i(\boldsymbol{d}_I) \geq 3\log^7 M$ for $i \in \{1, 2\}$, then Equation (4.10) implies that a.a.s. $n_i(\boldsymbol{d}_S) = \Omega(n_i(\boldsymbol{d}_I))$.

Thus, if $n_1(\boldsymbol{d}_I^*) \geq cM(\boldsymbol{d}_I^*)/\sqrt{M_2(\boldsymbol{d}_I^*)}$ or $n_2(\boldsymbol{d}_I^*) \geq c\sqrt{M(\boldsymbol{d}_I^*)^3/M_2(\boldsymbol{d}_I^*)}$ for some c such that $c = \Omega(1)$, then a.a.s. the corresponding inequality is holds for $\boldsymbol{d}_S^*$ if c is replaced with $c/2$, and vice versa. Therefore, $\boldsymbol{d}_I^*$ satisfies either condition (a) or (b) of Lemma 2.2.14 if and only if $\boldsymbol{d}_S^*$ a.a.s. satisfies the same condition. This concludes the proof. □

The remaining step is to prove Lemma 4.3.3(a). The idea for proving this part is similar to part (b), albeit slightly more complicated. Here we give an outline of the proof. Corollary 3.3.2 implies that the average degree $d(\boldsymbol{d}_S)$ is highly concentrated, and Lemma 3.3.3 implies that the maximum degree $\Delta(\boldsymbol{d}_S)$ is either concentrated or bounded (that is, $O(1)$). Furthermore, as mentioned in the proof of Lemma 4.3.3(b), $n_1(\boldsymbol{d}_S)$ and $n_1(\boldsymbol{d}_I)$ are a.a.s. either asymptotic or $O(\log^7 M)$; the same claim can be made about n_2. Since conditions (A1) – (A4) are only concerned with asymptotic values, if $n_1(\boldsymbol{d}_S) \sim n_1(\boldsymbol{d}_I)$ and $n_2(\boldsymbol{d}_S) \sim n_2(\boldsymbol{d}_I)$ then each condition is either satisfied by both sequences or satisfied by neither sequence.

Thus, the only issues that arise come from cases where $n_i(\boldsymbol{d}_I) = O(\log^7 M)$ for $i = 1$ or 2. For example, consider some hypothetical degree sequence $\boldsymbol{d}$ such that $\Delta^2/d = n^{1/4}/\log^2 n$ and $n_1 = O(\log^7 M)$. If $n_1(\boldsymbol{d}) = \log^7 M$, then $\boldsymbol{d}$ does not satisfy condition (A2), but if $n_1(\boldsymbol{d}) = 1$ then the condition is satisfied. This looks problematic, until we remember that there are three other conditions that $\boldsymbol{d}$ has to satisfy before Lemma 2.2.13 applies. In particular, under these circumstances $\boldsymbol{d}$ fails condition (A4): by choosing a sufficiently small ε and recalling that $\alpha_1 < \frac{1}{4}$, it follows that

$$\left(\frac{n_1}{n}\right)^{\alpha_1(1-\varepsilon)} \left(\frac{\Delta^2}{d}\right)^{2-\varepsilon} = \log^{O(1)} n n^{\frac{1}{4}(2-\varepsilon)-\alpha_1(1-\varepsilon)} = \omega(1),$$

where $\log^{O(1)} n$ refers to a non-specific function of n which is $O(\log^k n)$ for some constant $k \in \mathbb{R}$. The proof of Lemma 4.3.3(b) generalises this idea: if $n_i(\boldsymbol{d}_I)$ is large for one of $i \in \{1, 2\}$, then $n_i(\boldsymbol{d}_S) \sim n_i(\boldsymbol{d}_I)$ and all the desired results follow quickly. However, if $n_i(\boldsymbol{d}_I)$ is small (that is, $O(\log^7 M)$), then we cannot argue that a.a.s. $n_i(\boldsymbol{d}_S) \sim n_i(\boldsymbol{d}_I)$. Depending on the value of $\Delta(\boldsymbol{d})$, it is then possible that the concentration ranges given in Lemma 3.1.5 are not sufficient to argue that $\boldsymbol{d}_S^*$ a.a.s. satisfies a particular condition if $\boldsymbol{d}_I^*$ satisfies that condition. In this case, we show

that $\boldsymbol{d}_I^*$ and $\boldsymbol{d}_S^*$ a.a.s. must both fail one of the other conditions, or the constants (R_1, R_2, ε) can be changed such that both sequences a.a.s. satisfy the particular condition in question. To help with this, we define the notion of $\boldsymbol{d}_I$ "predicting" properties of $\boldsymbol{d}_S$: we write "$\boldsymbol{d}_I$ predicts property K for $\boldsymbol{d}_S$" if $\boldsymbol{d}_S$ a.a.s. has property K if and only if $\boldsymbol{d}_I$ has property K (recall that $\boldsymbol{d}_I$ is deterministic). This analogously carries over for $\boldsymbol{d}_I^*$ predicting properties of $\boldsymbol{d}_S^*$.

Proof of Lemma 4.3.3(a). First we give an outline of the steps. We show that $\boldsymbol{d}_I^*$ predicts (A1) for $\boldsymbol{d}_S^*$. We then show that either $\boldsymbol{d}_I^*$ predicts (A2) and (A3) for $\boldsymbol{d}_S^*$ or $\boldsymbol{d}_I$ predicts (A4) for $\boldsymbol{d}_S^*$ and does not satisfy it. Finally, we show that if $\boldsymbol{d}_I^*$ predicts and satisfies (A1) – (A3), then there exists a choice of $\varepsilon > 0$ such that $\boldsymbol{d}_I^*$ predicts (A4) for $\boldsymbol{d}_S^*$ with constants (R_1, R_2, ε). Thus, either $\boldsymbol{d}_I^*$ predicts (A1) – (A4) for $\boldsymbol{d}_S^*$ or they both fail at least one condition. Therefore, $\boldsymbol{d}_S^*$ satisfies (A1) – (A4) if and only if $\boldsymbol{d}_I^*$ satisfies (A1) – (A4).

Lemma 3.3.1 states that a.a.s. $M(\boldsymbol{d}_S^*) \sim M(\boldsymbol{d}_I^*)$, since $M(\boldsymbol{d}) = M(\boldsymbol{d}^*)$ for all sequences. Lemma 3.2.11 states that a.a.s. $n(\boldsymbol{d}_S^*) \sim n(\boldsymbol{d}_I^*)$. Therefore, a.a.s.

$$d(\boldsymbol{d}_S^*) = \frac{M(\boldsymbol{d}_S^*)}{n(\boldsymbol{d}_S^*)} \sim \frac{M(\boldsymbol{d}_I^*)}{n(\boldsymbol{d}_I^*)} = d(\boldsymbol{d}_I^*).$$

Lemma 3.3.3 states that a.a.s. $\Delta(\boldsymbol{d}_S) = \Theta(\Delta(\boldsymbol{d}_I))$. Thus, if there exist constants (R_1, R_2) such that $\boldsymbol{d}_I^*$ satisfies condition (A1), then a.a.s. $\boldsymbol{d}_S^*$ also satisfies (A1) with the same constants.

Remark 3.2.10 implies that if $n_i(\boldsymbol{d}_I) \geq 3\log^7 M$, then a.a.s. $n_i(\boldsymbol{d}_S) \geq (1-\varepsilon)n_i(\boldsymbol{d}_I) - \log^7 M$ for every $\varepsilon > 0$. It also implies that if $n_i(\boldsymbol{d}_I) < 3\log^7 M$, then a.a.s. $n_i(\boldsymbol{d}_S) < 5\log^7 M$. We use these as two cases to delineate the proof into four parts. First consider the case where $n_1(\boldsymbol{d}_I),\ n_2(\boldsymbol{d}_I) \geq 3\log^7 M$. Then, for $i \in \{1, 2\}$, each pair of terms $n_i(\boldsymbol{d}_I)$ and $n_i(\boldsymbol{d}_S)$ a.a.s. differ by at most a factor of 3. Since a.a.s. the values of $\Delta^2(\boldsymbol{d}_I^*)/d(\boldsymbol{d}_I^*)$ and $\Delta^2(\boldsymbol{d}_S^*)/d(\boldsymbol{d}_S^*)$ are also within a constant factor of each other, it follows immediately that $\boldsymbol{d}_I^*$ predicts each property (A1) – (A4) for $\boldsymbol{d}_S^*$ in this case. Therefore, $\boldsymbol{d}_S^*$ a.a.s. satisfies (A1) – (A4) if and only if $\boldsymbol{d}_I^*$ satisfies (A1) – (A4) in this case.

For the remaining parts of the proof we consider the case where $n_i(\boldsymbol{d}_I^*) < 3\log^7 M$ for at least one of the values $i \in \{1, 2\}$. First we show that if $n_1(\boldsymbol{d}_I^*) < 3\log^7 M$ (and thus a.a.s. $n_1(\boldsymbol{d}_S^*) < 5\log^7 M$ as well) and $\boldsymbol{d}_I^*$ does not predict whether $\boldsymbol{d}_S^*$ satisfies (A2), then $\boldsymbol{d}_I^*$ does not satisfy (A4) and a.a.s. $\boldsymbol{d}_S^*$ does not satisfy it either. For simplicity, define $\log^{O(1)} M$ to be a stand-in for arbitrary functions that asymptotically grow faster than $\log^{-C} M$ but slower than $\log^C M$, for some constant C. Suppose that condition (A2) is satisfied by $\boldsymbol{d}_I^*$ but a.a.s. not satisfied by $\boldsymbol{d}_S^*$, or vice versa. Then it must be the case that

$$\frac{\Delta^2 n_1^{1/2}}{dn^{1/4}} = \log^{O(1)} M$$

both for $\boldsymbol{d}_I^*$ and a.a.s. for $\boldsymbol{d}_S^*$. Since we are assuming that $n_1 < 5\log^7 M$ for both sequences, this in turn implies that $\Delta^2/d = n^{1/4}/\log^{O(1)} M$ for both sequences. Note that in this equation, as well as in the rest of this proof, if a function of a degree sequence (for example, Δ, n, d) does not specify a particular degree sequence, then we implicitly mean that the equation holds both for $\boldsymbol{d}_I^*$

and a.a.s. for $\boldsymbol{d}_S^*$. Now if we consider condition (A4) it follows that

$$\left(\frac{n_1}{n}\right)^{\alpha_1(1-\varepsilon)}\left(\frac{\Delta^2}{d}\right)^{2-\varepsilon} = n^{\frac{1}{2}-\alpha_1-\varepsilon(1/4-\alpha_1)}\log^{O(1)} M.$$

Since $\alpha_1 < 1/4$ and $\varepsilon < 1$ by definition, the power of n in this equation is bounded from below by

$$\frac{1}{2}-\alpha_1-\varepsilon(1/4-\alpha_1) > \frac{1}{2}-\alpha_1-1/4+\alpha_1 = \frac{1}{4}.$$

Thus, for both $\boldsymbol{d}_I^*$ and a.a.s. for $\boldsymbol{d}_S^*$, we get that

$$\left(\frac{n_1}{n}\right)^{\alpha_1(1-\varepsilon)}\left(\frac{\Delta^2}{d}\right)^{2-\varepsilon} = \omega\left(n^{1/4}\log^{O(1)} M\right) = \omega(1).$$

Therefore $\boldsymbol{d}_I^*$ does not satisfy condition (A4), and a.a.s. neither does $\boldsymbol{d}_S^*$.

Now we show that if $n_2(\boldsymbol{d}_I^*) < 3\log^7 M$ and $\boldsymbol{d}_I^*$ does not predict whether $\boldsymbol{d}_S^*$ satisfies (A3), then again $\boldsymbol{d}_I^*$ does not satisfy (A4) and a.a.s. $\boldsymbol{d}_S^*$ also does not satisfy this condition. Suppose that $\boldsymbol{d}_I^*$ does not predict (A3) for $\boldsymbol{d}_S^*$, that is, (A3) is satisfied by $\boldsymbol{d}_I^*$ but not a.a.s. satisfied by $\boldsymbol{d}_S^*$ (or vice versa). Then it must be true that

$$\frac{\Delta^2}{d}\left(\frac{n_2}{n}\right)^{\alpha_2/2} = \log^{O(1)} M$$

for $\boldsymbol{d}_I^*$ as well as a.a.s. for $\boldsymbol{d}_S^*$. Since $n_2 < 5\log^7 M$, this implies that

$$\frac{\Delta^2}{d} = n^{\alpha_2/2}\log^{O(1)} M.$$

Now if we consider condition (A4), it follows that

$$\left(\frac{n_2}{n}\right)^{\alpha_2(1-\varepsilon)}\left(\frac{\Delta^2}{d}\right)^{2-\varepsilon} = n^{\alpha_2(1-\frac{\varepsilon}{2}-(1-\varepsilon))}\log^{O(1)} M = n^{\alpha_2\varepsilon/2}\log^{O(1)} M,$$

which is $\omega(1)$ for every constant $\varepsilon > 0$, regardless of what the $\log^{O(1)} M$ function is. Thus, if $\boldsymbol{d}_I^*$ satisfies condition (A3) but $\boldsymbol{d}_S^*$ does not a.a.s. satisfy the condition (or vice versa), then $\boldsymbol{d}_I^*$ fails condition (A4), and $\boldsymbol{d}_S^*$ also a.a.s. fails (A4). Therefore, either $\boldsymbol{d}_I^*$ predicts (A1) – (A3) for $\boldsymbol{d}_S^*$ or $\boldsymbol{d}_I^*$ does not satisfy (A4) (for every constant $\varepsilon > 0$) and a.a.s. $\boldsymbol{d}_S^*$ does not satisfy (A4).

Now suppose that $\boldsymbol{d}_I^*$ predicts (A1) – (A3) for $\boldsymbol{d}_S^*$ and also satisfies these conditions. Naturally, if $\boldsymbol{d}_I^*$ predicts but does not satisfy (A4) for at least one of the values $i \in \{1,2\}$, then $\boldsymbol{d}_I^*$ does not satisfy (A1) – (A4) and a.a.s. neither does $\boldsymbol{d}_S^*$. Thus, we suppose that for each of $i \in \{1,2\}$, either $\boldsymbol{d}_I^*$ predicts (A4) for $\boldsymbol{d}_S^*$ and satisfies the condition, or $\boldsymbol{d}_I^*$ does not predict the condition. Again recall that if $n_1(\boldsymbol{d}_I^*) \geq 3\log^7 M$ and $n_2(\boldsymbol{d}_I^*) \geq 3\log^7 M$, then it immediately follows that $\boldsymbol{d}_I^*$ predicts (A4) for $\boldsymbol{d}_S^*$. Thus, we assume that at least one of $n_1(\boldsymbol{d}_I^*)$ and $n_2(\boldsymbol{d}_I^*)$ are less than $3\log^7 M$.

First suppose that $\boldsymbol{d}_I^*$ predicts and satisfies (A4) for $i = 2$. Also suppose that $n_1(\boldsymbol{d}_I^*) < 3\log^7 M$ and that there exists a triple of constants (R_1, R_2, ε) such that condition (A4) is satisfied by $\boldsymbol{d}_I^*$

but not a.a.s. satisfied by $\boldsymbol{d}_S^*$ (or vice versa) for this set of constants and $i = 1$. This implies that

$$\left(\frac{n_1}{n}\right)^{\alpha_1(1-\varepsilon)} \left(\frac{\Delta^2}{d}\right)^{2-\varepsilon} = \log^{O(1)} M,$$

both for $\boldsymbol{d}_I^*$ and a.a.s. for $\boldsymbol{d}_S^*$. Since $n_1 < 5\log^7 M$, this means that

$$\left(\frac{\Delta^2}{d}\right)^{2-\varepsilon} n^{-\alpha_1(1-\varepsilon)} = \left(\frac{\Delta^2/d}{n^{\alpha_1}}\right)^{1-\varepsilon} \left(\frac{\Delta^2}{d}\right) = \log^{O(1)} M. \tag{4.11}$$

This implies two things: firstly, $n^{\alpha_1} = \omega(\Delta^2/d)$, and secondly,

$$\frac{\Delta^2}{d} = n^{\alpha_1(1-\varepsilon)/(2-\varepsilon)} \log^{O(1)} M. \tag{4.12}$$

Since $\alpha_1 < 1/4$, this implies that $\Delta^2/d = o(n^{1/4-\delta})$ for some constant $\delta > 0$. This implies that by choosing $\varepsilon' = \varepsilon/2$, it follows that

$$\begin{aligned}\left(\frac{\Delta^2}{d}\right)^{2-\varepsilon/2} n^{-\alpha_1(1-\varepsilon/2)} &= \left(\frac{\Delta^2}{d}\right)^{2-\varepsilon} n^{-\alpha_1(1-\varepsilon)} \left(\frac{\Delta^2/d}{n^{\alpha_1}}\right)^{\varepsilon/2} \\ &= \left(\frac{n^{\alpha_1(1-\varepsilon)/(2-\varepsilon)}}{n^{\alpha_1}}\right)^{\varepsilon/2} \log^{O(1)} M \\ &= n^{-c} \log^{O(1)} M \end{aligned} \tag{4.13}$$

for some $c > 0$. Since this is $o(1)$, this means that $\boldsymbol{d}_I^*$ satisfies the $i = 1$ case of (A4) with constants $(R_1, R_2, \frac{1}{2}\varepsilon)$, and a.a.s. $\boldsymbol{d}_S^*$ does too. To see that the $i = 2$ case of (A4) is still satisfied, note that

$$\left(\frac{n_2}{n}\right)^{\alpha_2(1-\varepsilon/2)} \left(\frac{\Delta^2}{d}\right)^{2-\varepsilon/2} = \left(\frac{n_2}{n}\right)^{\alpha_2(1-\varepsilon)} \left(\frac{\Delta^2}{d}\right)^{2-\varepsilon} \left(\left(\frac{n_2}{n}\right)^{\alpha_2} \frac{\Delta^2}{d}\right)^{\varepsilon/2}. \tag{4.14}$$

Since we assume that (A4) is satisfied for $i = 2$ and constants (R_1, R_2, ε), it follows that

$$\left(\frac{n_2}{n}\right)^{\alpha_2(1-\varepsilon)} \left(\frac{\Delta^2}{d}\right)^{2-\varepsilon} = o(1).$$

Since we also assume that (A3) is a.a.s. satisfied for both sequences (and trivially $n/n_2 \geq 1$), it also follows that

$$\left(\left(\frac{n_2}{n}\right)^{\alpha_2} \frac{\Delta^2}{d}\right) = o(1).$$

Thus, $\boldsymbol{d}_I^*$ still predicts (A4) for $\boldsymbol{d}_S^*$ and satisfies it for $i = 2$ and constants $(R_1, R_2, \frac{1}{2}\varepsilon)$.

Now suppose that $n_2(\boldsymbol{d}_I^*) < 3\log^7 M$ and there exists a triple of constants (R_1, R_2, ε) such that condition (A4) is satisfied by $\boldsymbol{d}_I^*$ but not a.a.s. satisfied by $\boldsymbol{d}_S^*$ (or vice versa) for this set of constants and $i = 2$. We also suppose that $\boldsymbol{d}_I^*$ predicts and satisfies (A4) for $i = 1$. Then analogous equations to Equations (4.11) to (4.13) hold when α_1 is replaced with α_2. This immediately implies that $\boldsymbol{d}_I^*$ predicts and satisfies (A4) for constants $(R_1, R_2, \frac{1}{2}\varepsilon)$ for $i = 2$. If $\boldsymbol{d}_I^*$ does not predict the $i = 1$ case of (A4) for constants $(R_1, R_2, \frac{1}{2}\varepsilon)$, then by the case argued previously (where $\boldsymbol{d}_I^*$ only predicted (A4) for $i = 2$), it follows that $\boldsymbol{d}_I^*$ predicts (A4) for $\boldsymbol{d}_S^*$ and satisfies (A4) for both $i = 1$

and $i = 2$ with constants $(R_1, R_2, \frac{1}{4}\varepsilon)$.

Now suppose that $n_i(\boldsymbol{d}_I^*) < 3\log^7 M$ for both $i = 1$ and $i = 2$ and that there exists a triple of constants (R_1, R_2, ε) such that, for each value of i, condition (A4) is satisfied by $\boldsymbol{d}_I^*$ but not a.a.s. satisfied by $\boldsymbol{d}_S^*$ (or vice versa) for this set of constants. Then Equations (4.11) to (4.13) apply to both the $i = 1$ and $i = 2$ cases. Thus, $\boldsymbol{d}_I^*$ satisfies (A4) and also $\boldsymbol{d}_I^*$ predicts (A4) for $\boldsymbol{d}_S^*$ (for both $i = 1$ and $i = 2$ cases) with constants $(R_1, R_2, \frac{1}{2}\varepsilon)$. This completes the proof as outlined in the first paragraph. □

4.4 Giant components in $G[S]$

The final property of $G[S]$ that we analyse using $\boldsymbol{d}_I$ and Lemma 3.1.5 is the existence of giant components. We begin with the result given by Joos, Perarnau, Rautenbach, and Reed, rewritten slightly for our purposes. We recall several definitions from Theorem 2.2.15. For a sequence $\boldsymbol{d} := (d(1), \ldots, d(n))$, define

$$\widetilde{M}(\boldsymbol{d}) = \sum_{i \in [n], d(i) \neq 2} d(i). \tag{4.15}$$

For sequences such that $d(1) \leq d(2) \leq \cdots \leq d(n)$, define the following quantities:

$$\begin{aligned} j_{\boldsymbol{d}} &= \min\left(\left\{j : j \in [n] \text{ and } \sum_{i=1}^{j} d(i)(d(i)-2) > 0\right\} \cup \{n\}\right), \\ R(\boldsymbol{d}) &= \sum_{i=j_{\boldsymbol{d}}}^{n} d(i). \end{aligned} \tag{4.16}$$

We restrict ourselves to analysing non-negative sequences of integers. For convenience, we also want our sequences to be ordered in non-decreasing order. Let $\mathcal{D}$ be the set of all such sequences, and let $\mathcal{D}_n$ be the set of all such sequences with exactly n elements.

Here we discuss this result more explicitly in terms of sequences of sequences, and we borrow some extra definitions from Joos et al. [83] to do so. Call a sequence of sequences $(\boldsymbol{d})_{n\geq 1}$ *well-behaved* if $\widetilde{M}(\boldsymbol{d}) \to \infty$ as $n \to \infty$. Call $(\boldsymbol{d})_{n\geq 1}$ *lower bounded* if $R(\boldsymbol{d}) = \Omega(\widetilde{M}(\boldsymbol{d}))$. Call $(\boldsymbol{d})_{n\geq 1}$ *upper bounded* if $R(\boldsymbol{d}) = o(\widetilde{M}(\boldsymbol{d}))$. In most contexts, we omit the dependence on n from the sequence, and thus the statement "$\boldsymbol{d}$ is well-behaved" means that the sequence of sequences $(\boldsymbol{d})_{n\geq 1}$ is well-behaved. Here we give a version of the result by Joos et al. [83] in terms of sequences of sequences, which we use in this section.

Theorem 4.4.1. ([83], Theorem 3) Let $\boldsymbol{d} \in \mathcal{D}_n$ for each $n \geq 1$, and suppose each $\boldsymbol{d}$ has minimum degree of at least 1.

(a) If $\boldsymbol{d}$ is well-behaved and lower bounded, there is a $\gamma > 0$ such that the probability that $\mathcal{G}(\boldsymbol{d})$ has a component of order at least γn is $1 - o(1)$.

(b) If $\boldsymbol{d}$ is well-behaved and upper bounded, then for every $\gamma > 0$ the probability that $\mathcal{G}(\boldsymbol{d})$ has a component of order at least γn is $1 - o(1)$.

(c) If $\boldsymbol{d}$ is either not well-behaved or neither upper bounded nor lower bounded, then for every sufficiently small positive γ, there is a $0 < \delta < 1$ such that there are both arbitrarily large n

for which the probability that $\mathcal{G}(\boldsymbol{d})$ has a component of order at least γn is at least δ, and arbitrarily large n for which the probability that $\mathcal{G}(\boldsymbol{d})$ has a component of order at least γn is at most $1-\delta$.

This completely characterises the existence of giant components in random graphs with a fixed degree sequence. Using Theorem 4.4.1 we can predict the a.a.s. existence of a giant component in G for all well-behaved degree sequences. We use this theorem as a black box to show that the a.a.s. existence or non-existence of a giant component in $G[S]$ can be predicted by applying Theorem 4.4.1 to $\boldsymbol{d}_I^*$. We show in Lemma 4.4.11 that $\boldsymbol{d}_I^*$ is well-behaved, and $\boldsymbol{d}_S^*$ is a.a.s. well-behaved. Then we prove, under our standard assumptions on $(\boldsymbol{d}, S)$, that the induced graph $G[S]$ a.a.s. contains a giant component if and only if the degree sequence $\boldsymbol{d}_I^*$ is lower bounded.

When $\boldsymbol{d}_I$ is graphical, these previous two statements combine to show that $G[S]$ a.a.s. has a giant component if and only if a uniformly random graph with degree sequence $\boldsymbol{d}_I^*$ a.a.s. contains a giant component (if $\boldsymbol{d}_I$ is not graphical, the same statement can be made with $\boldsymbol{d}_I^f$ in its place). The reason for using $\boldsymbol{d}_I^*$ here is simply because Theorem 4.4.1 requires a minimum degree of 1. By Lemma 3.2.11 the values of $|S|$, $|S| - n_0(\boldsymbol{d}_I)$, and $|S| - n_0(\boldsymbol{d}_S)$ are a.a.s. all within a constant factor of each other. Thus, a giant component of $G(\boldsymbol{d}_S^*)$ is a.a.s. still a giant component of $G(\boldsymbol{d}_S)$ and vice versa. Also note that for a sequence $\boldsymbol{d}$ ordered in non-decreasing order, $R(\boldsymbol{d}) = R(\boldsymbol{d}^*)$, so we often simply refer to $\boldsymbol{d}_S$ and $\boldsymbol{d}_I$, rather than $\boldsymbol{d}_S^*$ and $\boldsymbol{d}_I^*$.

Theorem 4.4.2. Let $(\boldsymbol{d}(n))_{n\geq 1}$ be a sequence of sequences $\boldsymbol{d} = \boldsymbol{d}(n)$ such that $\boldsymbol{d} \in \mathcal{D}_n$ for each n, and $\boldsymbol{d}$ has minimum degree at least 1 and maximum degree at most $\frac{\sqrt{M}}{\log^7 M}$. Let $S \subset [n]$ be a subset for each n such that $d(S) = \Theta(M)$ and $d(\overline{S}) = \Theta(M)$. Then $\boldsymbol{d}_I$ is well-behaved, and a.a.s. $\boldsymbol{d}_S$ is well-behaved. Furthermore, a.a.s. $\boldsymbol{d}_S$ is upper bounded if and only if $\boldsymbol{d}_I$ is upper bounded, and $\boldsymbol{d}_S$ is lower bounded if and only if $\boldsymbol{d}_I$ is lower bounded. That is, $G[S]$ a.a.s. contains a component with at least γn vertices, for some constant $\gamma > 0$, if and only if $R(\boldsymbol{d}_I') \geq \varepsilon M$ for some constant $\varepsilon > 0$.

We wish to emphasise the distinction between $M(\boldsymbol{d})$, the total degree of $\boldsymbol{d}$, and $\widetilde{M}(\boldsymbol{d})$, as defined in Equation (4.16), which omits degree 2 vertices. It is entirely possible that $\widetilde{M}(\boldsymbol{d}) = o(M(\boldsymbol{d}))$, and for some such degree sequences it happens that $R(\boldsymbol{d}) = o(M(\boldsymbol{d}))$ and $\mathcal{G}(\boldsymbol{d})$ a.a.s. has a giant component. For our purposes, this is not an issue, and the distinction between $M(\boldsymbol{d})$ and $\widetilde{M}(\boldsymbol{d})$ is purely nominal. We show that a.a.s. $\widetilde{M}(\boldsymbol{d}_S) = \Theta(M)$ under our standard assumptions on $(\boldsymbol{d}, S)$; see Lemma 4.4.11 for the formal discussion of this. In terms of the following proofs, this result on $\widetilde{M}(\boldsymbol{d}_S)$ means that to prove Theorem 4.4.2 it is sufficient to show that $R(\boldsymbol{d}_S^*)$ a.a.s. lies within a window of size $o(M)$ around $R(\boldsymbol{d}_I')$. We rely heavily on this observation, as our previous concentration results such as Lemma 3.1.5 are all in terms of M, rather than $\widetilde{M}(\boldsymbol{d}_S)$ or $\widetilde{M}(\boldsymbol{d}_I^*)$.

A notable consequence of Theorem 4.4.2 is the following observation.

Observation 4.4.3. Suppose $\boldsymbol{d}$ is a graphical sequence and S is a subset of $[n]$ such that the conditions of Theorem 4.4.2 are satisfied and $M(\boldsymbol{d}) = \omega(|S|)$. Then $G[S]$ a.a.s. contains a giant component.

This observation is a consequence of Propositions 2.2.16 and 3.1.1. For more details, we direct the reader to Claim 7.2.3, where a similar result is proved and the argument is given in more detail.

Proof of Theorem 4.4.2

Now we show that slight perturbations to a degree sequence do not affect the asymptotically almost sure existence of a giant component. The following lemma formalises this in terms of $\widetilde{M}(\boldsymbol{d})$ and $R(\boldsymbol{d})$. The remaining lemmas in this section then argue that the concentration windows given in Lemma 3.1.5 are sufficiently tight to predict the asymptotically almost sure existence of a giant component in $G[S]$ by checking whether $\boldsymbol{d}_I$ satisfies Theorem 4.4.1. Note that in the following lemma, we do not require that $n(\boldsymbol{d}_1)$ and $n(\boldsymbol{d}_2)$ (recall that $n(\cdot)$ is the number of elements in a sequence) are equal, it is sufficient to assume that they are asymptotically equal.

Lemma 4.4.4. Let $\boldsymbol{d}_1, \boldsymbol{d}_2 \in \mathcal{D}$, and define $\widetilde{M}_1 = \widetilde{M}(\boldsymbol{d}_1)$, $\widetilde{M}_2 = \widetilde{M}(\boldsymbol{d}_2)$, $R_1 = R(\boldsymbol{d}_1)$, and $R_2 = R(\boldsymbol{d}_2)$. Suppose that $n(\boldsymbol{d}_1) \sim n(\boldsymbol{d}_2)$, $\widetilde{M}_1 \sim \widetilde{M}_2$, and that $|R_1 - R_2| = o(\widetilde{M}_1)$. Then $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ are either both well-behaved or both not. Furthermore, if they are well-behaved, then $\boldsymbol{d}_1$ is upper bounded if and only if $\boldsymbol{d}_2$ is upper bounded, and $\boldsymbol{d}_1$ is lower bounded if and only if $\boldsymbol{d}_2$ is lower bounded.

Proof. First note that $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ are either both well-behaved or both not: since $\widetilde{M}_2 \sim \widetilde{M}_1$, it follows immediately that $\widetilde{M}_1 \to \infty$ if and only if $\widetilde{M}_2 \to \infty$. Thus, $\boldsymbol{d}_1$ is well-behaved if and only if $\boldsymbol{d}_2$ is well-behaved. Now suppose both $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ are well-behaved. Suppose $\boldsymbol{d}_1$ is upper bounded. Then for every $\varepsilon > 0$, there is an n_ε such that $R_1 \leq \varepsilon \widetilde{M}_1$ for all $n > n_\varepsilon$. By the assumptions of this lemma, this implies for all $n > n_\varepsilon$ that

$$R_2 \leq R_1 + |R_1 - R_2| \leq \varepsilon \widetilde{M}_1 + o(\widetilde{M}_1) = \varepsilon \widetilde{M}_2(1 + o(1)) + o(\widetilde{M}_2) \leq \varepsilon(1 + o(1))\widetilde{M}_2.$$

Since this holds for every $\varepsilon > 0$, this implies that $\boldsymbol{d}_2$ is also upper bounded. Again, since this argument is symmetric in $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$, the converse follows immediately. Therefore, $\boldsymbol{d}_1$ is upper bounded if and only if $\boldsymbol{d}_2$ is upper bounded.

Similarly, without loss of generality suppose $\boldsymbol{d}_1$ is lower bounded. That is, for some $\beta > 0$, there exists an n_β such that for all $n > n_\beta$, $R_1 \geq \beta \widetilde{M}_1$. The assumptions on $\boldsymbol{d}_2$ then imply that, for all $n \geq n_\beta$,

$$R_2 \geq R_1 - |R_1 - R_2| \geq \beta \widetilde{M}_1 - o(\widetilde{M}_1) \geq \beta \widetilde{M}_2(1 - o(1)).$$

Thus, there exists some n'_β such that $R_2 \geq \frac{1}{2}\beta \widetilde{M}_2$ for all $n > n'_\beta$, and therefore $\boldsymbol{d}_2$ is also lower bounded. Again by symmetry the converse follows, and $\boldsymbol{d}_1$ is lower bounded if and only if $\boldsymbol{d}_2$ is lower bounded. This concludes the proof. □

The next step is to argue that $\boldsymbol{d}_S^*$ and $\boldsymbol{d}_I^*$ a.a.s. satisfy the conditions of Lemma 4.4.4. Comparing $\boldsymbol{d}_S^*$ and $\boldsymbol{d}_I^*$ is not too different to comparing $\boldsymbol{d}_S$ and $\boldsymbol{d}_I$: it follows from the definition of $M(\cdot)$ and $R(\cdot)$ that $M(\boldsymbol{d}^*) = M(\boldsymbol{d})$ and $R(\boldsymbol{d}^*) = R(\boldsymbol{d})$ for all sequences $\boldsymbol{d}$, and Lemma 3.2.11 implies that a.a.s. $n(\boldsymbol{d}_I^*) \sim n(\boldsymbol{d}_S^*)$. Thus, a.a.s. Lemma 4.4.4 applies to $\boldsymbol{d}_I^*$ and $\boldsymbol{d}_S^*$ if and only if it a.a.s. applies to $\boldsymbol{d}_I$ and $\boldsymbol{d}_S$.

Define $M_I = M(\boldsymbol{d}_I)$ and $M_S = M(\boldsymbol{d}_S)$, and analogously define R_I and R_S. Lemma 3.3.1 states that $M_I \sim \gamma^2 M$ (recall $\gamma = d(S)/M$), and also that a.a.s. $M_S \sim M_I$. Next we give conditions on two arbitrary ordered sequences $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ sufficient to ensure that $|R(\boldsymbol{d}_1) - R(\boldsymbol{d}_2)| = o(M(\boldsymbol{d}_1))$. Finally, we use Lemma 3.1.5 to show that a.a.s. $\boldsymbol{d}_S^*$ and $\boldsymbol{d}_I^*$ satisfy these conditions.

The next lemma gives a convenient bound on $|R(\boldsymbol{d}_1) - R(\boldsymbol{d}_2)|$ for two arbitrary (ordered) n-element sequences. Recall from (4.16) that

$$j_{\boldsymbol{d}} = \min\left(\left\{j : j \in [n] \text{ and } \sum_{i=1}^{j} d(i)(d(i)-2) > 0\right\} \cup \{n\}\right).$$

Lemma 4.4.5. Let $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ be two n-element sequences ordered in non-decreasing order, that is, $\boldsymbol{d}_1, \boldsymbol{d}_2 \in \mathcal{D}_n$. Suppose that $j_2 := j_{\boldsymbol{d}_2} \leq j_{\boldsymbol{d}_1} =: j_1$. Then

$$|R(\boldsymbol{d}_2) - R(\boldsymbol{d}_1)| \leq \left|\sum_{i=j_1}^{n} [d_2(i) - d_1(i)]\right| + \left|\sum_{i=1}^{j_1} d_2(i)(d_2(i)-2)\right|.$$

We emphasise here that the j_1 index in both summations is necessary. If the index in the first sum was replaced with j_1+1, then the claim would not hold for all pairs of sequences: for example, if $\boldsymbol{d}_1$ contains only zeroes and $\boldsymbol{d}_2$ has exactly one element equal to 2 with the rest equal to 0, then the modified claim does not hold. On the other hand, if the claim was modified to replace the index in the second summation with $j_1 - 1$, then there exist sequences $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ such that the left hand side is unbounded but the right hand side of the modified claim would be equal to 1. For example, suppose $\boldsymbol{d}_1 = (1, \ldots, 1, 3, \ldots, 3, k, k)$ such that $j_1 = n$ and $\sum_{i=1}^{n-1} d_1(i)(d_1(i)-2) = 0$. Then define $\boldsymbol{d}_2$ such that $\boldsymbol{d}_2(1) = 0$, and $d_2(i) = d_1(i)$ for all $i \geq 1$. Then $R(\boldsymbol{d}_2) = 2k$ but $\sum_{i=j_1}^{n} [d_2(i) - d_1(i)] = 0$ and $\sum_{i=1}^{j_1-1} d_2(i)(d_2(i)-2) = 1$.

Proof of Lemma 4.4.5. Note that if $j_2 = j_1$ then $R(\boldsymbol{d}_2) - R(\boldsymbol{d}_1) = \sum_{i=j_1}^{n} [d_2(i) - d_1(i)]$ and the lemma is trivially true. Thus, for the remainder of this proof, we assume that $j_2 < j_1$. Since $j_1 \leq n$, this immediately implies that $j_2 < n$. Thus it follows that $\sum_{i=1}^{j_2} d_2(i)(d_2(i)-2) > 0$, which in term means that $d_2(j_2) \geq 3$ and therefore $d_2(i) \geq 3$ for all $i \geq j_2$. We use these facts later in the proof.

The difference between $R(\boldsymbol{d}_2)$ and $R(\boldsymbol{d}_1)$ is

$$\begin{aligned} R(\boldsymbol{d}_2) - R(\boldsymbol{d}_1) &= \sum_{i=j_2}^{n} d_2(i) - \sum_{i=j_1}^{n} d_1(i) \\ &= \sum_{i=j_1}^{n} [d_2(i) - d_1(i)] + \sum_{i=j_2}^{j_1-1} d_2(i). \end{aligned} \tag{4.17}$$

Now we bound the second summation on the right hand side of Equation (4.17). Note that $x \leq x(x-2)$ for $x \geq 3$ and recall that $d_2(i) \geq 3$ for all $i \geq j_2$. Thus, we can employ the following bound:

$$\sum_{i=j_2}^{j_1-1} d_2(i) \leq \sum_{i=j_2}^{j_1-1} d_2(i)(d_2(i)-2) = \sum_{i=1}^{j_1-1} d_2(i)(d_2(i)-2) - \sum_{i=1}^{j_2-1} d_2(i)(d_2(i)-2). \tag{4.18}$$

By definition of j_2, we know that $\sum_{i=1}^{j_2-1} d_2(i)(d_2(i)-2) \in (-d_2(j_2)\,(d_2(j_2)-2)\,, 0]$. Since $j_2 < j_1$ and $d(j_2) \geq 3$, we also know that $d_2(j_2)(d_2(j_2)-2) \leq d_2(j_1)(d_2(j_1)-2)$, as $\boldsymbol{d}_2$ is ordered in

non-decreasing order. Thus, it follows from Equation (4.18) that

$$\sum_{i=j_2}^{j_1-1} d_2(i) \leq \sum_{i=1}^{j_1-1} d_2(i)(d_2(i)-2) + d_2(j_2)\,(d_2(j_2)-2) \leq \sum_{i=1}^{j_1} d_2(i)(d_2(i)-2). \tag{4.19}$$

Taking the absolute value of both sides of Equation (4.17), applying the triangle inequality, and then applying the bound given in (4.19) completes the proof. □

We next analyse $|R(\boldsymbol{d}_1)-R(\boldsymbol{d}_2')|$ where $\boldsymbol{d}_2$ is obtained from $\boldsymbol{d}_1$ by changing some of its elements. Note that $\boldsymbol{d}_2$ may not be in non-decreasing order after some of its elements are altered, which is why we compare $R(\boldsymbol{d}_1)$ to $R(\boldsymbol{d}_2')$. In the next lemma, we exclusively consider alterations that do not increase the value of $j_{\boldsymbol{d}}$. We then generalise to arbitrary alterations in Corollary 4.4.7 using a two-step process. It is also true that the types of alterations described in the following lemma do not decrease the value of $R(\boldsymbol{d})$, but as it is not necessary for our purposes we do not prove this here.

Lemma 4.4.6. Let $\boldsymbol{d}_1$ be an n-element sequence ordered in non-decreasing order and let $K \subset [n]$. Let $\boldsymbol{d}_2$ be an n-element sequence such that $d_1(i) = d_2(i)$ for all $i \notin K$ and for all $k \in K$, one of the following is true:

(a) $d_1(k) \geq 1$, $d_2(k) > d_1(k)$,

(b) $d_1(k) = 0$, $d_2(k) > 1$,

(c) $d_1(k) = 1$, $d_2(k) = 0$.

Then $j_{\boldsymbol{d}_2'} \leq j_{\boldsymbol{d}_1}$, and

$$|R(\boldsymbol{d}_1) - R(\boldsymbol{d}_2')| \leq \sum_{k\in K} \max\{d_1(k)^2, d_2(k)^2\} + \Delta(\boldsymbol{d}_1)^2.$$

Proof. We first prove that $j_{\boldsymbol{d}_2'} \leq j_{\boldsymbol{d}_1}$. Similarly to Lemma 4.4.5 define $j_1 := j_{\boldsymbol{d}_1}$ and $j_2 := j_{\boldsymbol{d}_2'}$. If $j_1 = n$ then the claim is trivially true. Thus, we may suppose that $j_1 < n$. By definition of j_1, this implies that $\sum_{i=1}^{j_1} d_1(i)(d_1(i)-2) > 0$. Thus it follows that $d_1(j_1) \geq 3$, and hence $d_2(j_1) \geq 3$, and thus both sequences have maximum degree at least 3. We now prove that $j_2 \leq j_1$ for two such sequences $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ that differ in a single element k - that is, in the case where $K = \{k\}$ - such that one of conditions (a), (b), or (c) are satisfied. Then the claim follows for $|K| > 1$ by repeatedly applying the claim for $|K| = 1$. Define $\sigma \in S_n$ such that $d_2(i) = d_2'(\sigma(i))$ for all $i \leq n$. We assume, as always, that σ maintains the relative order of elements where the values of $\boldsymbol{d}$ are equal.

Observe that if $k > j_1$, then since $d_1(k) \geq 3$ it must follow that $d_2(k) > d_1(k)$ and therefore $\sigma(k) > j_1$. Thus, the elements with index at most j_1 are identical between both sequences, and thus $j_2 = j_1$. If $k \leq j_1$ and $\sigma(k) > j_1$, then it follows that $d_2'(i) = d_1(i+1)$ for all $i \in [k, \sigma(k))$. That is, all the elements between k and $\sigma(k)$ get "pushed back" by one index when the sequence

is arranged in non-decreasing order. Thus, it follows that

$$\begin{aligned}\sum_{i=1}^{j_1} d_2'(i)(d_2'(i)-2) = \sum_{i=1}^{j_1} d_1(i)(d_1(i)-2) &+ \mathbb{1}_{\{\sigma(k)\le j_1\}} d_2(k)(d_2(k)-2) \\ &- \mathbb{1}_{\{k\le j_1\}} d_1(k)(d_1(k)-2) \\ &+ \mathbb{1}_{\{\sigma(k)>j_1\ge k\}} d_2(j_1+1)(d_2(j_1+1)-2).\end{aligned} \tag{4.20}$$

Since we assume that $j_1 < n$, it follows that the summation on the right hand side is positive. Thus, it only needs to be shown that the remaining terms on the right hand side are non-negative overall, and then it follows that $j_2 \le j_1$. We analyse based on the three cases stated in the lemma. In case (a), it follows that $d_1(k) < d_2(k)$ and $d_1(k) \le d_1(j_1+1)$ if $k \le j_1$, and since $d_2(k) \ne 1$ the claim follows in this case. Case (b) follows by identical reasoning. In case (c), it follows that the three terms sum to either 0 or 1, depending on whether $k \le j_1$. Thus, it follows that $j_2 \le j_1$ in all cases. This completes the proof of the first claim in the case where $|K| = 1$. By iteratively applying this result the claim holds for arbitrary $|K|$.

Now we prove the second claim of the lemma. By the definition of $\boldsymbol{d}_2$ and the first part of the lemma, it follows that $j_2 \le j_1$. The second claim is immediately true if either sequence has maximum element at most 2: if $\Delta(\boldsymbol{d}_1) \le 2$, then $R(\boldsymbol{d}_1) = d_1(n)$ and $R(\boldsymbol{d}_2) \le \sum_{k\in K} d_2(k) + d_1(n)$, and if $\Delta(\boldsymbol{d}_2) \le 2$ then it follows that $\Delta(\boldsymbol{d}_1) \le 2$. Thus, we may assume that each sequence contains at least one element that is at least 3. Therefore, $d_1(j_1) \ge 3$ and $d_2'(j_2) \ge 3$. If $j_2 = j_1 = n$, then $R(\boldsymbol{d}_2') - R(\boldsymbol{d}_1) = d_2'(n) - d_1(n)$ and the claim of the lemma follows. Thus, we may assume that $j_2 < n$. This implies that $\sum_{i=1}^{j_2} d_2'(i)(d_2'(i)-2) > 0$. Recall from the first part of the proof that $j_2 \le j_1$. Thus, it follows from Lemma 4.4.5 that

$$|R(\boldsymbol{d}_2') - R(\boldsymbol{d}_1)| \le \left|\sum_{i=j_1}^{n} (d_2'(i) - d_1(i))\right| + \left|\sum_{i=1}^{j_1} d_2'(i)(d_2'(i)-2)\right|. \tag{4.21}$$

We now examine each sum on the right hand side individually; we begin with the left sum. Since $d_1(j_1) \ge 3$, we know that $d_1(i) > 1$ for all $i \ge j_1$. Thus, for all i such that $d_1(i) > 1$, it follows from (a) that $d_2'(i) \ge d_1(i)$. This implies that the first summation in (4.21) is non-negative. Recall from earlier in the proof the definition of $\sigma \in S_n$, the permutation such that $d_2(i) = d_2'(\sigma(i))$ for all $i \le n$. Consider the case that $\sigma(i) \ge j_1$ and $i \notin K$ (and thus $d_2(i) = d_1(i)$). Then it follows that $\sigma(i) \le i$, since $d_2(j) \ge d_1(j)$ for all $j \ge j_1$. In this case, the two terms $d_2'(\sigma(i))$ and $d_1(i)$ are both present in the summation with opposite signs and cancel out. Thus, the summation can be re-expressed as

$$\begin{aligned}\sum_{i=j_1}^{n} (d_2'(i) - d_1(i)) &= \sum_{\sigma(i)\ge j_1} d_2(i) - \sum_{i\ge j_1} d_1(i) \\ &= \sum_{\sigma(i)\ge j_1,\ i\in K} d_2(i) - \sum_{i\ge j_1,\ i\in K} d_1(i).\end{aligned}$$

To bound this from above, note that $d_1(i) \ge 0$ for all $i \le n$, and thus the whole expression is bounded from above by $\sum_{k\in K} d_2(k)$. To bound this expression from below, recall that $d_2'(i) \ge$

$d_1(i)$ for all all $i \geq j_1$. Thus, the whole summation is bounded from below by 0. That is,

$$0 \leq \sum_{i=j_1}^{n} [d_2'(i) - d_1(i)] \leq \sum_{k \in K} d_2(k). \tag{4.22}$$

Now we consider the second summation in (4.21). For an arbitrary sequence $\boldsymbol{d}$, define

$$F(\boldsymbol{d}) = \sum_{i=1}^{j_1} d'(i)(d'(i) - 2);$$

note that the summation index j_1 does not depend on the sequence $\boldsymbol{d}$. First consider the case where $K = \{k\}$, that is, $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ differ on a single element. We apply (4.20) to bound $F(\boldsymbol{d}_2') - F(\boldsymbol{d}_1)$ by considering the indicator functions on the right hand side. As shown in the proof of the first claim (just after (4.20)), $F(\boldsymbol{d}_2') - F(\boldsymbol{d}_1) \geq 0$. Note that if $\sigma(k) > j_1 \geq k$, then $d_2(j_1 + 1) \leq d_2(k)$ and $d_2(j_1 + 1)(d_2(j_1 + 1) - 2) \leq d_2(k)(d_2(k) - 2)$. Also note that $d_1(k)(d_1(k) - 2) \geq 0$ unless $d_1(k) = 1$ by (a) and (b). Therefore, it follows from (4.20) that

$$0 \leq \sum_{i=1}^{j_1} d_2'(i)(d_2'(i) - 2) - \sum_{i=1}^{j_1} d_1(i)(d_1(i) - 2) \leq d_2(k)(d_2(k) - 2) + \mathbb{1}_{\{d_1(k)=1\}}. \tag{4.23}$$

Now for $|K| \geq 1$, we bound the second summation in (4.21) by changing $\boldsymbol{d}_1$ into $\boldsymbol{d}_2$ stepwise. Let $\{\boldsymbol{h}_j\}_{j \in \{0,\ldots,|K|\}}$ define a set of sequences such that $\boldsymbol{h}_0 = \boldsymbol{d}_1$ and for each $j \geq 1$ there is a unique $k_j \in K$ so that $h_{j-1}(k_j) = d_1(k_j)$, $h_j(k_j) = d_2(k_j)$, and $h_j(i) = h_{j-1}(i)$ for all $i \neq k_j$. That is, between each sequence the value of exactly one element with index in K is modified; thus it follows that $\boldsymbol{h}_{|K|} = \boldsymbol{d}_2$. Then a telescoping sum and $|K|$ applications of (4.23) gives

$$\begin{aligned} F(\boldsymbol{d}_2) &= F(\boldsymbol{d}_2) - F(\boldsymbol{h}_{|K|-1}) + F(\boldsymbol{h}_{|K|-1}) + \cdots + F(\boldsymbol{h}_1) - F(\boldsymbol{d}_1) + F(\boldsymbol{d}_1) \\ &\leq \sum_{k \in K} d_2(k)(d_2(k) - 2) + |\{k \in K \mid d_1(k) = 1\}| + \sum_{i=1}^{j_1} d_1(i)(d_1(i) - 2). \end{aligned}$$

Therefore, from (4.21), (4.22), and the definition of j_1, it follows that

$$|R(\boldsymbol{d}_2') - R(\boldsymbol{d}_1)| \leq \sum_{k \in K} d_2(k) + \sum_{k \in K} d_2(k)(d_2(k) - 2) + \sum_{i=1}^{j_1} d_1(i)(d_1(i) - 2) + |\{k \in K \mid d_1(k) = 1\}|,$$

noting that both sums on the right hand side of (4.21) are non-negative ((4.22) implies that the first sum is non-negative, and Equation (4.20) and the discussion thereafter implies that the second sum is non-negative). By definition we know that $\sum_{i=1}^{j_1} d_1(i)(d_1(i) - 2) \leq d_1(j_1)(d_1(j_1) - 2) \leq \Delta(\boldsymbol{d}_1)^2$. For the remaining terms, note by the definition of $\boldsymbol{d}_2$ that if $k \in K$ and $d_2(k) \geq 1$, then $d_2(k) > 1$, and thus $d_2(k)(d_2(k) - 1) + 1 \leq d_2(k)^2$. If $d_2(k) = 0$, then $d_1(k) = 1$, and thus $d_2(k)(d_2(k) - 2) + 1 = 1 = d_1(k)^2$. Therefore, it follows that

$$|R(\boldsymbol{d}_2') - R(\boldsymbol{d}_1)| \leq \sum_{k \in K} \max\{d_1(k)^2, d_2(k)^2\} + \Delta(\boldsymbol{d}_1)^2.$$

This completes the proof. □

We can apply the above lemma twice to give a more general bound when $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ differ by arbitrary substitutions.

Corollary 4.4.7. Let $\boldsymbol{d}_1$ be an n-element sequence ordered in non-decreasing order. Let $K \subset [n]$, and suppose $\boldsymbol{d}_2$ is a sequence such that $d_1(i) = d_2(i)$ for all $i \notin K$. Then

$$|R(\boldsymbol{d}_1) - R(\boldsymbol{d}_2')| \leq \sum_{k \in K} \max\{d_1(k)^2, d_2(k)^2\} + 2\Delta(\boldsymbol{d}_1)^2.$$

Proof. We can assume that $\Delta(\boldsymbol{d}_1) \geq 1$, otherwise Lemma 4.4.6 immediately gives the desired result. Without loss of generality, suppose K is minimal in the sense that $d_1(i) \neq d_2(i)$ for all $i \in K$. Partition K into K_1 and K_2 such that K_1 is the subset of $[n]$ where one of the following is satisfied:

(a) $d_1(i) \geq 1$, $d_2(i) > d_1(i)$,

(b) $d_1(i) = 0$, $d_2(i) > 1$,

(c) $d_1(i) = 1$, $d_2(i) = 0$.

Then it follows that K_2 is the subset of n such that one of the following is satisfied:

(a′) $d_1(i) > 1$, $d_2(i) < d_1(i)$,

(b′) $d_1(i) = 0$, $d_2(i) = 1$.

Define $\boldsymbol{k}$ to be the sequence such that $k(i) = d_1(i)$ for all $i \notin K_2$, and $k(i) = d_2(i)$ for all $i \in K_2$. Then observe that $\boldsymbol{d}_1$ and $\boldsymbol{k}$ satisfy the conditions of Lemma 4.4.6 (with $\boldsymbol{k}$ confusingly playing the role of $\boldsymbol{d}_1$ and K_2 playing the role of K in the lemma statement). Thus, applying Lemma 4.4.6 it follows that

$$|R(\boldsymbol{d}_1) - R(\boldsymbol{k}')| \leq \sum_{k \in K_2} \max\{d_1(k)^2, d_2(k)^2\} + \Delta(\boldsymbol{k})^2.$$

The sequences $\boldsymbol{k}$ and $\boldsymbol{d}_2$ also satisfy the conditions of Lemma 4.4.6 (with $\boldsymbol{k}$ again playing the role of $\boldsymbol{d}_1$ and K_1 playing the role of K in the lemma statement). Thus, it follows that

$$|R(\boldsymbol{d}_2') - R(\boldsymbol{k}')| \leq \sum_{k \in K_1} \max\{d_1(k)^2, d_2(k)^2\} + \Delta(\boldsymbol{k})^2$$

Applying the triangle inequality gives that

$$|R(\boldsymbol{d}_1) - R(\boldsymbol{d}_2')| \leq \sum_{k \in K_1} \max\{d_1(k)^2, d_2(k)^2\} + \sum_{k \in K_2} \max\{d_1(k)^2, d_2(k)^2\} + 2\Delta(\boldsymbol{k})^2.$$

Since (K_1, K_2) is a partition of K and $\Delta(k) \leq \Delta(\boldsymbol{d}_1)$, this completes the proof. □

The following two lemmas use Lemmas 4.4.5 and 4.4.6 to bound $|R(\boldsymbol{d}_1) - R(\boldsymbol{d}_2)|$ for two sequences that are "close" in some sense. In the first lemma, the ratio between the value of each term in $\boldsymbol{d}_2$ and the value of the corresponding term in $\boldsymbol{d}_1$ is close to 1. In the second lemma, the numbers of elements equal to k (for each $k \leq J$) only differ by a small amount between the two sequences. These results will then be used to compare $\boldsymbol{d}_S$ and $\boldsymbol{d}_I$.

Lemma 4.4.8. Let $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ be two n-element integer sequences ordered in non-decreasing order. Suppose that there exists a sequence $(\alpha_i)_{i\in[n]}$ of real numbers such that $\alpha := \max_i |\alpha_i| < 1$ and $\boldsymbol{d}_2 = (d_1(i)(1+\alpha_i))_{i\in[n]}$. Then

$$|R(\boldsymbol{d}_1) - R(\boldsymbol{d}_2)| < (1+7\alpha)\Delta(\boldsymbol{d}_1)^2 + 16\alpha M(\boldsymbol{d}_1).$$

Proof. Define $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ such that $k_1(i) = \lceil (1-\alpha)d_1(i)\rceil$ and $k_2(i) = \lfloor (1+\alpha)d_1(i)\rfloor$. Since $\boldsymbol{d}_1$, and $\boldsymbol{d}_2$ are integer sequences, it follows that $d_1(i), d_2(i) \in [k_1(i), k_2(i)]$ for all $i \in [n]$. It follows that all four sequences have an identical number of elements equal to 0. Since $\alpha < 1$, it also follows that $n_1(\boldsymbol{k}_2) \le n_1(\boldsymbol{d}_x) \le n_1(\boldsymbol{k}_1)$ for $x \in \{1,2\}$. Since they are all in non-decreasing order, the sequences themselves are identical for all i such that $d_1(i) < 2$. Thus, it follows that $j_{\boldsymbol{k}_2} \le j_{\boldsymbol{d}_x} \le j_{\boldsymbol{k}_1}$ and $R(\boldsymbol{k}_1) \le R(\boldsymbol{d}_x) \le R(\boldsymbol{k}_2)$ for $x \in \{1,2\}$. Therefore,

$$|R(\boldsymbol{d}_1) - R(\boldsymbol{d}_2)| \le R(\boldsymbol{k}_2) - R(\boldsymbol{k}_1).$$

We now bound $R(\boldsymbol{k}_2) - R(\boldsymbol{k}_1)$. First note that $R(\boldsymbol{k}_1) \ge 0$ and $R(\boldsymbol{k}_2) \le (1+\alpha)M(\boldsymbol{d}_1)$, and thus the claim of the lemma is trivially true if $1+\alpha \le 16\alpha$. Thus, for the rest of the proof, we assume that $\alpha < \frac{1}{15}$.

Lemma 4.4.5 implies that

$$R(\boldsymbol{k}_2) - R(\boldsymbol{k}_1) \le \sum_{i=j_{\boldsymbol{k}_1}}^{n} [k_2(i) - k_1(i)] + \sum_{i=1}^{j_{\boldsymbol{k}_1}} k_2(i)(k_2(i)-2), \tag{4.24}$$

where the absolute value signs are dropped as all summations are non-negative. To bound the first sum, the definitions of $\boldsymbol{k_1}$ and $\boldsymbol{k}_2$ immediately give that

$$\sum_{i=j_{\boldsymbol{k}_1}}^{n} [k_2(i) - k_1(i)] \le \sum_{i=j_{\boldsymbol{k}_1}}^{n} 2\alpha d_1(i) \le 2\alpha M(\boldsymbol{d}_1). \tag{4.25}$$

For the second summation in (4.24), first note that since $\alpha < \frac{1}{9}$, it follows that $k_2(i) \le (1+3\alpha)k_1(i)$ for all i. Since $x(x-2)$ achieves its minimum at $x = 1$ and is monotonically increasing for all $x \ge 1$, it follows that

$$\begin{aligned}
\sum_{i=1}^{j_{\boldsymbol{k}_1}} k_2(i)(k_2(i)-2) &\le \sum_{i=1}^{j_{\boldsymbol{k}_1}} (1+3\alpha)k_1(i)((1+3\alpha)k_1(i)-2) \\
&= \sum_{i=1}^{j_{\boldsymbol{k}_1}} k_1(i)(k_1(i)-2) + 6\alpha \sum_{i=1}^{j_{\boldsymbol{k}_1}} k_1(i)(k_1(i)-1) + 9\alpha^2 \sum_{i=1}^{j_{\boldsymbol{k}_1}} k_1(i)^2 \\
&\le \sum_{i=1}^{j_{\boldsymbol{k}_1}} k_1(i)(k_1(i)-2) + 7\alpha \sum_{i=1}^{j_{\boldsymbol{k}_1}} k_1(i)^2,
\end{aligned}$$

where the last inequality follows from the assumption that $\alpha < \frac{1}{9}$. By definition of $j_{\boldsymbol{k}_1}$,

$$\sum_{i=1}^{j_{\boldsymbol{k}_1}} k_1(i)(k_1(i)-2) \le k_1(j_{\boldsymbol{k}_1})^2 \le \Delta(\boldsymbol{k}_1)^2.$$

By adding $2\sum_{i=1}^{j_{\boldsymbol{k}_1}} k_1(i)$ to each side of the inequality, it also follows that

$$\sum_{i=1}^{j_{\boldsymbol{k}_1}} k_1(i)^2 \leq 2\sum_{i=1}^{j_{\boldsymbol{k}_1}} k_1(i) + k_1(j_{\boldsymbol{k}_1})^2 \leq 2M(\boldsymbol{k}_1) + \Delta(\boldsymbol{k}_1)^2.$$

Altogether, this implies that

$$\sum_{i=1}^{j_{\boldsymbol{k}_1}} k_1(i)(k_1(i)-2) + 7\alpha \sum_{i=1}^{j_{\boldsymbol{k}_1}} k_1(i)^2 \leq (1+7\alpha)\Delta(\boldsymbol{k}_1)^2 + 14\alpha M(\boldsymbol{k}_1). \tag{4.26}$$

Combining (4.25) and (4.26), and noting that $k_1(i) \leq d_1(i)$ for all i (and hence $M(\boldsymbol{k}_1) \leq M(\boldsymbol{d}_1)$), it follows that

$$R(\boldsymbol{k}_2) - R(\boldsymbol{k}_1) \leq (1+7\alpha)\Delta(\boldsymbol{d}_1)^2 + 16\alpha M(\boldsymbol{d}_1).$$

Since $|R(\boldsymbol{d}_1) - R(\boldsymbol{d}_2)| \leq R(\boldsymbol{k}_2) - R(\boldsymbol{k}_1)$, this completes the proof. □

Recall that $J = \log M \log\log M$. The next lemma argues that two sequences $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ that differ by small numbers of low-degree elements (elements less than or equal to J) have similar values for $R(\cdot)$. For brevity we define $M := M(\boldsymbol{d}_1)$.

Lemma 4.4.9. Let $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ be two sequences such that $n_i(\boldsymbol{d}_1) = n_i(\boldsymbol{d}_2)$ for all $i > J$ and $|n_i(\boldsymbol{d}_1) - n_i(\boldsymbol{d}_2)| \leq \frac{M}{i\log^3 M}$ for all $i \in \{1,\dots,J\}$. Then $|R(\boldsymbol{d}_1') - R(\boldsymbol{d}_2')| = O\left(\frac{M(\log\log M)^3}{\log M}\right)$.

Proof. First, note that the number of elements of the sequence that have their value changed is at most

$$\sum_{i=1}^{J} \frac{M}{i\log^3 M} = \frac{M}{\log^3 M}\sum_{i=1}^{J}\frac{1}{i} = \frac{M}{\log^3 M}(\log(J) + O(1)).$$

Also note that for all elements $i \in [n]$ that have their value changed, $\max\{d_1(i)^2, d_2(i)^2\} \leq J^2$. Let K be the set of elements that have their value changed. Applying Corollary 4.4.7 then implies that

$$\begin{aligned}
|R(\boldsymbol{d}_1) - R(\boldsymbol{d}_2)| &\leq \sum_{k\in K} \max\{d_1(i)^2, d_2(i)^2\} \leq J^2 + 2\Delta(\boldsymbol{d}_1)^2 \\
&\leq |K|J^2 + 2\Delta(\boldsymbol{d}_1)^2 \\
&\leq 2\frac{M}{\log^3 M}\log(J)J^2 + 2\Delta(\boldsymbol{d}_1)^2 \\
&= O\left(\frac{M(\log\log M)^3}{\log M}\right) + 2\Delta(\boldsymbol{d}_1)^2.
\end{aligned}$$

Since $\Delta(\boldsymbol{d}_1) \leq \sqrt{M}/\log^7 M$, this completes the proof. □

Now we combine the above two lemmas with Lemma 3.1.5 to show that the two degree sequences $\boldsymbol{d}_S$ and $\boldsymbol{d}_I$ a.a.s. satisfy the conditions of Lemma 4.4.4. For brevity, recall $R_S := R(\boldsymbol{d}_S)$, $R_I := R(\boldsymbol{d}_I)$, $M_S := M(\boldsymbol{d}_S)$, and $M_I := M(\boldsymbol{d}_I)$. Recall S_{big} is the set of all $i \in S$ such that $d(i) > J$. Recall that Y_i is the number of elements in $\boldsymbol{d}_S$ equal to i with index in S_{small}, and that y_i is the analogous term for $\boldsymbol{d}_I$.

Lemma 4.4.10. With probability $1 - o(1)$, $|R_S - R_I| = o(M_I)$.

Proof. Define an intermediate sequence $\boldsymbol{k}$, where $k(i) = d_S(i)$ for $i \in S_{\text{big}}$ and $k(i) = d_I(i)$ for $i \in S_{\text{small}}$. We prove the lemma by using $R(\boldsymbol{k})$ as an intermediate between R_I and R_S and applying the triangle inequality. By Lemma 3.1.5(a), a.a.s.

$$k(i) = d_I(i)\left(1 + O\left(\frac{1}{\sqrt{\log\log M}}\right)\right)$$

for all $i \in S_{\text{big}}$. For all $i \in S \backslash S_{\text{big}}$, by definition $k(i) = d_I(i)$. Thus, a.a.s. there exists a sequence $(\alpha_i)_{i\in[s]}$ such that $k(i) = d_I(i)(1+\alpha_i)$ for all $i \in [s]$ and $\alpha := \max_{i\in[s]} |\alpha_i| = O\left(\frac{1}{\sqrt{\log\log M}}\right)$. Thus, it follows from Lemma 4.4.8 that a.a.s.

$$|R_I - R(\boldsymbol{k})| = O\left(\frac{M_I}{\sqrt{\log\log M_I}}\right).$$

Recall from Lemma 3.3.1 that $M_I \sim \gamma^2 M$ where $\gamma = d(S)/M = \Theta(1)$. By Lemma 3.1.5(b) and Remark 3.2.10, a.a.s.

$$|Y_i - y_i| = O\left(\frac{y_i}{\log^5 M} + \log^7 M\right)$$

for all $i \le J$. By Lemma 4.4.9, this implies that a.a.s.

$$|R(\boldsymbol{k}) - R_S| = O\left(\frac{M_S(\log\log M_S)^3}{\log M_S}\right).$$

Also recall from Lemma 3.3.1 and that a.a.s. $M_S - M_I = o(M)$. Thus, by the triangle inequality, a.a.s.

$$|R_I - R_S| = \frac{M(\log\log M)^3}{\log M} + O\left(\frac{M_I}{\sqrt{\log\log M_I}}\right) = O\left(\frac{M_I}{\sqrt{\log\log M_I}}\right).$$

This completes the proof. □

As mentioned earlier, there is a small issue of counting degree 2 vertices in the definition of $M(\cdot)$. In the characterisation given in Theorem 4.4.1, degree 2 elements in a sequence do not contribute to the value of $\widetilde{M}(\boldsymbol{d})$. To apply Theorem 4.4.1, it needs to be checked that the concentration bounds previously found for M_I and M_S carry over to $\widetilde{M}(\boldsymbol{d}_I)$ and $\widetilde{M}(\boldsymbol{d}_S)$. However, we know from Lemma 3.1.5 that $\boldsymbol{d}_S$ and $\boldsymbol{d}_I$ a.a.s. have a similar number of degree 2 elements. We show in the following proof that the total degree of either sequence is not almost all from degree 2 elements. Thus, in the following lemma we show that $\boldsymbol{d}_I$ is well-behaved (as defined in Theorem 4.4.1) and thus $\boldsymbol{d}_S$ is a.a.s. also well-behaved.

Recall that $Z_j \sim \text{Bin}(j, \gamma)$, and recall from Equation (3.1) the definition of $\tilde{y}_i$, the sum of $\mathbb{P}\left(Z_{d(v)} = i\right)$ over all $v \in S_{\text{small}}$. Recall that $(y_i)_{i\le J}$, the number of degree i elements in the sequence $\boldsymbol{d}_I$ with index in S_{small}, is equal to the sequence $\{\tilde{y}_i\}_{i\le J}$ after cascade rounding.

Lemma 4.4.11. $\widetilde{M}(\boldsymbol{d}_I) = \Theta(M)$, and a.a.s. $\widetilde{M}(\boldsymbol{d}_I) - \widetilde{M}(\boldsymbol{d}_S) = o(M)$.

Proof. The first claim follows unless $M_I \sim 2y_2$; we prove that the definition of $\boldsymbol{d}_I$ means that this does not occur. Define $S_{\ge 3} := \{i \in S \mid d(i) \ge 3\}$. Consider two cases, one where $|S_{\ge 3}| = o(|S|)$

and one where $|S_{\geq 3}| = \Theta(|S|)$. In the first case, since $\boldsymbol{d}$ has minimum degree at least 1 it follows that S contains $|S|(1 - o(1))$ elements i such that $d(i) \in \{1, 2\}$. Then if $|S_{\geq 3}| = o(|S|)$, the definition of $\tilde{y}_1$ implies that

$$\tilde{y}_1 = \sum_{i \in S} \mathbb{P}\left(Z_{d(i)} = 1\right) \geq \sum_{i \in S \setminus S_{\geq 3}} \mathbb{P}\left(Z_{d(i)} = 1\right) \geq \gamma |S_1| + 2\gamma(1-\gamma)|S_2|$$
$$\geq |S|(1 - o(1)) \min\{\gamma, 2\gamma(1-\gamma)\}.$$

Since $\gamma = \Theta(1)$ and $y_2 \leq |S|$ and $y_1 \geq \tilde{y}_1 - 1$, it follows that

$$M_I - 2y_2 \geq y_1 = \Omega(y_2).$$

Thus, $M_I \nsim 2y_2$. Now consider the second case, where $S_{\geq 3} = \Theta(|S|)$. Then

$$\sum_{k \geq 3} \tilde{y}_k = \sum_{k \geq 3} \sum_{i \in S} \mathbb{P}\left(Z_{d(i)} = k\right) = \sum_{i \in S_{\geq 3}} \mathbb{P}\left(Z_{d(i)} \geq 3\right).$$

For all $i \in S_{\geq 3}$, $\mathbb{P}\left(Z_{d(i)} \geq 3\right) \geq c$ for some constant $c > 0$ independent of i. Thus, it follows that $\sum_{k \geq 3} \tilde{y}_k = \Theta(|S|)$. By similar reasoning to the previous case, this implies that $M_I \geq (2 + c)y_2$ for some $c > 0$, and thus $M_I \nsim 2y_2$. Thus, in either case, it follows that $\widetilde{M}(\boldsymbol{d}_I) = M_I - 2y_2 = \Theta(M_I) = \Theta(M)$. This proves the first claim in the lemma.

Now consider the second claim. Recall that Y_i is the number of degree i elements in $\boldsymbol{d}_S$ with index in S_{small}. By Lemma 3.1.5(a), there are a.a.s. no elements $v \in S_{\text{big}}$ such that $d_S(v) = 2$. Thus, a.a.s. the number of terms of degree 2 in $\boldsymbol{d}_S$ is exactly Y_2. By Lemma 3.1.5(b) and Remark 3.2.10, a.a.s.

$$|Y_2 - y_2| \leq |Y_2 - \mathbb{E}[Y_2]| + |\mathbb{E}[Y_2] - y_2| \leq \frac{y_2}{\log^5 M}(1 + o(1)) + \log^7 M.$$

Since $y_2 \leq \frac{1}{2} M_I$, it follows that a.a.s. $|Y_2 - y_2| \leq \frac{M_I}{\log^5 M} + \log^7 M$. By Lemma 3.3.1, this is $o(M)$, and therefore a.a.s.

$$\widetilde{M}(\boldsymbol{d}_S) - \widetilde{M}(\boldsymbol{d}_I) = M_S - M_I - Y_2 + y_2 = o(M).$$

This completes the proof. □

Since $\widetilde{M}(\boldsymbol{d}_I) = \Theta(M)$, all asymptotic results concerning M_I or M_S immediately carry over to $\widetilde{M}(\boldsymbol{d}_I)$ and $\widetilde{M}(\boldsymbol{d}_S)$. Lemmas 3.3.1, 4.4.4, 4.4.10, and 4.4.11 combine to give a short proof of Theorem 4.4.2.

Proof of Theorem 4.4.2. Lemmas 3.3.1 and 4.4.11 respectively imply that a.a.s. $|\widetilde{M}(\boldsymbol{d}_S) - \widetilde{M}(\boldsymbol{d}_I)| = o(M_I)$ and $\widetilde{M}(\boldsymbol{d}_S) = \Theta(M_I)$. This implies that a.a.s. $\boldsymbol{d}_S$ and $\boldsymbol{d}_I$ are both well-behaved. It also follows immediately that

$$|\widetilde{M}(\boldsymbol{d}_S) - \widetilde{M}(\boldsymbol{d}_I)| = o(\widetilde{M}(\boldsymbol{d}_I)).$$

Lemma 4.4.10 states that a.a.s. $|R_S - R_I| = o(M_I)$. Since $\widetilde{M}(\boldsymbol{d}_I) = \Theta(M_I)$ it follows that a.a.s.

$$|R_S - R_I| = o(\widetilde{M}(\boldsymbol{d}_I)).$$

By Lemma 4.4.4, this means that $\boldsymbol{d}_S$ is upper bounded if and only if $\boldsymbol{d}_I$ is upper bounded, and likewise for lower bounded. Therefore, $G[S]$ a.a.s. contains a giant component if and only if a.a.s. $R(\boldsymbol{d}'_S) \geq \alpha M$ for some constant $\alpha > 0$, which occurs if and only if $R(\boldsymbol{d}'_I) \geq \varepsilon M$ for some $\varepsilon > 0$. □

Chapter 5

Site percolation: random induced subgraphs

In this chapter we study the case where the subset S is chosen randomly by choosing each vertex independently with some constant probability p. The methods we apply are similar similar to the methods used in Chapters 3 and 4 to study the case where S is fixed: we define some sort of deterministic "average" sequence based on $\boldsymbol{d}$ and p and then show that the degree sequence of the induced graph is a.a.s. close to this sequence. Due to the extra randomness introduced by allowing S to vary, the definition of this average sequence is similar but slightly different to the definition of $\boldsymbol{d}_I$ given in Definition 3.1.2. The extra randomness also means that we lose some predictive power, and not all results from Chapter 4 carry over fully in this case. In this chapter, we characterise when $G[S]$ a.a.s. has a giant component, as well as when $G[S]$ is connected after deleting isolated vertices.

5.1 The site percolation model

Suppose that instead of S being a predetermined, fixed set, S is determined by taking each vertex in G with some probability p. The set S then corresponds to a random subset of $\{1, \dots, n\}$. This is more commonly known as *percolation*, more specifically site percolation of the random graph $\mathcal{G}(\boldsymbol{d})$ (this is in contrast to bond percolation, where edges are randomly deleted). For our purposes, we will always assume that $\min\{p, 1-p\} = \Theta(1)$. In this section we study our favourite graph properties on this random graph space, and show that we can still find thresholds for these properties in $G[S]$ based on looking at an "average" degree sequence (defined in Definition 5.2.1).

First we need to formally introduce the new probability space. In this chapter, the event space is the set of all possible induced subgraphs $G[S]$ for all graphs $G \in \mathcal{G}(\boldsymbol{d})$ and all subsets $S \subset [n]$. The probability that the percolated graph $G[S]$ is some particular (labelled) graph H is the probability that the correct subset S is chosen (which is simply $p^{|V(H)|}(1-p)^{n-|V(H)|}$) multiplied by the probability that some G is sampled from $\mathcal{G}(\boldsymbol{d})$ such that $G[S] = H$. More formally, if H is some graph with $V(H) \subseteq [n]$, then

$$\mathbb{P}\left(G[S] = H\right) = \frac{|\{G \in \mathcal{G}(\boldsymbol{d}) \mid H \subseteq G\}|}{|\mathcal{G}(\boldsymbol{d})|} p^{|V(H)|}(1-p)^{n-|V(H)|}.$$

The probability $\mathbb{P}(\cdot)$ and expectation $\mathbb{E}[\cdot]$ used in this chapter are with respect to this measure.

We call this model the *(site) percolated graph* $\mathcal{G}(\boldsymbol{d})$ *with survival probability* p. We define $\mathbb{E}[\cdot \mid S]$ to be the expectation conditional on the random subset being S. This conditional expectation corresponds to the definition of $\mathbb{E}[\cdot]$ used in Chapters 3 and 4, since S was predetermined. We also use many definitions from Chapter 3 but with an extra argument S when discussing things for a fixed S: for example, we write $\boldsymbol{d}_I(S)$ to mean the sequence $\boldsymbol{d}_I$ as defined in Definition 3.1.2 for a particular subset S. The sequence $\boldsymbol{d}_S$ is still defined to be the degree sequence of $G[S]$.

Throughout this chapter, we still assume that $\boldsymbol{d}$ has minimum degree at least 1 and maximum degree at most $\sqrt{M}/\log^7 M$. Naturally, we do not make assumptions about S or $d(S)$, since this is not a fixed object.

5.2 Concentration of the induced degree sequence

First we recall some relevant definitions from the previous chapters. Recall that $d(A)$, for some $A \subset [n]$, is the total degree in $\boldsymbol{d}$ of the set A, and recall that $M := M(\boldsymbol{d}) = \sum_{i=1}^{n} d(i)$ is the total degree of $\boldsymbol{d}$. Recall that $n_k(\boldsymbol{d})$ is the number of terms equal to k in some sequence $\boldsymbol{d}$. Again we define Y_i as in Lemma 3.1.5 to be the number of vertices in S_{small} such that $d_S(v) = i$. As we previously defined in Section 2.2.4, when we write "a.a.s. $a = O(b)$" (or equivalently "$\mathbb{P}(a = O(b)) = 1 - o(1)$") for two functions a and b (implicitly, $a(M)$ and $b(M)$), we mean that there exists a constant $C > 0$ such that $\mathbb{P}(|a| \le Cb) \to 1$ as $M \to \infty$ where the constant C is independent of M and bounded away from 0.

To study this model, we define an "idealised" or "average" degree sequence for this probability space as well, which we call $\boldsymbol{d}_A$. First, define

$$V_{\text{big}} = \{i \in [n] \mid d(i) > J\},$$

where $J = \log M \log\log M$ and $V_{\text{small}} = [n] \backslash V_{\text{big}}$. With this, we can also write S_{small} and S_{big} as $S \cap V_{\text{small}}$ and $S \cap V_{\text{big}}$ respectively. To allow us to use certain results from Chapters 3 and 4, we construct $\boldsymbol{d}_A$ by considering vertices in V_{big} and vertices in V_{small} separately.

Much like the definition of $\boldsymbol{d}_I$ back in Definition 3.1.2, we design $\boldsymbol{d}_A$ to represent an average degree sequence of $G[S]$ in our particular probability space. The way we define $\boldsymbol{d}_A$ is similar to how we defined $\boldsymbol{d}_I$ in Definition 3.1.2, particularly for low-degree terms: $\boldsymbol{d}_A$ contains the expected number of vertices with induced degree i under the binomial approximation, subject to cascade rounding. The treatment of high-degree terms in the definition of $\boldsymbol{d}_A$ is similar to the treatment of low-degree terms, which distinguishes it from $\boldsymbol{d}_I$.

The different treatment of high-degree vertices compared to $\boldsymbol{d}_I$ is due to the extra randomness caused by not fixing the set S. In the definition of $\boldsymbol{d}_I$, we assume that S is fixed, and so we can analyse each vertex in S_{big} precisely. In this case where the set S (and thus S_{big}) is not fixed, if V_{big} contains a wide range of degrees then some percolated graphs might have a much higher maximum degree than others. Thus, the behaviour of the set of high-degree vertices in the percolated graph $\mathcal{G}(\boldsymbol{d})$ is much less clear in this model than compared to the case where S is fixed. However, for a fixed set S and for each $v \in S_{\text{big}}$, an analogous lemma to Lemma 3.1.5 (specifically Lemma 5.2.2(c)) states that the degree of the corresponding vertex in $G[S]$ is well concentrated. We discuss the relationships between $\boldsymbol{d}_S$, $\boldsymbol{d}_A$, and $\boldsymbol{d}_I$ more in Section 5.3.1 and Remark 5.3.8.

Definition 5.2.1. Let $\boldsymbol{d} = (d(1), \dots, d(n))$ be an n-element graphical sequence ordered in non-decreasing order. Let $p \in (0,1)$ be a constant, and let $Z_j \sim \text{Bin}(j,p)$. For $k \in \{0, \dots, J\}$, define

$$\tilde{w}_k := p \sum_{i \in V_{\text{small}}} \mathbb{P}\left(Z_{d(i)} = k\right).$$

Let $(w_k)_{k=0}^J$ be the sequence $(\tilde{w}_k)_{k=0}^J$ after cascade rounding. Then $\boldsymbol{d}_A$ contains w_k elements equal to k. This makes up the first $\sum_{k \le J} w_k$ entries of $\boldsymbol{d}_A$. In addition to this, define

$$\tilde{z}_k := p \sum_{i \in V_{\text{big}}} \mathbb{P}\left(Z_{d(i)} = k\right).$$

Let $(z_k)_{k=0}^J$ be the sequence $(\tilde{z}_k)_{k=0}^J$ after cascade rounding. The sequence $\boldsymbol{d}_A$ also contains z_k more terms of degree k. These terms make up the remaining entries of $\boldsymbol{d}_A$.

The specific definition of z_k given in Definition 5.2.1 is not the only viable definition for a representative set of induced degrees of vertices in S_{big}. Rather, we just need some sequence of terms such that their total and average degree are not too probabilistically outlandish. We show (in Lemma 5.2.4 and the subsequent results that use it) that the behaviour of V_{big} in the site percolated graph is sufficiently predictable to yield useful results using this definition of $\tilde{z}_k$. Given the change to the treatment of V_{big}, one might ponder the reason for splitting the treatment of V_{small} and V_{big} and defining $\tilde{w}_k$ and $\tilde{z}_k$, rather than simply extending the definition of $\tilde{w}_k$ to all $v \in [n]$. Indeed this is possible, and might be more natural. However, as mentioned earlier, the split allows us to appeal to certain results from the previous sections, such as Lemma 3.1.5, in order to prove that $\boldsymbol{d}_A$ and $\boldsymbol{d}_S$ are similar for low-degree terms. These similarities are stated formally in Lemma 5.2.2.

As was the case with $\boldsymbol{d}_I$, we do not require that $\boldsymbol{d}_A$ is a graphical sequence to apply these results. We can define $\boldsymbol{d}_A^f$ analogously to $\boldsymbol{d}_I^f$ earlier (that is, the maximum degree element of $\boldsymbol{d}_A$ is reduced by 1 if $M(\boldsymbol{d}_A)$ odd, otherwise $\boldsymbol{d}_A^f = \boldsymbol{d}_A$); Lemma 3.1.3 then implies that $\boldsymbol{d}_A^f$ is a graphical sequence. Then one can think of a graph with degree sequence $\boldsymbol{d}_A^f$ in place of a graph with degree sequence $\boldsymbol{d}_A$ for intuition purposes.

One thing to note when defining sequences of degree sequences in the site percolation model is that indexing by $|S|$ does not make sense, since S does not have a fixed size. When comparing $\boldsymbol{d}_S$ and $\boldsymbol{d}_A$ for a fixed sequence $\boldsymbol{d}$ (and thus a fixed value of $n(\boldsymbol{d})$), this means that the two sequences do not necessarily have the same number of elements. However, this is no great concern: $\boldsymbol{d}_I^*$ also did not necessarily have the same number of elements as $\boldsymbol{d}_S^*$. Much like how Lemma 3.2.11 showed that $n(\boldsymbol{d}_I^*)$ and $n(\boldsymbol{d}_S^*)$ were a.a.s. close, we show later in Lemma 5.2.2 that the two sequences $\boldsymbol{d}_A$ and $\boldsymbol{d}_S$ have asymptotically equal numbers of terms. This is sufficient to argue that $G(\boldsymbol{d}_S)$ and $G(\boldsymbol{d}_A)$ (implicitly, $G(\boldsymbol{d}_A^f)$) have very similar properties. That is, if $\boldsymbol{d}_A$ has a particular property, then a.a.s. the site percolated graph $G[S]$ has some property.

Lemma 5.2.2. Let $\boldsymbol{d}$ be an n-element degree sequence with minimum degree at least 1 and maximum degree at most $\sqrt{M}/\log^7 M$. Let S and $G[S]$ be distributed as a site percolated $\mathcal{G}(\boldsymbol{d})$ with survival probability p for constant p. Then the following statements hold with probability $1 - o(1)$.

(a) $|S| = np\left(1 + O\left(\sqrt{\frac{\log n}{n}}\right)\right)$.

(b) $d(S) = pM\left(1 + O\left(\frac{1}{M^{1/4}}\right)\right)$.

(c) For all $v \in S_{\text{big}}$,

$$d_S(v) = pd(v)\left(1 + O\left(\frac{1}{\sqrt{\log\log M}}\right)\right).$$

(d) For all $i \leq J$,

$$|Y_i - w_i| \leq \left(\frac{w_i}{\log^5 M} + \log^7 M\right)(1 + o(1)).$$

The rough outline of the proof is as follows. A Chernoff bound shows that $|S|$ is a.a.s. equal to $np(1+o(1))$, and a basic vertex exposure martingale shows that $d(S)$ is a.a.s. equal to $pM(1+o(1))$. This proves the first two statements in the lemma. For each choice of S that satisfies parts (a) and (b) of this lemma, Lemma 3.1.5 applies. For terms in S_{big}, this means that the concentration window given in Lemma 3.1.5(a) carries over immediately for each of these "good" choices of S. This proves the third statement. For S_{small}, there is an extra level of complication. We know that the concentration result given in Lemma 3.1.5(b) applies for each fixed S, but this only yields good concentration of Y_i if $|S_j|$, the number of vertices in S with degree j in G, is sufficiently concentrated for each $j \leq J$. Thankfully, this is indeed the case, and a proof of the final statement follows from this. Because of the maximum degree constraint on $\boldsymbol{d}$, the concentration windows obtained this way are sufficiently small. Recall the notation that $a = b \pm c$ means that $a \in [b - c, b + c]$.

Proof of Lemma 5.2.2. Immediately we know that $\mathbb{E}[|S|] = np$, and linearity of expectation gives that

$$\mathbb{E}[d(S)] = \sum_{i \in [n]} d(i)\mathbb{P}(i \in S) = pM.$$

First we argue concentration of $|S|$. The Chernoff bound given in Lemma A.4 implies that

$$\mathbb{P}(||S| - \mathbb{E}[|S|]| \geq \varepsilon np) \leq 2\exp\left(-\frac{np\varepsilon^2}{3}\right).$$

Letting $\varepsilon = 3\sqrt{\log n/pn}$, it immediately follows that this probability is at most $2n^{-3}$. This completes the proof of (a). For part (b), we construct a martingale to show concentration of $d(S)$. At step i, for $i \in [n]$, reveal whether vertex i is in S. Define $M_i = d(S \cap [i])$. Then it follows immediately that $|M_i - M_{i-1}| \leq d(i)$ for all $i \leq n$ and $M_n = d(S)$. Thus, by Azuma's inequality (given in Theorem A.6),

$$\mathbb{P}(|d(S) - pM| \geq \alpha) \leq 2\exp\left(\frac{-\alpha^2}{2\sum_{i \in V(G)} d(i)^2}\right).$$

By assumption, $\Delta(\boldsymbol{d}) \leq \frac{\sqrt{M}}{\log^7 M}$, which implies that

$$\sum_{i \in V(G)} d(i)^2 \leq \Delta(\boldsymbol{d}) \sum_{i \in [n]} d(i) \leq \frac{M^{3/2}}{\log^7 M}.$$

Thus, by setting $\alpha = M^{3/4}$, this probability is at most $2M^{-3/2}$, which proves that S a.a.s. satisfies condition (b). For the remainder of this proof, we call a set S "good" if $|S| = pn\left(1 \pm 3\sqrt{\log n}/\sqrt{pn}\right)$ and $d(S) = pM\left(1 \pm 1/pM^{1/4}\right)$.

Now we focus on part (c). Let S be some arbitrary "good" subset of $[n]$, and recall that $\gamma = d(S)/M$. Then Lemma 3.1.5(a) implies that

$$\mathbb{P}\left(d_S(v) = pd(v)\left(1 \pm \frac{10}{\sqrt{\gamma \log\log M}}\right) \text{ for all } v \in S_{\text{big}} \,\middle|\, S\right) = 1 - o(1).$$

By parts (a) and (b) of this lemma, we know that S is good with probability $1 - o(1)$. Thus, the probability that there exists some $v \in S_{\text{big}}$ with induced degree that is not $pd(v)\left(1 + O\left(\frac{1}{\sqrt{\log\log M}}\right)\right)$ is at most

$$\mathbb{P}\left(S \text{ is not good}\right) + \mathbb{P}\left(d_S(v) \neq pd(v)\left(1 + O\left(\frac{1}{\sqrt{\log\log M}}\right)\right) \,\middle|\, S \text{ is good}\right)\mathbb{P}\left(S \text{ is good}\right) = o(1).$$

This proves part (c).

Finally, we focus on part (d) of the lemma. Recall that $n_j(\boldsymbol{d})$ is the number of elements of $\boldsymbol{d}$ equal to j, or equivalently (if $\boldsymbol{d}$ is graphical) the number of vertices of degree j in a graph with degree sequence $\boldsymbol{d}$. Let S_j be the set of vertices of degree j in S, for each $j \leq J$. We can express $|S_j|$ as a sum over all vertices in G with degree i:

$$|S_j| = \sum_{d(i)=j} \mathbb{1}_{\{i \in S\}},$$

where $\mathbb{1}_{\{i \in S\}}$ is an independent Bernoulli random variable with $\mathbb{P}\left(\mathbb{1}_{\{i \in S\}} = 1\right) = p$ for all $i \in [n]$. It follows from linearity of expectation that $\mathbb{E}\left[|S_j|\right] = pn_j(\boldsymbol{d})$. Since these indicators are independent random variables, McDiarmid's inequality (given in Theorem A.5) gives that

$$\mathbb{P}\left(||S_j| - pn_j(\boldsymbol{d})| > \alpha_j\right) \leq 2\exp\left(-\frac{2\alpha_j^2}{n_j(\boldsymbol{d})}\right)$$

for each $j \leq J$. Define $\alpha_j = (n_j(\boldsymbol{d}) \log\log M)^{1/2}$. Then

$$\exp\left(-\frac{2\alpha_j^2}{n_j(\boldsymbol{d})}\right) = \exp\left(-2\log\log M\right).$$

Performing the union bound over J such events implies that the probability that each S_j is within $(n_j(\boldsymbol{d}) \log\log M)^{1/2}$ of its expectation is at most

$$J\exp\left(-2\log\log M\right) = \frac{\log\log M}{\log M} = o(1).$$

Suppose for some particular j that $n_j(\boldsymbol{d}) \geq \log^{11} M$. Then, with probability $1 - o(1)$ it follows

that

$$||S_j| - pn_j(\boldsymbol{d})| \leq \alpha_j = (n_j(\boldsymbol{d}) \log\log M)^{1/2} = o\left(\frac{n_j(\boldsymbol{d})}{\log^{5.25} M}\right).$$

Now suppose that $n_j(\boldsymbol{d}) < \log^{11} M$. Then it follows that

$$||S_j| - pn_j(\boldsymbol{d})| \leq \alpha_j = (n_j(\boldsymbol{d}) \log\log M)^{1/2} = o(\log^{5.75} M).$$

Thus, it follows that a.a.s. for all $j \leq J$

$$||S_j| - pn_j(\boldsymbol{d})| = o\left(\frac{n_j(\boldsymbol{d})}{\log^{5.25} M} + \log^{5.75} M\right). \tag{5.1}$$

Suppose that S is an arbitrary good subset of $[n]$ which also satisfies the concentration bounds given in Equation (5.1) for all $j \in \{0, \ldots, J\}$. Since S is good, the conditions of Lemma 3.2.5 are met. Recall that Y_i is the number of vertices in S_{small} with induced degree i, for $i \leq J$. The expectation of Y_i, conditional on this set S, is thus given by

$$\mathbb{E}\left[Y_i \middle| S\right] = \sum_{j \leq J} |S_j| \mathbb{P}\left(Z_j = i\right) \left(1 + O\left(\frac{\Delta^2 J}{M}\right)\right).$$

Since we assume that S satisfies the concentration inequalities given in Equation (5.1), it follows that

$$\begin{aligned}
\sum_{j \leq J} |S_j| \mathbb{P}\left(Z_j = i\right) &= \sum_{j \leq J} \mathbb{P}\left(Z_j = i\right) \left[pn_j(\boldsymbol{d}) \pm \left(\frac{n_j(\boldsymbol{d})}{\log^{5.25} M} + \log^{5.75} M\right)\right] \\
&= \sum_{j \leq J} \mathbb{P}\left(Z_j = i\right) pn_j(\boldsymbol{d}) \left(1 + O\left(\frac{1}{\log^{5.25} M}\right)\right) \pm \sum_{j \leq J} \mathbb{P}\left(Z_j = i\right) \log^{5.75} M \\
&= w_i + o\left(\frac{w_i}{\log^5 M} + \log^7 M\right).
\end{aligned}$$

Since $\Delta^2 J / M = o(\log^{-12} M)$, it follows that, conditional on the aforementioned set S,

$$\mathbb{E}\left[Y_i \middle| S\right] = w_i + o\left(\frac{w_i}{\log^5 M} + \log^7 M\right). \tag{5.2}$$

Since we assume that S is good, $(\boldsymbol{d}, S)$ also satisfies the conditions of Lemma 3.1.5(b). This implies that a.a.s.

$$|Y_i - \mathbb{E}\left[Y_i \middle| S\right]| \leq \frac{\mathbb{E}\left[Y_i \middle| S\right]}{\log^5 M} + \log^7 M. \tag{5.3}$$

A combination of the bounds given in (5.2) and (5.3) and the triangle inequality implies that, conditional on the event that S is good and also satisfies the concentration bounds given in Equation (5.1), a.a.s.

$$|Y_i - w_i| \leq \left(\frac{w_i}{\log^5 M} + \log^7 M\right)(1 + o(1)).$$

Since S satisfies these conditions a.a.s., this proves part (d). □

Remark 5.2.3. Recall that $\boldsymbol{d}^*$ is the sequence $\boldsymbol{d}$ with all elements equal to 0 removed. An analogous argument to Lemma 3.2.11 implies that $n(\boldsymbol{d}_A^*) = \Theta(n)$. Since Lemma 5.2.2(a) implies that a.a.s. $n(\boldsymbol{d}_S) = |S| \sim pn$, and part (d) of the lemma states that a.a.s. $Y_0 - w_0 = o(n)$, it follows that a.a.s. $n(\boldsymbol{d}_S^*) \sim n(\boldsymbol{d}_A^*)$. That is, a.a.s. $\boldsymbol{d}_S$ and $\boldsymbol{d}_A$ both have $pn(1+o(1))$ terms, and a.a.s. $\boldsymbol{d}_S^*$ and $\boldsymbol{d}_A^*$ both have $cn(1+o(1))$ terms for some constant $c > 0$.

Next we show that the total induced degree of vertices in S_{big} is a.a.s. concentrated. Due to the nature of the site percolation model, the specific induced degrees of elements in S_{big} might vary wildly. However, Lemma 5.2.2(c) states that a.a.s. each element of S_{big} has induced degree near its expectation. Our assumption on the maximum degree of $\boldsymbol{d}$ is then sufficient to argue that $d(S_{\text{big}})$ is concentrated. Recall from Definition 5.2.1 the definition of z_k, the number of terms in $\boldsymbol{d}_A$ equal to k that correspond to vertices in V_{big}.

Lemma 5.2.4. Deterministically, $\sum_{k\geq 0} kz_k = p^2 d(V_{\text{big}}) \pm \frac{1}{2}\Delta(\boldsymbol{d})^2$. Also, if $d_S(S_{\text{big}})$ is the total induced degree of S_{big}, then a.a.s.

$$d_S(S_{\text{big}}) = p^2 d(V_{\text{big}})(1+o(1)) \pm 2d(V_{\text{big}})^{1/2}\Delta(\boldsymbol{d})^{1/2}\log\log M.$$

Proof. The first result simply follows from the definition of z_k, given in Definition 5.2.1:

$$\begin{aligned}
\sum_{k\geq 0} kz_k &= \sum_{k\geq 0} k\tilde{z}_k \pm \sum_{k\geq 0} k \\
&= \sum_{k\geq 0} kp \sum_{i\in V_{\text{big}}} \mathbb{P}\left(Z_{d(i)} = k\right) \pm \frac{1}{2}\Delta(\boldsymbol{d})^2 \\
&= p\sum_{i\in V_{\text{big}}} \sum_{k\geq 0} k\mathbb{P}\left(Z_{d(i)} = k\right) \pm \frac{1}{2}\Delta(\boldsymbol{d})^2 \\
&= p^2 d(V_{\text{big}}) \pm \frac{1}{2}\Delta(\boldsymbol{d})^2.
\end{aligned}$$

We show that $d(S_{\text{big}})$ is concentrated using a similar method to that used to prove that $d(S)$ is concentrated. Let $V_{\text{big}} = \{v_1, \ldots, v_{|V_{\text{big}}|}\}$ where $d(v_1) \leq d(v_2) \leq \cdots \leq d(v_{|V_{\text{big}}|})$. If D_t is the total degree (in G) of $\{v_1, \ldots, v_t\} \cap S$ for $t \leq |V_{\text{big}}|$, then $|D_t - D_{t-1}| \leq d(v_t)$. Noting that $\mathbb{E}\left[d(S_{\text{big}})\right] = pd(V_{\text{big}})$, Azuma's inequality implies that

$$\mathbb{P}\left(|d(S_{\text{big}}) - pd(V_{\text{big}})| > \alpha\right) \leq 2\exp\left(-\frac{\alpha^2}{2\sum_{i\in V_{\text{big}}} d(i)^2}\right) \leq 2\exp\left(-\frac{\alpha^2}{2\Delta(\boldsymbol{d})d(V_{\text{big}})}\right).$$

Choosing $\alpha = d(V_{\text{big}})^{1/2}\Delta(\boldsymbol{d})^{1/2}\log\log M$ implies that a.a.s. $d(S_{\text{big}}) = pd(V_{\text{big}}) \pm \alpha$. Lemma 5.2.2 implies that a.a.s. $d_S(v) = pd(v)(1+o(1))$ for all $v \in S_{\text{big}}$. Thus, a.a.s. $d_S(S_{\text{big}}) = \sum_{v\in S_{\text{big}}} d_S(v) = p(1+o(1))\sum_{v\in S_{\text{big}}} d(v) = p^2 d(V_{\text{big}})(1+o(1)) \pm 2d(V_{\text{big}})^{1/2}\Delta(\boldsymbol{d})^{1/2}\log\log M$. This proves the second claim. □

Combining this previous lemma with Lemma 5.2.2, it follows that the total degree of $M(\boldsymbol{d}_S)$ is a.a.s. asymptotically equal to $M(\boldsymbol{d}_A)$, which is approximately $p^2 M$.

Lemma 5.2.5. $M(\boldsymbol{d}_A) = p^2 M \pm \Delta(\boldsymbol{d})^2$ always, and a.a.s. $M(\boldsymbol{d}_S) \sim M(\boldsymbol{d}_A)$.

Proof. The first claim follows almost identically to the first claim in Lemma 5.2.4, and thus we

omit some details:

$$\begin{aligned} M(\boldsymbol{d}_A) &= \sum_{k\le J} kw_k + \sum_{k\ge 0} kz_k \\ &= \sum_{k\le J} k\tilde{w}_k + \sum_{k\ge 0} k\tilde{z}_k \pm 2\sum_{k=0}^{\Delta} k \\ &= p^2 M \pm \Delta^2. \end{aligned}$$

To prove the second claim, we apply Lemmas 5.2.2 and 5.2.4. Suppose $\boldsymbol{d}_S$ satisfies (a) – (d) of Lemma 5.2.2, which occurs with probability $1 - o(1)$. Under these conditions, it follows that

$$\begin{aligned} M(\boldsymbol{d}_S) &= \sum_{k\le J} kY_k + d_S(S_{\text{big}}) \\ &= \sum_{k\le J} kw_k + d_S(S_{\text{big}}) \pm \sum_{k\le J} k\left[\frac{w_k}{\log^5 M} + \log^7 M\right](1+o(1)) \\ &= p^2 d(V_{\text{small}}) + p^2 d(V_{\text{big}})(1+o(1)) + o(M) \\ &= p^2 M(1+o(1)). \end{aligned}$$

This proves the second claim. □

5.3 Properties of random induced subgraphs

The concentration results proved in the previous section allow us to prove results about the properties of $G[S]$ analogous to the case where S is fixed. We focus on analysing the connectivity of the percolated graph, as well as the existence of giant components in such a graph. It is noteworthy that an analogous result to Theorem 4.2.1 about the chromatic number of $G[S]$, or more specifically Lemma 3.3.4, does not immediately carry over to the site percolation model. For a naive counterexample, consider the degree sequence $\left(3,\dots,3,\frac{\sqrt{M}}{\log^7 M}\right)$. With probability p, the degree sequence $\boldsymbol{d}_S$ contains vertex n. If $n \in S$, a.a.s. this vertex has induced degree $\Theta(\sqrt{M}/\log^7 M) = \Theta(\sqrt{n}/\log^7 n)$ by Lemma 5.2.2. Recall that $D_k(\boldsymbol{d}) := D_k(\boldsymbol{d},1) = \sum_{i=0}^{k-1} d(n(\boldsymbol{d}) - i)$. Then condition (c) of Lemma 2.2.10 is a.a.s. not satisfied for $k = 1$:

$$D_1(\boldsymbol{d}_S) = \Theta\left(\frac{\sqrt{n}}{\log^7 n}\right) = \omega\left(dn^{1-\alpha}\right) = \Omega\left(d|S|^{1-\alpha}\right),$$

since $\alpha > \frac{1}{2}$ and $d = O(1)$ for this sequence. If $n \notin S$, then the degree sequence $\boldsymbol{d}_S$ has maximum degree 3. Thus, $D_k(\boldsymbol{d}_S) \le 3k$ for all $k \le |S|$. It follows by taking $\alpha = \frac{3}{4}$ and $K_0 = 2$ that

$$D_k(\boldsymbol{d}_S) \le 3k \le 6n^{1/4}k^{3/4} = K_0 n\left(\frac{k}{n}\right)^{\alpha},$$

which holds for all $k \le |S|$. Thus, condition (c) is satisfied for every $\varepsilon \le 1$. Therefore, we cannot directly apply Lemma 2.2.10 to determine the chromatic number of the percolated graph $G[S]$, since $\boldsymbol{d}_S$ may or may not satisfy the conditions of the lemma with non-trivial probability. However, there is an obvious caveat to fix this: the addition of a single high-degree vertex to $G[S]$ can only affect the chromatic number of the graph by at most 1, since we can simply colour this vertex

with a unique colour.

Speaking more generally, an essential feature of this class of counterexamples is the presence of a bounded number of vertices of degree $\sqrt{n}/\text{polylog n}$. This means that there is then a non-trivial probability of having none of these vertices, as well as a non-trivial probability of having at least one. If instead we consider a graphs with a slightly lower maximum degree, it is possible that we can get a similar result to Theorem 4.2.1 for this smaller range of degree sequences. If we impose the condition that $\Delta \leq \log M \log\log M$, an analogous argument to the deterministic case should show that the satisfiability of condition (c) is predictable using $\boldsymbol{d}_A$, as in the deterministic case with $\boldsymbol{d}_I$. However, better results are likely not too hard to prove. One possible conjecture is given below.

Conjecture 5.3.1. Suppose that the maximum degree is at most $n^{1/2-\delta}$ for some constant $\delta > 0$. With probability $1 - o(1)$, the claims of Theorem 4.2.1 hold if the set S is chosen by taking each vertex with probability p (where $p \in (0, 1)$ is independent of n) and $\boldsymbol{d}_I$ is replaced with $\boldsymbol{d}_A$.

The idea here is based in a simple concentration argument: if the maximum degree is $n^{1/2-\delta}$, then by choosing $\alpha = \frac{1}{2}(1 + \delta)$, at least $n^{\delta/2}$ vertices of degree $n^{1/2-\delta}$ must be summed over in $D_k(\boldsymbol{d}_S)$ for condition (c) of Lemma 2.2.10 to not be satisfied. If a degree sequence $\boldsymbol{d}$ contains this many big elements, then the number of such elements in S must be concentrated around its average. Thus, the numbers of terms of this size in $\boldsymbol{d}_S$ and $\boldsymbol{d}_A$ would be a.a.s. close, and thus their effects on the summations $D_k(\boldsymbol{d}_S)$ and $D_k(\boldsymbol{d}_A)$ would be similar.

5.3.1 Connectivity of the percolated graph

Here we give an analogous result to Lemma 4.1.1 for the case where S is chosen randomly. One notable difference in this model, compared to the case where S is fixed, is that $\mathbb{E}[Y_i]$ is much harder to determine asymptotically. In the case where S is fixed, Lemma 3.1.5 implies that $\mathbb{E}[Y_0 | S] \sim \tilde{y}_0(S)$, which we used to prove Lemma 4.1.1. However, the extra randomness from choosing S probabilistically means that the value of $\mathbb{E}[Y_i]$ is not as easy to determine as $\mathbb{E}[Y_i | S]$. In particular, the proof of Lemma 5.2.2 shows that S is "good" (in the sense defined in the proof, that is, that $|S| = pn(1 \pm 3\sqrt{\log n}/\sqrt{pn})$ and $d(S) = pM(1 \pm 1/pM^{1/4})$) with probability at least $1 - 4M^{-3/2}$ (noting that $M \leq n^2$). We can apply this and sum over all possible subsets S to bound $\mathbb{E}[Y_i]$:

$$\mathbb{E}[Y_i] = \sum_{S \in \mathcal{P}(n)} \mathbb{E}[Y_i | S] \mathbb{P}(S) = \sum_{\text{good } S} \mathbb{E}[Y_i | S] \mathbb{P}(S) + \sum_{\text{bad } S} \mathbb{E}[Y_i | S] \mathbb{P}(S).$$

Since S is good with probability at least $1 - 4M^{-3/2}$, and $Y_0 \leq |S| \leq M$, the second summation is between 0 and $M^{-1/2}$. Recall from Definition 5.2.1 that w_i is the number of elements in $\boldsymbol{d}_A$ with degree i and index at most ℓ, and serves as an approximation for the number of vertices in S_{small} with degree i in $G[S]$. Then Equation (5.2) implies that

$$\mathbb{E}[Y_i] = w_i(1 + o(1)) + o(\log^7 M) + O\left(M^{-1/2}\right) = w_i(1 + o(1)) + o(\log^7 M). \tag{5.4}$$

This means that we can easily determine $\mathbb{E}[Y_i]$ asymptotically if $\mathbb{E}[Y_i] \geq c \log^7 M$, as this implies that $\mathbb{E}[Y_i] \sim w_i$. Unfortunately, it also means that $\tilde{w}_i = o(1)$ does not immediately imply that $\mathbb{E}[Y_i] = o(1)$, and thus results such as Lemma 4.1.1(a) and (b) do not carry over to the percolation

model. With this in mind, we give the following result about whether $G[S]$ is a.a.s. connected or not.

Theorem 5.3.2. Let $\boldsymbol{d}$ be an n-element degree sequence with minimum degree at least 1 and maximum degree at most $\sqrt{M}/\log^7 M$. Let S and $G[S]$ be distributed as a site percolated $\mathcal{G}(\boldsymbol{d})$ with survival probability p for constant p. Then

(a) If $w_1 = o(\sqrt{M})$, then $G[S]$ is a.a.s. connected after deleting the degree 0 vertices.

(b) If $w_1 = \Theta(\sqrt{M})$, then $G[S]$ is a.a.s. disconnected, and there exists some constant $c > 0$ such that the probability that $G[S]$ is connected after deleting the degree 0 vertices is at least c.

(c) If $w_1 = \omega(\sqrt{M})$, then $G[S]$ is a.a.s. disconnected even after deleting the degree 0 vertices.

Proof. First we focus on each set S that is "good". Then, for each good set S, we apply Lemma 4.1.1 to the probability space conditional on this set S. Since S is a.a.s. good, it is sufficient to only focus on such sets S. If S is good, then by definition we know that $d(S) = \Theta(M)$ and $d(\overline{S}) = \Theta(M)$. Furthermore, if S is good, then Equation (5.2) implies that $\tilde{y}_i(S) = w_i(1+o(1))+o(\log^7 M)$ for $i \in \{0, 1\}$.

Suppose that $w_1 = o(\sqrt{M})$, and define A to be the event that $G[S]$ is connected except for isolated vertices. Then

$$\mathbb{P}(A) = \sum_{S \in \mathcal{P}([n])} \mathbb{P}(A|\, S)\, \mathbb{P}(S) \geq \sum_{\text{good } S} \mathbb{P}(A|\, S)\, \mathbb{P}(S).$$

Since $w_1 = o(\sqrt{M})$, for each good S we know that $\tilde{y}_1(S) = o(\sqrt{M})$. Thus, Lemma 4.1.1 implies that $\mathbb{P}(A|\, S) = 1 - o(1)$ for each good set S. This proves part (a). The other two cases follow similarly, noting that $\tilde{w}_0 \leq C\tilde{w}_1$ for a sufficiently large constant $C > 0$. □

5.3.2 Giant components in the percolated graph

Here we prove an analogous result to Theorem 4.4.2 in the case where $\Delta(\boldsymbol{d}) \leq \sqrt{M}/\log^7 M$ and S is chosen randomly by taking each vertex independently with probability p, where $p \in (0, 1)$ is a constant. We show that $R(\boldsymbol{d}_A) = \Theta(M)$ is a threshold for the existence of a giant component in the random graph with given degree sequence $\boldsymbol{d}$ after site percolation under these assumptions.

Theorem 5.3.3. Let $\boldsymbol{d}$ be an n-element degree sequence with minimum degree at least 1 and maximum degree at most $\sqrt{M}/\log^7 M$. Let S and $G[S]$ be distributed as a site percolated $\mathcal{G}(\boldsymbol{d})$ with survival probability p for constant p. Then $G[S]$ a.a.s. has a giant component if and only if $R(\boldsymbol{d}'_A) \geq \varepsilon M$ for some constant $\varepsilon > 0$.

Specifically, we prove the following result, which is analogous to Theorem 4.4.2. Recall that $\boldsymbol{d}^*$ is the sequence $\boldsymbol{d}$ with all terms equal to 0 removed.

Lemma 5.3.4. Let $\boldsymbol{d}$ be an n-element degree sequence with minimum degree at least 1 and maximum degree at most $\sqrt{M}/\log^7 M$. Let S and $G[S]$ be distributed as a site percolated $\mathcal{G}(\boldsymbol{d})$ with survival probability p for constant p. Then $\boldsymbol{d}^*_A$ is well-behaved and a.a.s. $\boldsymbol{d}^*_S$ is also well-behaved. Furthermore, $\boldsymbol{d}^*_S$ is a.a.s. upper bounded if and only if $\boldsymbol{d}^*_A$ is upper bounded, and $\boldsymbol{d}^*_S$ is a.a.s. lower bounded if and only if $\boldsymbol{d}^*_A$ is lower bounded.

The proof that $\boldsymbol{d}_A$ functions as a predictor for the existence of a giant component in the site percolated $\mathcal{G}(\boldsymbol{d})$ is similar to the proof of the case where S is fixed. The idea is again that the degree sequence $\boldsymbol{d}_S$ is sufficiently concentrated to a.a.s. predict the value of $R(\boldsymbol{d}_S)$ within a $o(M)$ error margin, and also that a.a.s. $\widetilde{M}(\boldsymbol{d}_A) = \Theta(M(\boldsymbol{d}_A))$ (and thus $\widetilde{M}(\boldsymbol{d}_A) = \Theta(M)$) and a.a.s. $\widetilde{M}(\boldsymbol{d}_S) \sim \widetilde{M}(\boldsymbol{d}_A)$. This statement implies that $\boldsymbol{d}_S$ is a.a.s. well-behaved. Then we apply Lemma 4.4.4 to $\boldsymbol{d}_A$ and $\boldsymbol{d}_S$ to get the desired result. Now we give the details.

Lemma 5.3.5. $\widetilde{M}(\boldsymbol{d}_A) = \Theta(M(\boldsymbol{d}_A))$, and a.a.s. $\widetilde{M}(\boldsymbol{d}_A) - \widetilde{M}(\boldsymbol{d}_S) = o(M(\boldsymbol{d}_A))$.

Proof. The proof is analogous to Lemma 4.4.11. Define $V_{\geq 3} := \{i \in V \mid d(i) \geq 3\}$. Consider two cases, one where $|V_{\geq 3}| = o(n)$ and one where $|V_{\geq 3}| = \Theta(n)$. First consider the case where $|V_{\geq 3}| = o(n)$. The sequence $\boldsymbol{d}$ has minimum degree at least 1, and thus in this case $G \sim \mathcal{G}(\boldsymbol{d})$ has $n(1 - o(1))$ vertices with degree either 1 or 2. Let V_1 be the set of degree 1 vertices in G, and let V_2 be the set of degree 2 vertices. The definition of $\tilde{w}_1$ then implies that

$$\tilde{w}_1 = p \sum_{i \in V_{\text{small}}} \mathbb{P}\left(Z_{d(i)} = 1\right) \geq p^2|V_1| + 2p^2(1-p)|V_2| \geq cn$$

for some constant $c > 0$, since $\min\{p, 1-p\} = \Theta(1)$. Since $w_2 \leq n$, it follows that $\widetilde{M}(\boldsymbol{d}_A) = M(\boldsymbol{d}_A) - 2w_2 \geq w_1 \geq cn - 1 = \Omega(w_2)$. Thus, $M(\boldsymbol{d}_A) \nsim 2w_2$, and so $\widetilde{M}(\boldsymbol{d}_A) = \Theta(M(\boldsymbol{d}_A))$. Now consider the second case, where $|V_{\geq 3}| = \Theta(n)$. Then

$$\sum_{k \geq 3} \tilde{w}_k = p \sum_{i \in V_{\text{small}}} \mathbb{P}\left(Z_{d(i)} \geq 3\right) \geq c|V_{\geq 3}|$$

for some constant $c > 0$. Thus, $M(\boldsymbol{d}_A) \geq 2w_2 + c|V_{\geq 3}|$, which implies that $M(\boldsymbol{d}_A) \nsim 2w_2$. Therefore, in either case it follows that $\widetilde{M}(\boldsymbol{d}_A) = \Theta(M(\boldsymbol{d}_A))$.

Now we prove the second claim. Lemma 5.2.2(d) implies that a.a.s. $|Y_2 - w_2| = o(w_2 + M)$, and Lemma 5.2.5 states that a.a.s. $M(\boldsymbol{d}_S) \sim M(\boldsymbol{d}_A) = p^2 M(1 + o(1))$. Since $w_2 \leq \frac{1}{2}M(\boldsymbol{d}_A)$, this implies that a.a.s.

$$\widetilde{M}(\boldsymbol{d}_S) - \widetilde{M}(\boldsymbol{d}_A) = M(\boldsymbol{d}_S) - M(\boldsymbol{d}_A) - 2Y_2 + 2w_2 = o(M).$$

This completes the proof. □

Now we show that $R(\boldsymbol{d}_S)$ and $R(\boldsymbol{d}_A)$ a.a.s. differ by $o(M)$. The proof of this is more involved than the deterministic case due to the lack of a fixed set S_{big}. The low-degree elements of both $\boldsymbol{d}_S$ and $\boldsymbol{d}_A$ are handled in a way that is much more reminiscent of the case where S is deterministic, and as such we leverage some intermediate results from that proof. In particular, recall Lemma 4.4.9 from Section 4.4 which (roughly) states that adding and removing small numbers of low-degree terms to a sequence changes the value of $R(\cdot)$ by $o(M)$.

The high-degree elements pose more of an issue here than in the case where S is fixed, since the high-degree elements of $\boldsymbol{d}_S$ and $\boldsymbol{d}_A$ may not be asymptotic, or even particularly close. Despite this, the value of $R(\boldsymbol{d}_S)$ is sufficiently concentrated. We prove a new lemma to show that the difference in high-degree terms between $\boldsymbol{d}_A$ and $\boldsymbol{d}_S$ is a.a.s. not enough to perturb $R(\cdot)$ by more than $o(M)$. Recall that $\boldsymbol{d}'$ is the sequence $\boldsymbol{d}$ ordered in non-decreasing order.

Lemma 5.3.6. Let $\boldsymbol{d}_1$ be a sequence of non-negative integers ordered in non-decreasing order with total degree M. Let $f(M)$ be some function going to infinity arbitrarily slowly. Suppose $\boldsymbol{d}_2$ is a sequence of non-negative integers such that there exists some index k that satisfies

(a) $d_1(i) = d_2(i)$ for all $i \leq k$,

(b) $d_1(j),\ d_2(j) \geq f(M)$ for all $j > k$,

(c) $d_2(i) \leq d_2(j)$ for all $i \leq k < j$, and

(d) $M(\boldsymbol{d}_2) = M(1 + o(1))$.

Then $|R(\boldsymbol{d}_1) - R(\boldsymbol{d}_2')| = o(M)$.

Proof. Define $j_{\boldsymbol{d}_i} := j_i$ for brevity, where $i \in \{1, 2\}$. By definition of $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$, we know that

$$\sum_{i=1}^{k} d_1(i)(d_1(i) - 2) = \sum_{i=1}^{k} d_2'(i)(d_2'(i) - 2).$$

This immediately implies that $j_1 \leq k$ if and only if $j_2 \leq k$, in which case the claim of the lemma immediately follows. Thus, we suppose that $j_1,\ j_2 > k$ for the rest of the proof. Note that since $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ agree on all terms up to the k^{th} term, condition (d) implies that

$$\sum_{i=k+1}^{n(\boldsymbol{d}_1)} d_1(i) = \sum_{i=k+1}^{n(\boldsymbol{d}_2)} d_2(i) + o(M).$$

If $\sum_{i=k+1}^{n(\boldsymbol{d}_1)} d_1(i) = o(M)$, then $R(\boldsymbol{d}_1),\ R(\boldsymbol{d}_2') = o(M)$, and the claim of the lemma immediately follows in this case too. Now we suppose that $j_1,\ j_2 > k$ and $\sum_{i=k+1}^{n(\boldsymbol{d}_1)} d_1(i) = cM$ for some $c = \Theta(1)$. In this case, we show that $R(\boldsymbol{d}_1), R(\boldsymbol{d}_2) = cM(1 - o(1))$. Suppose there exists some partition (X_1, X_2) of the set $\{k + 1, \ldots, n(\boldsymbol{d}_1)\}$ such that $X_1 = \{k + 1, \ldots, x_1\}$ and $d_1(X_1) \geq 2M/f(M)$ and $d_1(X_1) \leq M/\sqrt{f(M)}$. Then

$$\sum_{i=1}^{x_1} d_1(i)(d_1(i) - 2) > -M + (f(M) - 2)2M/f(M) > 0.$$

This implies that $R(\boldsymbol{d}_1) \geq \sum_{i=x_1}^{n(\boldsymbol{d}_1)} d_1(i) - d_1(X_1) = cM(1 - o(1))$. Now if no such partition (X_1, X_2) exists, then $d_1(x) > M/2\sqrt{f(M)}$ for some $x \in \{k + 1, n(d_1)\}$. Then

$$\sum_{i=1}^{x} d_1(i)(d_1(i) - 2) > -M + \frac{1}{3}M\sqrt{f(M)} > 0.$$

This implies that $R(\boldsymbol{d}_1) \geq \sum_{i=x}^{n(\boldsymbol{d}_1)} d_1(i) > cM - 2M/f(M) = cM(1 - o(1))$. Since $d_2(j) \geq f(M)$ for all $j > k$, an identical argument shows that $R(\boldsymbol{d}_2') \geq cM(1 - o(1))$. This completes the proof. □

Lemma 5.3.7. With probability $1 - o(1)$, $|R(\boldsymbol{d}_S) - R(\boldsymbol{d}_A)| = o(M(\boldsymbol{d}_A))$.

Proof. Define a sequence $\boldsymbol{k}$ which contains w_k elements equal to k for each $k \leq J$ (in non-decreasing order), as well as containing the term $d_S(i)$ for each $i \in S_{\text{big}}$ (in the same order as they appear in $\boldsymbol{d}_S$). This sequence matches $\boldsymbol{d}_A$ for the "small" terms and $\boldsymbol{d}_S$ for the "big" terms. Suppose

that $\boldsymbol{d}_S$ satisfies conditions (a) – (d) of Lemma 5.2.2, which occurs with probability $1-o(1)$. This implies that a.a.s.

$$|Y_i - w_i| \leq \left(\frac{y_i}{\log^5 M} + \log^7 M\right)(1+o(1))$$

for all $i \leq J$. Thus, by Lemma 4.4.9, it follows that a.a.s.

$$|R(\boldsymbol{d}_S) - R(\boldsymbol{k})| = O\left(\frac{M(\log\log M)^3}{\log M}\right) = o(M). \tag{5.5}$$

Next we show that a.a.s. $|R(\boldsymbol{d}_A) - R(\boldsymbol{k})| = o(M)$. From this result, the claim of the lemma follows via the triangle inequality and the fact that $M(\boldsymbol{d}_A) = \Theta(M)$ (from Lemma 5.2.5). We now show that $\boldsymbol{d}_A$ and $\boldsymbol{k}$ a.a.s. satisfy the conditions of Lemma 5.3.6, and thus a.a.s. $|R(\boldsymbol{d}_A) - R(\boldsymbol{k})| = o(M(\boldsymbol{d}_A)) = o(M)$. Choose k to be the largest index in $\boldsymbol{d}_A$ such that $d_A(k) \leq \frac{1}{2}pJ$. Note that $\tilde{z}_k = 0$ for all $k \leq \frac{1}{2}pJ$ by definition (and standard binomial concentration results), and a.a.s. $G[S]$ contains no vertex $v \in S_{\text{big}}$ such that $d_S(v) \leq \frac{1}{2}pJ$ by Lemma 5.2.2. Thus, $\boldsymbol{d}_A$ and $\boldsymbol{k}$ a.a.s. satisfy conditions (a) – (c) of Lemma 5.3.6.

Now we consider condition (d). Note that since $\boldsymbol{d}_A$ and $\boldsymbol{k}$ agree on the first $\sum_j w_j$ elements, it suffices to show that $d_S(S_{\text{big}})$ and $\sum_{k\geq 0} kz_k$ are a.a.s. within $o(M)$ of each other. Since $\Delta(\boldsymbol{d}) \leq \sqrt{M}/\log^7 M$, Lemma 5.2.4 implies that a.a.s. $d_S(S_{\text{big}}) = p^2 d(V_{\text{big}}) \pm o(M)$. Since $\boldsymbol{k}$ contains the term $d_S(i)$ for each $i \in S_{\text{big}}$, it follows that a.a.s. $|M(\boldsymbol{d}_A) - M(\boldsymbol{d}_S)| = o(M) = o(M(\boldsymbol{d}_A))$. Thus, Lemma 5.3.6 implies that a.a.s. $|R(\boldsymbol{d}_A) - R(\boldsymbol{k})| = o(M)$, and the claim of the lemma follows by the triangle inequality. □

To finish off the proof of Lemma 5.3.4 we apply Lemma 4.4.4. Since this lemma only requires that $n(\boldsymbol{d}_1) \sim n(\boldsymbol{d}_2)$, it is not an issue that $\boldsymbol{d}_A$ and $\boldsymbol{d}_S$ do not necessarily have the same number of terms.

Proof of Lemma 5.3.4. Lemma 5.2.2(a) and Remark 5.2.3 imply that a.a.s. $n(\boldsymbol{d}_S) \sim n(\boldsymbol{d}_A)$ and $n(\boldsymbol{d}_S^*) \sim n(\boldsymbol{d}_A^*)$. Lemma 5.3.5 states that a.a.s. $\widetilde{M}(\boldsymbol{d}_S) \sim \widetilde{M}(\boldsymbol{d}_A) = \Theta(M)$. This implies that $\boldsymbol{d}_A^*$ is well-behaved and that $\boldsymbol{d}_S^*$ is a.a.s. well-behaved. Lemma 5.3.7 then implies that a.a.s. $R(\boldsymbol{d}_S^*) - R(\boldsymbol{d}_A^*) = R(\boldsymbol{d}_S) - R(\boldsymbol{d}_A) = o(\widetilde{M}(\boldsymbol{d}_A))$. Therefore, the conditions of Lemma 4.4.4 are a.a.s. satisfied by $\boldsymbol{d}_S^*$ and $\boldsymbol{d}_A^*$, and the claim of the lemma follows. □

Proof of Theorem 5.3.3. Theorem 5.3.3 follows from Lemmas 5.3.4 and 5.3.5. Recall that $R(\boldsymbol{d}^*) = R(\boldsymbol{d})$ for all sequences $\boldsymbol{d}$. If there exists some $\varepsilon > 0$ such that $R(\boldsymbol{d}_A) \geq \varepsilon M$, then by definition $\boldsymbol{d}_A^*$ is lower bounded. Then a.a.s. $\boldsymbol{d}_S^*$ is lower bounded. Thus, a.a.s. $G[S]$ has a giant component by Theorem 4.4.1. Otherwise, if $R(\boldsymbol{d}_A) = o(M)$, then $\boldsymbol{d}_A^*$ is upper bounded, and thus a.a.s. $\boldsymbol{d}_S^*$ is upper bounded. Therefore, $G[S]$ a.a.s. does not contain a giant component. Therefore, $R(\boldsymbol{d}_A) \geq \varepsilon M$ is a threshold for the existence of a giant component in $\mathcal{G}(\boldsymbol{d})$ subject to site percolation under our assumptions on p and $\boldsymbol{d}$, which proves Theorem 5.3.3. □

Remark 5.3.8. Another method of studying the site percolated $\mathcal{G}(\boldsymbol{d})$ random graph would be to study $\boldsymbol{d}_I(S)$ for all "good" sets S. Ideally in this method, one could show that if S is good then $\boldsymbol{d}_I(S)$ is close to $\boldsymbol{d}_A$ in a sense that implies that $R(\boldsymbol{d}_A) - R(\boldsymbol{d}_I(S)) = o(M)$. Then we could apply Theorem 4.4.2 to compare $\boldsymbol{d}_I(S)$ and $\boldsymbol{d}_S$ for each fixed, good set S. Since S is a.a.s. good, this

would imply that the percolated graph $G[S]$ a.a.s. has a giant component if and only $R(\boldsymbol{d}_A) \geq \varepsilon M$, as desired.

We do not give the details here, but this method does work: if S is good, then Equations (5.2) and (5.3) imply that $\mathbb{E}[Y_i | S]$ is close to w_i, and the value of $d_I(S_{\text{big}})$ can be shown to be a.a.s. $p^2 d(V_{\text{big}}) + o(M)$ (and thus the definition of "good" could be modified to include concentration of $d(S_{\text{big}})$). Thus, the proof of Lemma 5.3.7 would carry over to show that $|R(\boldsymbol{d}_A) - R(\boldsymbol{d}_I(S))| = o(M(\boldsymbol{d}_A))$ for every good set S. We do not use this method though, as the work required to compare $\boldsymbol{d}_A$ and $\boldsymbol{d}_I$ is basically the same as the work required to compare $\boldsymbol{d}_A$ and $\boldsymbol{d}_S$ directly, since results akin to Lemmas 5.3.5 and 5.3.6 are still required. Thus, the proof using this method would be slightly more cumbersome, as it would be very similar to the current proof but with an extra step, since we must compare $\boldsymbol{d}_A$ and $\boldsymbol{d}_I(S)$ as well as comparing $\boldsymbol{d}_I(S)$ and $\boldsymbol{d}_S$ for good sets S.

Chapter 6

Exploration approach in the configuration model

In this chapter we again consider the existence of giant components in induced subgraphs, but using a different method. In Chapter 2 we discussed a result by Joos et al. [83] which determined the threshold for giant components in $\mathcal{G}(\boldsymbol{d})$. Their proof, as well as many previous results (such as Molloy and Reed [113]) used an exploration process on the random graph to study the component structure. We adapt the modified breadth-first search random process used by Joos et al. to determine the existence of giant components in induced subgraphs. This is the basis of what we referred to in Chapter 1 as the exploration method, since our analysis revolves heavily around studying the exploration process.

In this chapter, we focus on doing so in configuration model random graphs $\mathcal{C}(\boldsymbol{d})$. For a given sequence $\boldsymbol{d}$ and $S \subset [n]$, we characterise the asymptotic threshold for when the subgraph of a uniformly random pairing with degree sequence $\boldsymbol{d}$ induced on the set of vertices S a.a.s. has a giant component. In the following chapter, we then adapt this framework to the random graph model $\mathcal{G}(\boldsymbol{d})$. We begin our analysis in the configuration model since it is much more straightforward to analyse than $\mathcal{G}(\boldsymbol{d})$, and thus we can gain intuition in this model without getting mired in switchings and other technicalities associated with $\mathcal{G}(\boldsymbol{d})$.

In the next section, we describe our main results and the threshold we obtain in the configuration model. We also lay out some basic implications of the configuration model that are crucial to our analysis in this setting. In Section 6.2, we describe the exploration process that we use to prove our main results, and determine the threshold for the a.a.s. existence of a component in $G[S]$ with $\Theta(M)$ edges. In Section 6.3, we show that the a.a.s. (non-)existence of a component with $\Theta(M)$ edges implies the a.a.s. (non-)existence of a component with $\Theta(|S|)$ vertices. That is, if $G[S]$ a.a.s. contains a component of linear size, then it also a.a.s. contains a component of linear order.

6.1 Main results

Let $\boldsymbol{d}$ be an n-element sequence of non-negative integers with even sum. Since we are considering pairings rather than graphs in this chapter, we do not need this sequence to satisfy the Erdős-Gallai criterion for graphical sequences given in Theorem 2.2.6. Instead, we call a sequence *pseudographical* if it is a sequence of non-negative integers with even sum. Let $S \subset [n]$ and

$\overline{S} = [n] \backslash S$, and define

$$\mathsf{d}_S(w) = \begin{cases} d(w) & w \in S, \\ 1 & w \notin S. \end{cases} \tag{6.1}$$

We stress here that this definition is not related to the sequence $\boldsymbol{d}_S$ as defined and used in Chapters 3 to 5, which was the degree sequence of the induced graph $G[S]$. We also stress that we do not use $\boldsymbol{d}_S$ to mean the degree sequence of $G[S]$ in Chapters 6 and 7.

Without loss of generality we assume that the sequence $\boldsymbol{d}$ is ordered in non-decreasing order with respect to $d(w)$ ($\mathsf{d}_S(w) - 2$). Thus, $(\overline{S}, S)$ is a partition of $[n]$ such that $\overline{S} = \{1, 2, \dots, n - |S|\}$ and $S = \{n - |S| + 1, \dots, n\}$. For such a sequence $\boldsymbol{d}$ and set S, define

$$\begin{aligned} j_S(\boldsymbol{d}) &= \min\left(\left\{ j : j \in [n] \text{ and } \sum_{i \le j} d(i)\left(\mathsf{d}_S(i) - 2\right) > 0 \right\} \cup \{n\}\right), \\ R_S(\boldsymbol{d}) &= \sum_{i=j_S}^{n} d(i), \end{aligned} \tag{6.2}$$

and recall that $M(\boldsymbol{d}) = \sum_{i=1}^{n} d(i)$ is the total degree of $\boldsymbol{d}$. We abbreviate these by omitting the $\boldsymbol{d}$ in most cases. Note that $j_S(\boldsymbol{d})$ can be defined equivalently as

$$j_S(\boldsymbol{d}) = \min\left(\left\{ j : j \ge |\overline{S}| + 1 \text{ and } \sum_{i=|\overline{S}|+1}^{j} d(i)\left(d(i) - 2\right) > d(\overline{S}) \right\} \cup \{n\}\right),$$

where $[j] = \{1, \dots, j\}$. Much like in Chapters 3 and 4, we emphasise the distinction between $M(\boldsymbol{d})$ and $\widetilde{M}(\boldsymbol{d})$ (where $\widetilde{M}(\boldsymbol{d})$ is defined in Theorem 2.2.15): if $n_2(\boldsymbol{d})$ is the number of elements in $\boldsymbol{d}$ equal to 2, then $M(\boldsymbol{d}) = \widetilde{M}(\boldsymbol{d}) + 2n_2(\boldsymbol{d})$. We also recall from these chapters the definition of the *total degree* of a set $A \subset [n]$ as

$$d(A) := \sum_{v \in A} d(v).$$

Let $\mathcal{C}(\boldsymbol{d})$ be the space of all pairing graphs with degree sequence $\boldsymbol{d}$, and let P be a pairing from $\mathcal{C}(\boldsymbol{d})$ chosen uniformly at random. Recall from Section 2.2.4 the various definitions of $O(\cdot)$, $o(\cdot)$, $\Omega(\cdot)$, $\omega(\cdot)$, and $\Theta(\cdot)$.

Theorem 6.1.1. Let $\boldsymbol{d}$ be a pseudographical sequence. Consider a uniformly random element $P \in \mathcal{C}(\boldsymbol{d})$, and let $S \subset [n]$. Then:

(a) If $R_S(\boldsymbol{d}) = o(M)$ and $d(\overline{S}) = \Theta(M)$, the probability that $P[S]$ contains a component of size $\Theta(M)$ is $o(1)$.

(b) If $R_S(\boldsymbol{d}) = \Theta(M)$, the probability that $P[S]$ contains a component of size $\Theta(M)$ is $1 - o(1)$.

We call case (a) of Theorem 6.1.1 the *subcritical* case, and case (b) the *supercritical* case. Theorem 6.1.2 translates the statement of Theorem 6.1.1 from a result about edges to a result about vertices.

Theorem 6.1.2. Let $\boldsymbol{d}$ be a pseudographical sequence. Consider a uniformly random element $P \in \mathcal{C}(\boldsymbol{d})$, and let $S \subset [n]$. Suppose that $d(S) = \Theta(M)$. Then Theorem 6.1.1 still holds if "of size $\Theta(M)$" is replaced with "of order $\Theta(|S|)$".

This result is strongly reminiscent of Theorem 2.2.15 about the existence of giant components in $\mathcal{G}(\boldsymbol{d})$. The set of deleted vertices $\overline{S}$ have the same effect in this threshold as degree 1 vertices do in the threshold given in Theorem 2.2.15, as well as more classical thresholds such as the one by Molloy and Reed [113]. This is natural, given the similarities in the proof method, but there are more intuitive reasons for this similarity. Degree 1 vertices and vertices in $\overline{S}$ have a similar effect on the exploration process: every time a degree 1 vertex or a vertex in $\overline{S}$ is discovered, the number of unpaired half edges in the component currently being explored decreases by 1.

In comparison to the result of Janson [74], the results given in Theorems 6.1.1 and 6.1.2 expand the ranges sequences for which the existence a giant component is predictable in the configuration model. There are no restrictions on the maximum degree of a vertex or on the proportion of vertices with a given degree. However, the most notable drawback in this result is that there is no prediction for the order of the largest component beyond whether it is linear or sublinear, while the Janson result gives the order up to an error of $o(n)$.

In the supercritical case of Theorem 6.1.2, the condition that $d(S) = \Theta(M)$ is implied by the fact that $R_S = \Theta(M)$. However, the condition that $d(S) = \Theta(M)$ cannot be relaxed in the subcritical case of Theorem 6.1.2. We give an example of a pair $(\boldsymbol{d}, S)$ where $R_S = o(M)$ and $P[S]$ a.a.s. contains a component with $\Theta(|S|)$ vertices. Lemmas 6.3.1 and 6.3.3 are the specific lemmas that allow us to convert the statement about the size of the largest component, given in Theorem 6.1.1, into the statement of Theorem 6.1.2 about the order of the largest component.

Example 6.1.3. Suppose that S contains $n^{2/5}$ vertices of degree $n^{1/2}$, and $\overline{S}$ contains $n - n^{2/5}$ vertices of bounded degree. Then $M = \Theta(n)$ and $d(S) = o(n)$, so $R_S \leq (S) = o(M)$. Thus, Theorem 6.1.1 applies and says that a.a.s. the induced pairing $P[S]$ has no component of size $\Theta(M)$. In fact, this is always true: the maximum number of edges in $P[S]$ is $\binom{n^{2/5}}{2} \sim \frac{1}{2}n^{4/5}$. However, a result we give later (specifically Lemma 6.2.2) implies that if (U, W) is a partition of S, then the probability that (U, W) induces an empty cut (that is, P contains no edges of the form uv where $u \in U$ and $w \in W$) is bounded from above by

$$\exp\left(-\frac{1}{2M}d(U)(n^{9/10} - d(U))\right)$$

for every non-trivial partition (U, W) of S. Since every vertex in U has degree $n^{1/2}$, this probability is at most $\exp\left(-Cn^{2/5}|U|\right)$ for some constant $C > 0$. Thus,

$$\mathbb{P}\left(\text{some } (U, V) \text{ induces an empty cut}\right) < \sum_{i=1}^{n^{2/5}} \binom{n^{2/5}}{k} \exp\left(-Ckn^{2/5}\right) = o(1).$$

Therefore, the induced pairing $P[S]$ is a.a.s. connected, despite satisfying the subcritical criterion of Theorem 6.1.1.

We prove Theorem 6.1.1 by modifying the exploration process used by Joos et al. [83] to the problem at hand. To describe this exploration process more neatly, we "explode" vertices outside S into degree 1 vertices while maintaining the total degree of $\overline{S}$. This does not affect the induced

pairing $P[S]$ at all, but makes the analysis of the exploration process much more intuitive. To define this formally, let $\boldsymbol{k} = (k(1), \dots, k(n))$ be a pseudographical sequence and let $S \subset [n]$. Define a new sequence $\boldsymbol{d} := \boldsymbol{d}(\boldsymbol{k}) = (d(1), \dots, d(n'))$ by replacing each vertex $v \in \overline{S}$ with $d(v)$ vertices of degree 1 (ordering $\boldsymbol{d}$ such that $d(j) \leq d(j+1)$ for all $j < n'$). Note that if $\boldsymbol{k}$ was an n-element sequence, then $\boldsymbol{d}$ has more than n elements, and thus $P \in \mathcal{C}(\boldsymbol{d})$ has more than n vertices. However, the sub-pairing $P[S]$ still has exactly $|S|$ vertices. The following lemma implies that we only need to consider sequences of this type. It shows that for all pseudographical sequences $\boldsymbol{k}$ there exists a corresponding (pseudographical) sequence $\boldsymbol{d}$ such that $P[S]$ has the same distribution in both $\mathcal{C}(\boldsymbol{k})$ and $\mathcal{C}(\boldsymbol{d})$.

Lemma 6.1.4. For every pseudographical sequence $\boldsymbol{k}$, the induced pairings $P(\boldsymbol{k})[S]$ and $P(\boldsymbol{d}(\boldsymbol{k}))[S]$ have the same distribution.

Proof. Define a map $f : \mathcal{C}(\boldsymbol{k}) \to \mathcal{C}(\boldsymbol{d})$ as follows: for each $P \in \mathcal{C}(\boldsymbol{k})$, suppose the labels of the half edges in $B(i)$, the bin corresponding to vertex for some $i \in \overline{S}$, are labelled $i_1, \dots, i_{k(i)}$. Then for each $i \in \overline{S}$, delete it and replace it with $k(i)$ vertices of degree 1 (where vertices are labelled by any rule), each containing a different half edge from vertex i. For every pair of half edges $u, v \in [M]$, the probability that u and v are paired in $\mathcal{C}(\boldsymbol{k})$ is $\frac{1}{M-1}$, and the same holds for $f(P) \in \mathcal{C}(\boldsymbol{d})$. Thus, define a coupling between uniform generation of pairings in $\mathcal{C}(\boldsymbol{k})$ and $\mathcal{C}(\boldsymbol{d})$ by making the same half edge pairings (that is, using the same permutation to determine the pairings for P and $f(P)$). Deleting $\overline{S}$ and all incident edges in both P and $f(P)$ then yields the same induced pairing. Thus, the induced pairings $P(\boldsymbol{k})[S]$ and $P(\boldsymbol{d})[S]$ have the same probability of yielding any given sub-pairing $P[S]$. □

6.2 The configuration model exploration process and analysis

Theorem 6.1.1 is proved by analysing a modified breadth-first search process on a uniformly random element of $\mathcal{C}(\boldsymbol{d})$. In this section we describe the process in full, along with how we use it to analyse a uniform random pairing. By Lemma 6.1.4, we may assume that the degree sequence $\boldsymbol{d}$ is such that $d(w) = 1$ for all $w \notin S$. Note that this means $\boldsymbol{d}$ may have more than n components due to "exploding" vertices in $\overline{S}$, despite the original pairing of interest having exactly n vertices. However, as mentioned above, $P[S]$ still has exactly $|S|$ vertices.

6.2.1 The configuration model exploration process

The initial input to the exploration process is a pairing P with (n-element) pseudographical sequence $\boldsymbol{d}$ and a subset $S \subset [n]$ such that $d(i) = 1$ for all $i \notin S$. The algorithm also fixes a uniformly random ordering on the half edges in the pairing, which we denote by $\sigma \in S_M$. When we refer to the "lowest indexed" half edge, it is with respect to this permutation. Call a pairing $P \in \mathcal{C}(\boldsymbol{d})$ with a half edge ordering an *input*. The algorithm performs the exploration process on this input in a manner that we describe below. The process is deterministic for a given input, but can also be defined in a randomised way without reference to a permutation: instead of picking the relevant half edge with the lowest index, simply pick one uniformly at random from the set of relevant half edges.

At time $t = 0$ of the exploration process, initialise $T_0 = (V_0, E_0, \mathcal{X}_0)$. The set V_0 is a subset of the vertex set $V(P)$ (the specific subset is defined differently for the subcritical and supercritical

cases, at the start of Sections 6.2.3 and 6.2.4 respectively). This set V_0 is called the *preprocessing set*. The edge set E_t, for all $t \geq 0$, is the edge set of the induced partial pairing $P[V_t]$. The set $\mathfrak{X}_t$ is a set of half edges i such that the parent vertex u_i is in V_t and the parent vertex of $m(i)$ is not in V_t. We call these half edges *open edges at time t*. For all $t \geq 0$, define $M_t := M - d(V_t)$, the total degree of the vertices not in T_t. As the process progresses it produces a nested set of partial sub-pairings

$$T_0 \subset T_1 \subset T_2 \subset \cdots \subset (V(P), E(P), \emptyset)$$

where $T_t \subset T_{t+1}$ means that $V_t \subset V_{t+1}$ and $E_t \subset E_{t+1}$. We call T_t the *partially explored pairing at time t*, or simply the *partial pairing at time t*.

After time $t = 0$ the process is defined iteratively. For each t, let $X_t = |\mathfrak{X}_t|$ be the number of open edges in T_t. At time $t \geq 1$, if $X_{t-1} > 0$, then the there exists an open edge at time $t-1$, that is, a half edge in whose parent vertex is in V_{t-1} and whose mate is not in V_{t-1}. The process takes the lowest-indexed open edge (whose parent vertex is denoted v_t) in $\mathfrak{X}_{t-1}$ and pairs it to its mate. Call the parent vertex of this mate w_t. Then $V_t = V_{t-1} \cup \{w_t\}$, and $E_t = E_{t-1} \cup \{im(i) \mid i \in B(w_t), m(i) \in \mathfrak{X}_t\}$. The edges added to E_t at this step, other than $im(i)$, are called *back edges*. On the other hand, if $X_{t-1} = 0$, that is, all half edges in T_{t-1} are paired and there are no open edges, then choose the vertex containing the lowest-indexed half edge in S to be w_t. This starts a new component in $P[S]$. Let $d'_{S,t}(w_t)$ be the number of back edges between w_t and T_{t-1}. Then for all $t > 0$, we have that

$$X_t = X_{t-1} + d(w_t) - 2 - 2d'_{S,t}(w_t) \text{ and } M_t = M_{t-1} - d(w_t). \tag{6.3}$$

This then iterates for each step of the exploration process.

The vertex sets of T_t and T_{t-1} differ by a single vertex w_t. We say that w_t is *discovered* at time t. The edge sets only differ by edges incident to w_t, as do the sets of half edges. We call the component of T_{t-1} containing v_t the *active component*. Define $S_t = S \backslash V_t$ and $\overline{S}_t = \overline{S} \backslash V_t$, the analogous set for $\overline{S}$. Let $d : \mathcal{P}(V) \to \mathbb{N}$ to be the degree of a vertex (or set of vertices), where $d(U) = \sum_{u \in U} d(u)$. Then define $M_{t-1} = d(S_{t-1}) + d(\overline{S}_{t-1})$.

To analyse this process on a random element of $\mathcal{C}(\boldsymbol{d})$, we use the method of deferred decisions to generate the pairing as the process evolves. In this case, the input can be considered an unpaired set of M points, distributed among n labelled bins as per the degree sequence $\boldsymbol{d}$. At each step $t > 0$, an open edge i_t is chosen uniformly at random from all open edges in $\mathfrak{X}_t$, and its mate $m(i_t)$ is chosen uniformly at random from all half edges belonging to vertices $v \notin V_{t-1}$. In this way, we generate a uniformly random pairing stepwise. For each t, define $\mathbb{P}_t(\cdot) = \mathbb{P}(\cdot \,|\, T_t, \ldots, T_0)$ to be the probability measure conditional on the deferred decision exploration process up to time t. This conditional probability space is the uniform probability space of all pairings $P \in \mathcal{C}(\boldsymbol{d})$ such that the induced pairing $P[V_t]$ contains exactly the edges in E_t. Similarly, define $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot \,|\, T_t, \ldots, T_0]$, the expectation conditional on the deferred decision exploration process up to time t. We say that vertices and edges in T_t have been *revealed* or *explored*.

The proof of Theorem 6.1.1 then revolves around keeping track of X_t throughout this deferred decision process. Specifically, we show that at each step the expected change in X_t is slightly positive or negative, depending on whether the degree sequence satisfies the supercritical or subcritical

criterion of Theorem 6.1.1. Then we show that X_t, as a sum of many random small variables, is concentrated for large values of t. We show that in the supercritical regime there exists some t such that a.a.s. $X_t > \beta M$ for some constant $\beta > 0$. We then argue that this implies that the corresponding component in $P[S]$ a.a.s. has $\Theta(M)$ edges, as well as $\Theta(|S|)$ vertices. On the other hand, in the subcritical regime we prove that a.a.s. $X_t = 0$ for some $t = o(M)$, and from there argue that a.a.s. all components of $P[S]$ contain at most $o(n)$ vertices.

In order to make X_t sufficiently concentrated throughout this process, we "preprocess" high-degree vertices to remove them from the exploration process, placing them into V_0. This keeps $X_t - X_{t-1}$ small even when the maximum degree of $\boldsymbol{d}$ is arbitrarily large. The exact definition of the preprocessing step is slightly different depending on whether we are analysing the subcritical or supercritical case, and is defined formally in the respective sections of the proof. However, the general idea is the same, borrowed from the process defined by Joos et al. [83].

6.2.2 Adjacency of vertices in the configuration model

One of the simplicities of analysing the configuration model, in comparison to the random graph model $\mathcal{G}(\boldsymbol{d})$, is that determining the probability that two vertices are adjacent is much more straightforward. More generally, the probability that two disjoint subsets of vertices U and W have an edge between them is the same as the probability that two vertices of degree $\sum_{u\in U} d(u)$ and $\sum_{w\in W} d(w)$ are adjacent. To analyse $\mathcal{C}(\boldsymbol{d})$ and the exploration process, we consistently use the following observation.

Observation 6.2.1. Let A be a set of pairs of half edges in $[M]$ such that no half edge appears in more than one pair. Conditional on any such set A, the remaining half edges are paired uniformly at random. Conditional on the event that a half edge u is paired with some half edge in a set $K \subset [M]$ (where $u \notin K$), the probability that u is paired with some particular $k \in K$ is exactly $1/|K|$.

This is simply because the half edges are all paired uniformly at random, so any subset is also uniformly random. This observation implies that, conditional on an arbitrary partial pairing T_{t-1} for some $t \geq 1$, the probability that $w_t = w$, for some $w \notin V_{t-1}$, is exactly $\frac{d(w)}{M_{t-1}}$. This forms the crux of our analysis of the exploration process, and obtaining an analogous result to this is the main hurdle in adapting this result to the random graph model in Chapter 7. Observation 6.2.1 is also the main tool used in the proof of the following lemma, which is used many times throughout the proofs of the main results.

Lemma 6.2.2. Let U and W be sets of vertices with total degree $d(U)$ and $d(W)$ respectively. The probability that there are no edges between U and W is at most

$$\exp\left(-\frac{1}{2M}d(U)d(W)\right).$$

Proof. Consider a process of stepwise checking for edges between U and W. At each step, choose a half edge in W and reveal whether it is paired to a half edge in U. Suppose that out of K half edges in W checked so far, no edge between U and W has been revealed. Conditional on this, the probability that the next half edge is paired to a half edge in U is $\frac{d(U)}{M-1-2K}$, which is bounded below by $\frac{d(U)}{M}$. Since revealing the mate of each half edge consumes at most two half edges in W,

at least $\frac{1}{2}d(W)$ half edges in W must have their mate revealed before the mate of each half edge in W is known. Thus, the probability that the partition (U, W) induces an empty cut is at most

$$\left(1 - \frac{d(U)}{M}\right)^{\frac{1}{2}d(W)}.$$

Taking the exponential of the logarithm of both sides gives that

$$\left(1 - \frac{d(U)}{M}\right)^{\frac{1}{2}d(W)} = \exp\left(\frac{1}{2}d(W)\log\left(1 - \frac{d(U)}{M}\right)\right) < \exp\left(-\frac{1}{2}d(W)\frac{d(U)}{M}\right),$$

where the second inequality follows from the Taylor expansion of the logarithm. □

6.2.3 Analysis in the supercritical regime

In the supercritical criterion, we assume that there exists some $\varepsilon > 0$ such that $R_S \geq \varepsilon M$. Define

$$H^+ = \left\{j \in [n] : d(j) > \frac{\sqrt{M}}{\log M}\right\}.$$

As we only consider pairs $(\boldsymbol{d}, S)$ such that $d(i) = 1$ for all $i \notin S$, it follows that $H^+ \subset S$. The proof of the supercritical case splits into two cases, depending on whether $d(H^+) = \Theta(M)$ or $d(H^+) = o(M)$. In both cases, we show that there exists some set of vertices that are connected in $P[S]$ and have total degree $\Theta(M)$. It then remains to be shown that a.a.s. a positive fraction of the half edges in vertices belonging to this component in $P[S]$ pair with other half edges in S.

Lemma 6.2.3. If $d(H^+) = cM$ for some c bounded away from 0, then the partial pairing induced by the vertices of H^+ is a.a.s. connected.

Proof. First suppose that H^+ contains only one vertex, which would be vertex n. Then immediately $P[H^+]$ is connected. From here we suppose that $|H^+| \geq 2$. Consider the probability that some partition (U, W) of H^+ is a cut of $P[H^+]$. Without loss of generality we may suppose $d(U) \leq d(W)$. Since $d(H^+) = cM$, Lemma 6.2.2 implies that the probability that the partition (U, W) induces an empty cut is

$$\mathbb{P}\left(U \nsim W\right) < \exp\left(-\frac{1}{2M}d(U)d(W)\right).$$

Since every vertex in H^+ has degree greater than $\frac{\sqrt{M}}{\log M}$, we know that

$$|H^+| \leq \frac{cM}{\sqrt{M}/\log M} = c\sqrt{M}\log M.$$

Furthermore, if we assume that $|U| = k$, it also follows that

$$d(U) \geq k\frac{\sqrt{M}}{\log M}.$$

First, consider the case where U contains at most $M^{\frac{2}{3}}$ half edges, that is, $d(U) \leq M^{2/3}$. The probability that U and W induce an empty cut is then at most $\exp\left(-\frac{1}{2}c'd(U)\right)$, for some constant

$c' < c$. Since $d(U) \leq M^{2/3}$ by assumption, it must be true that $k \leq M^{1/6} \log M$. Thus, the probability that any partition (U, W) where $d(U) \leq M^{2/3}$ induces an empty cut is less than

$$\sum_{k=1}^{M^{1/6}\log M} \binom{M^{1/6}\log M}{k} e^{-\frac{c'k}{2}\frac{\sqrt{M}}{\log M}} \leq \sum_{k=1}^{M^{1/6}\log M} \left(\frac{M^{1/6}\log M}{k} e^{1-\frac{c'\sqrt{M}}{2\log M}}\right)^k.$$

Asymptotically the summand is $o(\frac{1}{M})$ and thus the sum is $o(1)$.

Now consider the case where both parts contain more than $M^{\frac{2}{3}}$ half edges. Then W has total degree at least bM for $b = \frac{c}{3}$ (since we assume that $d(U) \leq d(W)$) and U has total degree at least $M^{\frac{2}{3}}$. Lemma 6.2.2 then implies that

$$\mathbb{P}(U \nsim W) \leq \exp\left(-\frac{1}{2} bM^{\frac{2}{3}}\right)$$

for any partition (U, W) where $d(U) \leq d(W)$ and $d(U) > M^{2/3}$. The total number of vertices in H^+ is less than $\sqrt{M}\log M$, since each vertex has degree greater than $\frac{\sqrt{M}}{\log M}$. Thus, there are less than $2^{\sqrt{M}\log M}$ possible partitions (U, W), since each of the $\sqrt{M}\log M$ vertices in H^+ can be placed in either part. Then Markov's inequality implies that the probability that some partition (U, W) forms an empty cut is at most

$$2^{\sqrt{M}\log M} e^{-bM^{2/3}} < e^{\sqrt{M}\log M - bM^{\frac{2}{3}}} = o(1).$$

Combining these cases, the probability that the sub-pairing induced by H^+ is not connected is $o(1)$. □

If some constant fraction of the half edges in H^+ pair with other half edges in vertices in H^+, then the component of $P[H^+]$ has $\Theta(M)$ edges, and thus $P[S]$ has a component with $\Theta(M)$ edges. Otherwise, $P[H^+]$ is a.a.s. a partial pairing that contains a connected component with αM open edges, for some constant $\alpha > 0$. In Lemma 6.2.9, we show that this also implies that a.a.s. $P[S]$ has a component with $\Theta(M)$ edges.

Thus, for the remainder of the proof, we suppose that $d(H^+) = o(M)$. If H^+ is non-empty and $d(H^+) \leq \delta M$ for some $\delta \to 0$, as part of the preprocessing step we move H^+ to $\overline{S}$, exploding the vertices into $d(H^+)$ degree 1 vertices, creating a new sequence $\boldsymbol{d}'$ and redefining S. We prove that the induced sub-pairing $P[S \backslash H^+]$ a.a.s. has a component with $\Theta(M)$ edges. This implies the existence of a component with $\Theta(M)$ edges in the pairing $P[S]$, since adding the vertices in S^+ back into the sub-pairing can only increase the size and order of the largest component. We still use H^+ to refer to the set of vertices in S with degree greater than $\sqrt{M}/\log M$, as we need to check that deleting these vertices does not affect the evolution of the exploration process significantly.

Set $V_0 = \{v\}$, where v is the vertex containing the lowest-indexed half edge in S. We consider the process up to some time τ, where τ is the minimum time t at which either $X_t > \beta M$ or $M_t \leq \left(1 - \frac{\varepsilon}{4}\right) M$, for $\beta = 10^{-6}\varepsilon^2$. The next two lemmas show that moving the vertices in H^+ to $\overline{S}$ does not significantly affect the exploration process. That is, since $d(H^+)$ is small, this the expected number of open edges gained at each step remains positive even for this restricted S.

Lemma 6.2.4. Suppose that $R_S \geq \varepsilon M$ and $d(H^+) \leq \frac{\varepsilon M}{100}$ for some fixed $\varepsilon > 0$. Suppose $U \subset S$

is a set containing H^+ such that $d(U) < \frac{\varepsilon}{4}M + d(H^+)$. Then

$$\sum_{i \in V \setminus U} d(i)\,(d(i)-2) \geq \frac{2}{3} R_S (1 - o(1)).$$

Proof. We first prove by contradiction that, under the assumptions of this lemma, $d(j_S) \leq \frac{\sqrt{M}}{\log M}$. If $d(j_S) > \frac{\sqrt{M}}{\log M}$, then vertex j_S and all vertices of greater degree are in U. Thus, $d(U) \geq R_S$, which violates the assumptions on U. Therefore $d(j_S) \leq \frac{\sqrt{M}}{\log M}$. We write

$$\begin{aligned}\sum_{i \in V \setminus U} d(i)\,(d(i)-2) = \sum_{i \leq j_S} d(i)(d(i)-2) - \sum_{i \in U, i \leq j_S} d(i)\,(d(i)-2) \\ + \sum_{i \notin U, i > j_S} d(i)(d(i)-2).\end{aligned}$$

To bound the first sum on the right hand side, note that by definition $\sum_{i \leq j_S} d(i)(d(i)-2) > 0$. Let $K = \sum_{i \in U, i \leq j_S} d(i)$, and note that $\frac{\varepsilon}{4}M + d(H^+) \leq \frac{R_S}{3}$. Since $R_S \geq \varepsilon M$, it also follows that $d(j_S) \geq 3$. Thus,

$$\begin{aligned}\sum_{i \in V \setminus U} d(i)\,(d(i)-2) &\geq -\sum_{i \in U, i \leq j_S} d(i)\,(d(i)-2) + (d(j_S)-2) \sum_{i \notin U, i > j_S} d(i) \\ &\geq -K(d(j_S)-2) + (d(j_S)-2)\left(\sum_{i > j_S} d(i) - \sum_{i \in U, i > j_S} d(i)\right) \\ &\geq -K(d(j_S)-2) + (d(j_S)-2)\left(R_S - d(j_S) - \left(\frac{1}{3}R_S - K\right)\right) \\ &\geq \frac{2}{3}\,(d(j_S)-2)\,R_S - d(j_S)^2.\end{aligned}$$

Since $d(j_S) \in \left[3, \frac{\sqrt{M}}{\log M}\right]$ by previous claims, and $R_S \geq \varepsilon M$ by assumption, it follows that $\frac{2}{3}\,(d(j_S)-2)\,R_S - d(j_S)^2 \geq \frac{2}{3}R_S(1-o(1))$. □

The next lemma shows that $(X_t)_{t \geq 0}$ is expected to increase as t increases for all $t \leq \tau$, regardless of the steps that the process has taken up to time t. Observation 6.2.1 implies that the probability that $w = w_t$ is $\frac{d(w)}{M_{t-1}}$ for all $w \notin V_{t-1}$ (and if $w \in V_{t-1}$, the probability that $w_t = w$ is 0). Thus,

$$\mathbb{E}_{t-1}\left[d(w_t) - 2\right] = \sum_{w \notin V_{t-1}} \frac{d(w)}{M_{t-1}}\,(d(w)-2)\,. \tag{6.4}$$

Lemma 6.2.5. Suppose that $R_S \geq \varepsilon M$ and $d(H^+) \leq \frac{\varepsilon M}{100}$ for some fixed $\varepsilon > 0$. For all $t \leq \tau$, $\mathbb{E}_{t-1}\left[d(w_t) - 2\right] \geq \frac{1}{3}\varepsilon$.

Proof. Assume that $X_{t-1} > 0$. If this is not the case, and $X_{t-1} = 0$, then the following forms a lower bound, which is sufficient for the proof. Since $t \leq \tau$, we know that $(1 - \frac{\varepsilon}{4})M \leq M_{t-1} \leq M$.

Thus, we can bound the summation in Equation (6.4) from below:

$$\begin{aligned}
\sum_{w\notin V_{t-1}} \frac{d(w)}{M_{t-1}}\left(d(w)-2\right) &= \sum_{w\notin V_{t-1},d(w)=1} \frac{d(w)}{M_{t-1}}\left(d(w)-2\right) + \sum_{w\notin V_{t-1},d(w)\geq 2} \frac{d(w)}{M_{t-1}}\left(d(w)-2\right) \\
&\geq \sum_{w\notin V_{t-1},d(w)=1} \frac{d(w)}{\left(1-\frac{\varepsilon}{4}\right)M}\left(d(w)-2\right) + \sum_{w\notin V_{t-1},d(w)\geq 2} \frac{d(w)}{M}\left(d(w)-2\right) \\
&= \sum_{w\notin V_{t-1}} \frac{d(w)}{M}\left(d(w)-2\right) - \left(\frac{1}{1-\frac{\varepsilon}{4}}-1\right) \sum_{w\notin V_{t-1},d(w)=1} \frac{d(w)}{M} \\
&= \sum_{w\notin V_{t-1}} \frac{d(w)}{M}\left(d(w)-2\right) - \left(\frac{\varepsilon}{4}+\frac{\varepsilon^2}{16}+O\left(\varepsilon^3\right)\right)(1-\varepsilon) \\
&\geq \sum_{w\notin V_{t-1}} \frac{d(w)}{M}\left(d(w)-2\right) - \frac{\varepsilon}{4}.
\end{aligned}$$

Recall that H^+ has been moved to $\overline{S}$, and thus contributes negatively to this summation. Applying Lemma 6.2.4 with $U = V_{t-1} \cup H^+$ gives that

$$\begin{aligned}
\sum_{w\notin V_{t-1}} \frac{d(w)}{M}\left(d(w)-2\right) - \frac{\varepsilon}{4} &= \sum_{w\notin U} \frac{d(w)}{M}\left(d(w)-2\right) - \frac{d(H^+)}{M} - \frac{\varepsilon}{4} \\
&\geq \frac{2}{3}\frac{R_S}{M}(1-o(1)) - \frac{\varepsilon}{4} - \frac{d(H^+)}{M} \\
&\geq \frac{1}{3}\varepsilon,
\end{aligned}$$

since $d(H^+) \leq \frac{\varepsilon M}{100}$ and $R_S \geq \varepsilon M$. □

At each step of the exploration process, all edges between w_t and V_{t-1} are revealed. Recall that $d'_{S,t}(w)$ is the number of edges between w and V_{t-1} other than $v_t w$, that is, the number of back edges revealed at step t. As mentioned in Equation (6.3), the number of open edges gained at step t is not necessarily $d(w_t)-2$, but is instead $d(w_t)-2-2d'_{S,t}(w_t)$, since the process reveals back edges at each step. However, on average, these back edges only consume a small fraction of the open edges gained at each step. This is proved in the following lemma.

Lemma 6.2.6. For all $t \leq \tau$, $\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right] \leq \frac{1}{12}\mathbb{E}_{t-1}\left[d(w_t)-2\right]$.

Proof. Suppose that $w_t = w$ for an arbitrary $w \notin V_{t-1}$. Upon revealing w, it contains $d(w)-1$ unpaired half edges (since $v_t w$ is an edge). For each unpaired half edge $p \in B(w)$ (the set of half edges in the bin corresponding to vertex w) and each open edge $q \in \mathfrak{X}_{t-1}$, let $A(p,q)$ denote the event that p and q are paired. For each of these events, Observation 6.2.1 implies that

$$\mathbb{P}_{t-1}\left(A(p,q)\middle|\, w_t = w\right) = \frac{1}{M_{t-1}},$$

since q cannot be adjacent to another open edge by definition of the exploration process. For a given half edge $p \in B(w)$, the events $A(p,q_1)$ and $A(p,q_2)$ are mutually exclusive for every pair

$q_1 \neq q_2$ of open edges in $\mathfrak{X}_{t-1}$. Thus, for each $p \in B(w)$,

$$\mathbb{P}_{t-1}\left(\bigcup_{q\in\mathfrak{X}_{t-1}} A(p,q)\middle| w_t = w\right) = \frac{X_{t-1}}{M_{t-1}}.$$

If $t \leq \tau$, this is at most $\frac{\beta M}{M}\frac{100}{99} = \frac{100\beta}{99}$, since $X_{t-1} \leq \beta M$ and $M_{t-1} \geq \frac{99}{100}M$ by the definition of τ (and a sufficiently small choice of ε). There are $d(w)-1$ half edges in $B(w)$ that could be paired to form back edges (since $v_t w$ does not count as a back edge). Thus,

$$\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\middle| w_t = w\right] = \frac{100\beta}{99}(d(w)-1).$$

Finally, comparing this to $\mathbb{E}_{t-1}[d(w_t)-2]$ gives that

$$\begin{aligned}\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right] &= \sum_{w\notin V_{t-1}} \mathbb{E}_{t-1}\left[d'_{S,t}(w)|w_t = w\right]\mathbb{P}\left(w_t = w\right)\\ &= \frac{100}{99}\beta\sum_{w\notin V_{t-1}}\frac{d(w)}{M_{t-1}}(d(w)-1)\\ &= \frac{100}{99}10^{-6}\varepsilon^2\left(\mathbb{E}_{t-1}\left[d(w_t)-2\right]+1\right).\end{aligned}$$

Since $\varepsilon < 1$ and $\mathbb{E}_{t-1}[d(w_t)-2] \geq \frac{1}{3}\varepsilon$, this implies that $\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right] \leq \frac{1}{12}\mathbb{E}_{t-1}[d(w_t)-2]$. □

It still needs to be shown that X_t is sufficiently close to its expectation for all $t \leq \tau$. The proof of this is almost identical to the analogous proof by Joos et al. [83, Lemma 25]. The proof follows from an application of Azuma's inequality, which is given in Theorem A.6. Define $A_t = d(w_t) - \mathbb{E}_{t-1}[d(w_t)]$ and $B_t = d'_{S,t}(w_t) - \mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right]$.

Lemma 6.2.7. With probability $1-o(1)$, $|\sum_{t'\leq t} A_{t'}|$ and $|\sum_{t'\leq t} B_{t'}|$ are both strictly less than $\frac{M}{\log\log M}$ for all $t \leq M$.

Proof. The argument is identical for $|\sum_{t'\leq t} A_{t'}|$ and $|\sum_{t'\leq t} B_{t'}|$, and thus we only give the argument for the former. Note that $\mathbb{E}_{t-1}[A_t] = 0$ for all t. It is straightforward to show that $\sum_{t'\leq t} A_{t'}$ is a martingale with respect to the filtration of partial pairings $(T_t)_{t\geq 0}$:

$$\mathbb{E}_{t-1}\left[\sum_{t'\leq t} A_{t'}\right] = \mathbb{E}_{t-1}[A_t] + \sum_{t'\leq t-1} A_{t'} = \mathbb{E}_{t-1}\left[d(w_t) - \mathbb{E}_{t-1}[d(w_t)]\right] + \sum_{t'\leq t-1} A_{t'} = \sum_{t'\leq t-1} A_{t'}.$$

It follows that $\mathbb{E}\left[\sum_{t'\leq t} A_{t'}\right] = 0$ for all t. Since the maximum degree of a vertex not in H^+ is at most $\frac{\sqrt{M}}{\log M}$, it follows that $|A_t|,\ |B_t| \leq \frac{\sqrt{M}}{\log M}$ for all $t \leq M$. Thus, by Azuma's inequality with $c_k = \frac{\sqrt{M}}{\log M}$ for all $k \leq t$,

$$\mathbb{P}\left(\left|\sum_{t'\leq t} A_t\right| > \frac{M}{\log\log M}\right) \leq 2\exp\left(-\frac{M\log^2 M}{2t(\log\log M)^2}\right) < \exp\left(-\log^{3/2} M\right),$$

where the last inequality holds for all $t \leq M$. Taking the union bound over all such t then gives

that

$$\mathbb{P}\left(\left|\sum_{t'\le t} A_t\right| > \frac{M}{\log\log M} \text{ for some } t \le M\right) < M\exp\left(-\log^{3/2} M\right) = M^{1-\sqrt{\log M}} = o(1).$$

This completes the proof. □

Recall that τ is the minimum time t such that either $X_t > \beta M$ or $M_t \le \left(1-\frac{\varepsilon}{4}\right)M$, for $\beta = 10^{-6}\varepsilon^2$. We show that a.a.s. $X_\tau > \beta M$. Since all open edges belong to the same component of $P[S]$ (specifically, they belong to the active component of the process at time τ), this gets us most of the way to finding a component with $\Theta(M)$ edges in $P[S]$.

Lemma 6.2.8. With probability $1-o(1)$, $X_\tau > \beta M$.

Proof. Recall that for a given input, that is for a given pairing P and permutation of the half edges $\sigma \in S_M$, the process is deterministic. Consider the behaviour of X_t on a given input ω. For all ω, it follows that

$$X_\tau(\omega) \ge \sum_{t\le\tau}(d(w_t(\omega)) - 2) - 2\sum_{t\le\tau} d'_{S,t}(w_t(\omega)), \tag{6.5}$$

where $w_t(\omega)$ is the vertex revealed at step t for input ω. Assume for contradiction that $X_\tau(\omega) \le \beta M$. Then by the definition of τ, it follows that $M_\tau(\omega) \le \left(1-\frac{\varepsilon}{4}\right)M$. Then adding and subtracting $\sum_{t\le\tau}\mathbb{E}_{t-1}\left[d(w_t)-2\right] - 2\sum_{t\le\tau}\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right]$ from the right hand side of (6.5) gives that

$$\begin{aligned} X_\tau(\omega) &\ge \sum_{t\le\tau}(d(w_t)-2) - 2\sum_{t\le\tau} d'_{S,t}(w_t) \\ &= \sum_{t\le\tau}\mathbb{E}_{t-1}\left[d(w_t)-2\right] - 2\sum_{t\le\tau}\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right] + \sum_{t\le\tau} A_t - 2\sum_{t\le\tau} B_t. \end{aligned}$$

By Lemmas 6.2.5 and 6.2.6, it follows that

$$\sum_{t\le\tau}\mathbb{E}_{t-1}\left[d(w_t)-2\right] - 2\sum_{t\le\tau}\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right] \ge \frac{5}{6}\frac{1}{3}\varepsilon\tau = \frac{5\varepsilon\tau}{18}.$$

Let Ω be the set of inputs ω for which $|\sum_{t'\le t} A_{t'}(\omega)|$ and $|\sum_{t'\le t} B_{t'}(\omega)|$ are each less than $\frac{M}{\log\log M}$ for all $t \le M$. By Lemma 6.2.7, $\mathbb{P}(w\in\Omega) = 1-o(1)$. For all $\omega\in\Omega$,

$$X_\tau(\omega) \ge \frac{5\varepsilon\tau}{18} - \frac{3M}{\log\log M} \ge \frac{\tau\varepsilon}{4},$$

where the last inequality holds for sufficiently large M. Recall that for a fixed input ω, $\tau(\omega)$ is deterministic. We partition Ω based on the size of $\tau(\omega)$. First suppose $\tau(\omega) \ge \frac{3}{64}M$. Then it immediately follows that $X_\tau > \beta M$. Now suppose that $\tau(\omega) < \frac{3}{64}M$ and suppose for contradiction that $X_\tau \le \beta M$. This implies that $M_\tau \le \left(1-\frac{\varepsilon}{4}\right)M$, which then implies that $\sum_{t\le\tau} d(w_t) \ge \frac{\varepsilon}{4}M$.

Thus,

$$\begin{aligned} X_\tau(\omega) &\geq \sum_{t\leq\tau}(d(w_t)-2) - 2\sum_{t\leq\tau} d'_{S,t}(w_t) \\ &\geq \sum_{t\leq\tau}(d(w_t)-2) - 2\sum_{t\leq\tau}\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right] - 2\sum_{t\leq\tau} B_t \\ &= \frac{5}{6}\sum_{t\leq\tau}(d(w_t)-2) + \frac{1}{6}\sum_{t\leq\tau}\mathbb{E}_{t-1}\left[d(w_t)-2\right] - 2\sum_{t\leq\tau}\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right] + \frac{1}{6}\sum_{t\leq\tau} A_t - 2\sum_{t\leq\tau} B_t. \end{aligned} \tag{6.6}$$

Lemma 6.2.6 implies that $\frac{1}{6}\sum_{t\leq\tau}\mathbb{E}_{t-1}\left[d(w_t)-2\right] - 2\sum_{t\leq\tau}\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right] \geq 0$, and thus

$$X_\tau(\omega) \geq \frac{5}{6}\sum_{t\leq\tau}\left[d(w_t)-2\right] + \frac{1}{6}\sum_{t\leq\tau} A_t - 2\sum_{t\leq\tau} B_t.$$

For all $\omega\in\Omega$, we know that $\frac{1}{6}\sum_{t\leq\tau} A_t - 2\sum_{t\leq\tau} B_t > -\frac{3M}{\log\log M}$. Then since $\sum_{t\leq\tau} d(w_t) \geq \frac{\varepsilon}{4}M$ it follows that

$$X_\tau(\omega) \geq \frac{5}{6}\left[\sum_{t\leq\tau} d(w_t) - 2\tau\right] - \frac{3M}{\log\log M} \geq \frac{5}{6}M\left(\frac{\varepsilon}{4} - \frac{\varepsilon}{32}\right) - \frac{3M}{\log\log M} > \beta M.$$

Therefore, $X_\tau > \beta M$ for all $\omega\in\Omega$. Since $\mathbb{P}(\omega\in\Omega) = 1-o(1)$, this completes the proof. □

At this point we have proved that a.a.s. there exists a partial pairing of $P[S]$ that contains a connected component with $\Theta(M)$ open edges. Finally, we show that a positive fraction of these open edges are not paired with half edges in vertices in $\overline{S}$, and thus correspond to edges in $P[S]$.

Lemma 6.2.9. If a partial pairing $T=(V,E,\mathcal{X})$ has at least βM open edges, then a.a.s. at least $\beta^2 M$ of these open edges pair with half edges with parent vertices in S.

Proof. Since $R_S \geq \varepsilon M$, there exists some $c\in[\varepsilon,1)$ such that $d(S) = cM$. Take an arbitrary subset K of $\mathcal{X}$ of size βM. We analyse a process of checking the mate of each half edge in this subset sequentially, similar to that used in Lemma 6.2.2. Note however that half edges in $\mathcal{X}$ cannot be paired with other half edges in $\mathcal{X}$, as these would have been revealed as back edges during the exploration process. Suppose that less than $\beta^2 M$ of the previously revealed mates of half edges in K belong to vertices in S_τ, the set of vertices in S that have not been explored at time τ. The definition of τ implies that $d(S_\tau) \geq cM - \frac{\varepsilon}{4}M - d(w_\tau)$. Thus, the probability that the mate of the next half edge checked is not in S is at most

$$1 - \frac{d(S_\tau) - \beta^2 M}{M_\tau} < 1 - \frac{2c}{3},$$

noting that $c\geq\varepsilon$ by definition and $\beta = 10^{-6}\varepsilon^2$. Thus, the probability that less than $\beta^2 M$ of the βM half edges have a pair in S is less than

$$\sum_{i=1}^{\beta^2 M}\binom{\beta M}{i}\left(1-\frac{2c}{3}\right)^{(\beta M - i)} < 2^{\beta M}\exp\left(-\frac{1}{2}\beta c M\right) = o(1).$$

Therefore, a.a.s. at least $\beta^2 M$ of the βM edges have both endpoints in S. □

Proof of Theorem 6.1.1(b). Lemma 6.2.8 states that the partial pairing T_τ a.a.s. has a connected component with more than βM open edges. Lemma 6.2.9 then implies that the corresponding component in $P[S]$ a.a.s. has at least $\beta^2 M$ edges. Therefore, $P[S]$ a.a.s. contains a component with at least $\beta^2 M$ edges. □

6.2.4 Analysis in the subcritical regime

Suppose that $R_S = \sum_{i=j_S}^{n} d(i) \leq \delta M$ for some $\delta \to 0$ as $n \to \infty$. Define $U \subset S$ to be the smallest set of highest degree vertices in S such that $\sum_{i \in U} d(i) > 5\delta^{1/4} M$. That is, define

$$u = \max\left\{ v \in [n] : \sum_{i=v}^{n} d(i) > 5\delta^{1/4} M \right\} \quad \text{and} \quad U = \{u, u+1, \ldots, n\}. \tag{6.7}$$

Since $5\delta^{1/4} M > \delta M$ and $R_S \leq \delta M$, it immediately follows that $u < j_S$. Then define $V_0 = U \cup \{v\}$ for an arbitrary $v \in [n]$. To analyse $(X_t)_{t \geq 0}$ in the subcritical case, we bound it from above using a simpler random variable:

$$X'_t = X'_0 + \sum_{i=1}^{t} [d(w_i) - 2] \tag{6.8}$$

where $X'_0 = \sum_{u \in U \cup \{v\}} d(u)$. This is an upper bound on X_t for all t, since it does not take into account back edges or edges between vertices in V_0. Thus, if $X'_t \leq 0$, then there must exist some time $t' < t$ such that $X_{t'} = 0$. To prove the subcritical part of Theorem 6.1.1, we show that for each choice of v (such that $V_0 = U \cup \{v\}$), $X'_t < 0$ for some $t = o(M)$ with probability $1 - o(M^{-1})$. Therefore, by the union bound a.a.s. there is no $v \in V(P)$ that lies in a component with $\Theta(M)$ edges. The following observation gives us a straightforward but useful lower bound on the size of M_t throughout the exploration process.

Observation 6.2.10. For all $t \geq 1$, $M_t \geq d(\overline{S}_{t-1}) \geq d(\overline{S}) - t$.

This observation simply stems from the idea that at each step of the process, $M_t - M_{t-1} - d(w_t)$, and that everything in $\overline{S}$ has degree 1 by assumption. Thus, for all $t \geq 0$, either $d(\overline{S}_t) = d(\overline{S}_{t-1}) - 1$ or $d(\overline{S}_t) = d(\overline{S}_{t-1})$.

We first prove that the degrees of the vertices in S outside of V_0 are bounded by a slowly growing function of δ; this is formalised in the following lemma.

Lemma 6.2.11. The maximum degree of a vertex in S_{t-1} is at most $\delta^{-1/4}$ for all $t \geq 1$. Furthermore, $X'_0 \leq 7\delta^{1/4} M$.

Proof. We show that there exists a vertex in U with degree at most $\delta^{-1/4}$. Since each vertex outside U cannot have a higher degree than any vertex inside U, this proves the claim. Recall that u is the vertex in U with lowest index in $\boldsymbol{d}$, and thus lowest degree. For the sake of contradiction suppose that the claim of the lemma is not true, that is, assume $d(u) > \delta^{-\frac{1}{4}}$. Then note that $\delta^{-\frac{1}{4}} - 2 \geq \frac{3}{4}\delta^{-\frac{1}{4}}$ asymptotically, and also recall that $U \subset S$ and $u < j_S$. Thus,

$$\sum_{i=u}^{j_S - 1} d(i)\,(d(i) - 2) \geq (d(u) - 2) \sum_{i=u}^{j_S - 1} d(i) \geq \frac{3}{4}\delta^{-1/4} \sum_{i=u}^{j_S - 1} d(i) \geq \frac{3}{4}\delta^{-1/4}(5\delta^{1/4} M - \delta M) \geq 3M.$$

But this means that

$$\sum_{i=1}^{j_S-1} d(i)\,(d(i)-2) = \sum_{i=1}^{u-1} d(i)(d(i)-2) + \sum_{i=u}^{j_S-1} d(i)\,(d(i)-2) \geq -M + 3M = 2M,$$

since the sum $\sum_{i=1}^{u-1} d(i)(\mathsf{d}_S(i)-2)$ is at least $-M$. This is a contradiction, since we know $\sum_{i=1}^{j_S-1} d(i)\,(\mathsf{d}_S(i)-2) \leq 0$ from the definition of j_S. This proves the first claim. The second claim follows from the fact that $\delta^{-1/2} \leq M$ (since $R_S \leq \delta M$ and $R_S \geq 1$), and U contains a vertex of degree at most $\delta^{-1/4}$. Together this implies that

$$X_0' = d(U \cup \{v\}) \leq 5\delta^{1/4}M + 2\delta^{-1/4} \leq 7\delta^{1/4}M.$$

This completes the proof. □

This implies that $d(w_t) \leq \delta^{-1/4}$ for all $t \geq 1$. Now we study how $(X_t')_{t\geq 0}$ changes over time. From the definition of X_t', we know that $X_t' - X_{t-1}' = d(w_t) - 2$. In a similar fashion to the proof of the supercritical case, we study $\mathbb{E}_{t-1}[d(w_t)-2]$ throughout the process. Despite not having to keep track of back edges throughout this process, the argument in this case is more nuanced. This is due to the fact that we do not obtain a bound on $\mathbb{E}_{t-1}[d(w_t)-2]$ that is uniform for all (sufficient) t. This is in contrast to the proof of the supercritical case, where Lemma 6.2.5 bounds $\mathbb{E}_{t-1}[d(w_t)-2]$ from below by a positive constant. In this case, we show that $\mathbb{E}_{t-1}[d(w_t)-2]$ grows more and more negative as the process evolves. The first step is to show that $\mathbb{E}_0[d(w_1)-2]$, the expected number of open edges gained in the first time step, is negative. This forms the base case of our later analysis of how $\mathbb{E}_{t-1}[d(w_t)-2]$ changes over time.

Lemma 6.2.12. If $R_S \leq \delta M$ for $\delta \to 0$, then $\mathbb{E}_0[d(w_1)-2] \leq -4\delta^{1/4}$.

Proof. By definition, V_0 contains U and thus has total degree greater than $5\delta^{1/4}M$. Recall that u is defined to be the lowest-indexed vertex in U. First suppose that U contains only vertices of degree at least 3. Then

$$\sum_{i=u}^{j_S-1} d(i)(d(i)-2) \geq \sum_{i=u}^{j_S-1} d(i).$$

Since $\sum_{i=1}^{j_S-1} d(i)(d(i)-2) \leq 0$, this implies that

$$\sum_{i=1}^{u-1} d(i)(d(i)-2) = \sum_{i=1}^{j_S-1} d(i)(d(i)-2) - \sum_{i=u}^{j_S-1} d(i)(d(i)-2) \leq -5\delta^{1/4}M + \delta M.$$

Recall that $\mathbb{E}_{t-1}[d(w_t)-2]$ is given by the expression in Equation (6.4). Thus, we obtain an upper bound for $\mathbb{E}_0[d(w_1)-2]$:

$$\begin{aligned}
\mathbb{E}_0[d(w_1)-2] &= \sum_{i=1}^{u-1} \frac{d(i)(d(i)-2)}{M_0} - \frac{d(v)}{M_0}(d(v)-2)\mathbb{1}_{\{v\leq u-1\}} \\
&\leq -\frac{9}{2}\delta^{1/4} - \frac{d(v)}{M_0}(d(v)-2)\mathbb{1}_{\{v\leq u-1\}} \\
&\leq -4\delta^{1/4}.
\end{aligned}$$

Now suppose that U contains a vertex of degree 2 or lower. If this is the case, then every vertex outside of U must have degree at most 2. Thus,

$$\begin{aligned}\mathbb{E}_0\left[d(w_1)-2\right] &= \sum_{i=1}^{u-1}\frac{d(i)(d(i)-2)}{M_0} - \frac{d(v)(d(v)-2)}{M_0}\mathbb{1}_{\{v\le u-1\}}\\ &\le -\frac{d(\overline{S})}{M_0} - \frac{d(v)(d(v)-2)}{M_0}\mathbb{1}_{\{v\le u-1\}}\\ &\le -c\end{aligned}$$

for some constant $c>0$. This concludes the proof. □

This bound on its own is insufficient to prove the desired result. Aside from the obvious problem, that this lemma only applies to time $t=1$, we also require that the expected number of open edges decreases more quickly than this bound implies. If $\mathbb{E}_{t-1}\left[d(w_t)-2\right] = -\Theta(\delta^{1/4})$ for all t, then on average it would take $\Theta(M)$ steps to guarantee that $X_t' \le 0$. This is too long, as it is possible that in that time the process revealed a component with $\Theta(M)$ edges or $\Theta(n)$ vertices. However, intuitively we expect that $\mathbb{E}_{t-1}\left[d(w_t)-2\right]$ should decrease over time. This is because vertices with high-degree are more likely to be explored at each step (since the probability of discovering a vertex at time t is proportional to its degree), and thus over time we expect fewer and fewer high-degree vertices to be left unexplored to lift this expectation. On the other hand, Observation 6.2.10 states that $d(\overline{S}_t)$ stays close to its initial value for all $t = o(M)$. We show that this intuition is roughly correct, and that $\mathbb{E}_{t-1}\left[d(w_t)-2\right]$ a.a.s. decreases over time. The following lemma states this idea formally, and is the critical lemma we use to prove the subcritical case of Theorem 6.1.1. Lemma 6.2.12 acts as a base case for the proof of this lemma. For notational convenience we also define

$$f_t := \mathbb{E}_{t-1}\left[d(w_t)-2\right] = \sum_{w\notin V_{t-1}}\frac{d(w)}{M_{t-1}}\left(d(w)-2\right). \tag{6.9}$$

Note that this can also be written as

$$f_t = -1 + \sum_{w\in S_{t-1}}\frac{d(w)}{M_{t-1}}(d(w)-1). \tag{6.10}$$

Both forms are useful at various points in the following proofs. We also define $G_t = f_{t+1} - f_t - \mathbb{E}_{t-1}\left[f_{t+1}-f_t\right]$, to help analyse how f_t changes throughout this process.

Lemma 6.2.13. Define $c := \frac{d(\overline{S})}{8M}$ and suppose $t \le \delta^{1/19}M$. If $(T_{t'})_{t'<t}$ is such that $\left|\sum_{t'<t} G_{t'}\right| \le \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{t}$ for all $t'<t$, then $\mathbb{E}_{t-1}\left[d(w_t)-2\right] \le -\frac{ct}{M_{t-1}}$.

We defer the proof briefly, as it is rather long. Note that since we assume $d(\overline{S}) = \Theta(M)$, it follows that c is bounded away from 0. The following lemma shows that $\left|\sum_{t'<t} G_{t'}\right|$ is a.a.s. sufficiently small to apply Lemma 6.2.13 for all possible t. That is, a.a.s. the expected number of open edges gained at each step decreases at the expected rate.

Lemma 6.2.14. The probability that $\left|\sum_{t'\le t} G_{t'}\right| > \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{t}$ for some $t \le M$ is $o(M^{-2})$.

Proof. $\sum_{t'\le t} G_{t'}$ is a martingale with expectation 0 (the proof of this statement is analogous to the ideas in Lemma 6.2.7). To bound $|G_t|$ for all t, we first express f_{t+1} in terms of f_t for arbitrary

t:

$$f_{t+1} = \frac{1}{M_t} \sum_{w \notin V_t} d(w)\,(d(w)-2) = \frac{M_{t-1}}{M_t} \sum_{w \notin V_t} \frac{d(w)}{M_{t-1}}\,(d(w)-2)\,.$$

Since $V_t = V_{t-1} \cup \{w_t\}$, it follows that

$$f_{t+1} = \frac{M_{t-1}}{M_t}\left[f_t - \frac{d(w_t)(d(w_t)-2)}{M_{t-1}}\right].$$

Thus, the difference can be written as

$$f_{t+1} - f_t = \frac{d(w_t)}{M_t}\left(f_t - (d(w_t)-2)\right). \tag{6.11}$$

Since $d(w_t) \le \delta^{-1/4}$ for all $t \ge 1$, it follows that $|G_t| \le \frac{2\delta^{-1/2}}{M_t}$. Applying Azuma's inequality gives that

$$\mathbb{P}\left(\left|\sum_{t' \le t} G_{t'}\right| > \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{t}\right) < \exp\left(-3\log M\right).$$

Applying a union bound over all $t \le M$ completes the proof. □

Much like in the supercritical case, it also needs to be shown that $d(w_t)-2$ is actually close to $\mathbb{E}_{t-1}\left[d(w_t)-2\right]$, at least when averaged over a large enough time scale. The proof is very similar to that of Lemma 6.2.7, as well as the analogous result from Joos et al. [83, Lemma 25]. The lemma is given here for completeness. Recall the variable $A_t = d(w_t) - \mathbb{E}_{t-1}\left[d(w_t)\right]$ from the proof of the supercritical case. Note that since X_t' does not account for back edges, we do not need to prove the analogous result about the concentration of the number of back edges revealed over time.

Lemma 6.2.15. The probability that $\left|\sum_{t' \le t} A_t\right| > M^{2/3}$ for some $t \le M$ is less than $2e^{-M^{1/4}}$.

Proof. Recall that $\sum_t A_t$ is a martingale with expectation 0 (see Lemma 6.2.7). By Lemma 6.2.11, $|A_t| \le \delta^{-1/4}$ for all $t \le M$. Then by Azuma's inequality and the union bound,

$$\mathbb{P}\left(\left|\sum_{t' \le t} A_t\right| > M^{2/3} \text{ for some } t \le M\right) < 2Me^{-\frac{M^{4/3}}{2\delta^{-1/2}t}} < e^{-M^{1/4}}.$$

□

Proof of Lemma 6.2.13

The proof of this lemma is not entirely straightforward. The main step is given in Lemma 6.2.18, which gives a bound on how $\mathbb{E}_{t-1}\left[d(w_t)-2\right]$ changes at each step in expectation. However, the proof of Lemma 6.2.18 requires bounds on the magnitude of $\mathbb{E}_{t-1}\left[d(w_t)-2\right]$ which are not necessarily true for all $t \le \delta^{1/19}M$ a priori. These bounds are true for a small initial window of time after $t = 0$. Furthermore, applying these lemmas to the first window of time, and recalling the assumption from the lemma about $\sum_{t'} G_{t'}$, ensures that the necessary bounds are true for

the next window of time. Thus, Lemmas 6.2.14, 6.2.16, and 6.2.18 form the base case for a proof by induction of Lemma 6.2.13. The inductive step is then given in Lemma 6.2.19. Define $\lambda = \frac{1}{2}\delta^{1/4}d(\overline{S})$, and given λ define $\mathcal{I}_x = \{t \in \mathbb{Z} \mid (x-1)\lambda < t \leq x\lambda\}$ for $x \in \{1, \dots, \delta^{-1/5}\}$. The sequence $(\mathcal{I}_x)_{x \leq \delta^{-1/5}}$ corresponds to the windows of time mentioned prior. Note that the assumption that $d(\overline{S}) = \Theta(M)$ implies that $\lambda = \Theta(\delta^{1/4}M)$.

Lemma 6.2.16. For all $t \in \mathcal{I}_1$, $f_t < -2\delta^{1/4}$ and $f_{t+1} - f_t \leq \frac{1}{M_{t-1}}$.

Proof. We prove this lemma by induction on t. Note by Lemma 6.2.12 that $f_1 \leq -4\delta^{1/4}$. This proves the first part of the base case where $t = 1$. Now we prove the second half of the statement for time $t = 1$. Since $f_1 \leq -4\delta^{-1/4}$, it follows from (6.11) that

$$\frac{d(w_1)}{M_1}\left(f_1 - (d(w_1) - 2)\right) \leq -\frac{d(w_1)(d(w_1)-2)}{M_0}.$$

This is maximised when $d(w_1) = 1$. Therefore, the second part of the claim holds for $t = 1$. This completes the proof of the base case. Now assume that the statement holds for all $t' < t$. By the induction hypothesis, the definition of $\mathcal{I}_1$, and Observation 6.2.10, it follows that $f_t < -2\delta^{1/4}$:

$$f_t \leq f_{t-1} + \frac{1}{M_{t-1}} \leq f_1 + \sum_{i=1}^{t-1}\frac{1}{M_i} < f_1 + \frac{t-1}{M_{t-1}} < -4\delta^{1/4} + \frac{\delta^{1/4}d(\overline{S})}{2d(\overline{S})}(1 + o(1)) < -2\delta^{1/4}.$$

This proves the first part of the statement, that $f_t < -2\delta^{1/4}$ for all $t \in \mathcal{I}_1$. It follows from (6.11) that

$$\frac{d(w_t)}{M_t}\left(f_t - (d(w_t) - 2)\right) \leq -\frac{d(w_t)(d(w_t)-2)}{M_{t-1}}, \tag{6.12}$$

as $f_t < -2\delta^{1/4}$. The right hand side of (6.12) is also maximised when $d(w_t) = 1$. This completes the proof of the second part of the claim for all $t \in \mathcal{I}_1$. □

We mentioned earlier the intuitive idea that, on average, we expect $\mathbb{E}_{t-1}\left[d(w_t) - 2\right]$ to decrease as t increases. This is formalised in Lemma 6.2.18. Note that Lemma 6.2.16 implies that $f_t < -2\delta^{1/4}$ for all $t \in \mathcal{I}_1$. The following observation is useful in the proof of this lemma.

Observation 6.2.17. For a vector of non-negative integers $\boldsymbol{d}$, it follows that $\sum_w d(w)^2(d(w)-2) \geq \frac{3}{2}\sum_w d(w)(d(w)-2)$.

Lemma 6.2.18. Let $c := \frac{d(\overline{S})}{8M}$ and $t \leq \delta^{1/19}M$. If $f_t < -2\delta^{1/4}$, then either $\mathbb{E}_{t-1}\left[f_{t+1} - f_t\right] \leq -\frac{c}{M_{t-1}}$ or $f_t < -c$.

Proof. We assume that $f_t < -2\delta^{1/4}$. Thus, it follows from (6.12) that, for all possible choices of w_t,

$$f_{t+1} - f_t = -\frac{d(w_t)}{M_t}(d(w_t) - 2) + \frac{d(w_t)}{M_t}f_t < -\frac{d(w_t)}{M_{t-1}}(d(w_t) - 2).$$

Taking the conditional expectation with respect to the partial pairing T_{t-1} gives that

$$\begin{aligned}\mathbb{E}_{t-1}\left[f_{t+1}-f_t\right] &< \mathbb{E}_{t-1}\left[-\frac{d(w_t)}{M_{t-1}}(d(w_t)-2)\right]\\ &= \sum_{w\notin V_{t-1}} \mathbb{P}_{t-1}\left(w_t=w\right)\mathbb{E}_{t-1}\left[-\frac{d(w_t)}{M_{t-1}}(d(w_t)-2)\middle| w_t=w\right]\\ &= \sum_{w\notin V_{t-1}} \frac{d(w)}{M_{t-1}}\left(-\frac{d(w)}{M_{t-1}}(d(w)-2)\right)\\ &= -\sum_{w\in S_{t-1}} \frac{d(w)^2}{M_{t-1}^2}(d(w)-2)+\frac{d(\overline{S}_{t-1})}{M_{t-1}^2}.\end{aligned}$$

We now split the proof into two cases. First suppose that $\sum_{w\in S_{t-1}} d(w)^2(d(w)-2) \geq \frac{5}{4}d(\overline{S}_{t-1})$. Then this implies that

$$\mathbb{E}_{t-1}\left[f_{t+1}-f_t\right] < -\frac{d(\overline{S}_{t-1})}{4M_{t-1}^2} \leq -\frac{c}{M_{t-1}}$$

since Observation 6.2.10 implies that $d(\overline{S}_{t-1})=\Theta(M)$ by the assumptions of Theorem 6.1.1 and the fact that $t=o(M)$. Thus the claim of the lemma holds in this case. Now consider the case where $\sum_{w\in S_{t-1}} d(w)^2(d(w)-2) \leq \frac{5}{4}d(\overline{S}_{t-1})$. Then, by Observation 6.2.17, f_t is bounded by

$$\begin{aligned}f_t=\mathbb{E}_{t-1}\left[d(w_t)-2\right] &= \frac{1}{M_{t-1}}\left[\sum_{w\in S_{t-1}} d(w)(d(w)-2)-d(\overline{S}_{t-1})\right]\\ &\leq \frac{1}{M_{t-1}}\left[\frac{2}{3}\sum_{w\in S_{t-1}} d(w)^2(d(w)-2)-d(\overline{S}_{t-1})\right]\\ &\leq \frac{1}{M_{t-1}}\left[\frac{5}{6}d(\overline{S}_{t-1})-d(\overline{S}_{t-1})\right]\\ &< -c.\end{aligned}$$

This completes the proof. □

Lemma 6.2.16 implies that Lemma 6.2.18 applies for all $t\in\mathcal{I}_1$. The next step is to show that this lemma applies for all $t\leq\delta^{1/19}M$. We do so by analysing each set of $\lambda := \frac{1}{2}\delta^{1/4}d(\overline{S})$ time steps inductively. We give the first step here to help build some intuition. Recall that $\mathcal{I}_x=\{t\in\mathbb{Z}\mid (x-1)\lambda<t\leq x\lambda\}$. Suppose that there exists some $t\in\mathcal{I}_1$ such that $f_t<-c$, for c defined in Lemma 6.2.13. Then Lemma 6.2.16 implies that, for all $s\in[t,\delta^{1/20}M]$,

$$f_s \leq f_t+\sum_{i=t}^{s-1}\frac{1}{M_{i-1}} \leq -c+\frac{\delta^{1/4}M}{M_s} = -(c-o(1)). \tag{6.13}$$

Thus, it follows that $f_{\lfloor\lambda\rfloor+1}<-4\delta^{1/4}$. On the other hand, if there does not exist some $t\in\mathcal{I}_1$ such that $f_t<-c$, then Lemma 6.2.18 implies that $\mathbb{E}_{t-1}\left[f_{t+1}-f_t\right]\leq-\frac{c}{M_{t-1}}$ for all $t\leq\lambda$. Then, since

we assume that $\left|\sum_{t\le\lambda} G_t\right|$ is small, we bound $f_{\lfloor\lambda\rfloor+1}$ by

$$\begin{aligned}
f_{\lfloor\lambda\rfloor+1} &= f_{\lfloor\lambda\rfloor+1} - f_{\lfloor\lambda\rfloor} + f_{\lfloor\lambda\rfloor} - \cdots - f_1 + f_1 \\
&= \sum_{t=1}^{\lfloor\lambda\rfloor}(f_{t+1} - f_t) + f_1 + \sum_{t=1}^{\lfloor\lambda\rfloor}\mathbb{E}_{t-1}\left[f_{t+1} - f_t\right] - \sum_{t=1}^{\lfloor\lambda\rfloor}\mathbb{E}_{t-1}\left[f_{t+1} - f_t\right] \\
&= \sum_{t=1}^{\lfloor\lambda\rfloor}\mathbb{E}_{t-1}\left[f_{t+1} - f_t\right] + \sum_{t=1}^{\lfloor\lambda\rfloor} G_t + f_1 \\
&\le -\sum_{t=1}^{\lfloor\lambda\rfloor}\frac{c}{M_{t-1}} + \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{\delta^{1/4}M} + f_1 \\
&< -4\delta^{1/4},
\end{aligned}$$

as $f_1 \le -4\delta^{1/4}$ and the rest of the terms are negative overall. Thus, $f_t \le -2\delta^{1/4}$ and $f_{t+1} - f_t \le 1/M_{t-1}$ for all $t \in \mathcal{I}_2$ (using the same proof as Lemma 6.2.16), and thus Lemma 6.2.18 applies to each $t \in \mathcal{I}_2$. By assuming $\left|\sum_{t'\le t} G_{t'}\right|$ is small for all $t \le M$, we can apply the same ideas to $\mathcal{I}_3$. We apply this idea inductively to get the following lemma, which shows that the claims made about f_t for $t \in \mathcal{I}_1$ extend to all $t \le \frac{1}{2}\delta^{1/20}d(\overline{S})$.

Lemma 6.2.19. If $(T_t)_{t\ge1}$ is such that $\left|\sum_{t'<t} G_{t'}\right| \le \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{t}$ for all $t' < t$, then the inequalities $f_t \le -2\delta^{1/4}$ and $f_{t+1} - f_t \le \frac{1}{M_{t-1}}$ hold for all $t \le \frac{1}{2}\delta^{1/20}d(\overline{S})$.

Proof. Let $x \in \{1, \dots, \delta^{-1/5}\}$. We prove this lemma by proving the following statement by induction on x: if $\left|\sum_{t'<t} G_{t'}\right| \le \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{t}$, then $f_{\lfloor(x-1)\lambda\rfloor+1} \le -4\delta^{1/4}$ and $f_{t+1} - f_t \le \frac{1}{M_{t-1}}$ for all $t \in \mathcal{I}_x$ for all $x \le \delta^{-1/5}$. Lemma 6.2.16 forms the base case $x = 1$.

Now suppose the claim holds for all $x' \le x$ for some $x \ge 1$; we next prove that the claim holds for $x+1$. The assumption that the claim holds for $x' \le x$ implies that $f_t \le -2\delta^{1/4}$ for all $t \le x\lambda$. Thus, we can apply Lemma 6.2.18. Writing $f_{\lfloor x\lambda\rfloor+1}$ as a telescoping sum and applying our assumption on $\sum_{t'<t} G_{t'}$ gives that

$$\begin{aligned}
f_{\lfloor x\lambda\rfloor+1} &= f_{\lfloor x\lambda\rfloor+1} - f_{\lfloor x\lambda\rfloor-1} + f_{\lfloor x\lambda\rfloor-1} - \cdots - f_1 + f_1 \\
&= \sum_{t=1}^{\lfloor x\lambda\rfloor} G_t + \sum_{t=1}^{\lfloor x\lambda\rfloor}\mathbb{E}_{t-1}\left[f_{t+1} - f_t\right] + f_1 \\
&\le \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{(x+1)\lambda} - \sum_{t=1}^{\lfloor x\lambda\rfloor}\frac{c}{M_{t-1}} + f_1 \\
&\le -4\delta^{1/4},
\end{aligned} \tag{6.14}$$

since the first two terms are non-negative overall and Lemma 6.2.12 states that $f_1 \le -4\delta^{1/4}$. Since $f_{\lfloor x\lambda\rfloor+1} \le -4\delta^{1/4}$, an identical proof to Lemma 6.2.16 (applying (6.12)) then shows that $f_{t+1} - f_t \le 1/M_{t-1}$ for all $t \in \mathcal{I}_{x+1}$. This completes the proof. □

With this, we finally prove Lemma 6.2.13.

Proof of Lemma 6.2.13. Suppose $(T_t)_{\ge1}$ is such that $\left|\sum_{t'<t} G_{t'}\right| \le \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{t}$ for all t. We consider two cases based on the cases outlined in Lemma 6.2.18. Suppose $f_s < -c$ for some $s \le t$, where $c = \Theta(1)$ is as defined in the lemma statement. By Lemmas 6.2.16 and 6.2.19, it follows that

$f_t \le f_s + o(1)$ for $t \le \delta^{1/20}M$, and the claim of the lemma follows immediately. Now suppose that there is no $s \le t$ such that $f_s < -c$. Then by Lemmas 6.2.18 and 6.2.19, it follows that

$$f_t \le \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{t} - 4\delta^{1/4} - \frac{ct}{M_{t-1}}.$$

Since our choice of δ is such that $\delta = \omega(M^{-1/2})$, we know that $\frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{\delta^{1/20}M} < 4\delta^{1/4}$. Thus, the claim of the lemma follows. □

Lemma 6.2.20. With probability $1 - o(M^{-2})$, there exists some time $t \le \delta^{1/18}M$ such that $X_t = 0$.

Proof. Consider the set of inputs (pairings $P \in \mathcal{C}(\boldsymbol{d})$ with half edge orderings $\sigma \in S_M$) such that for all $t \le M$,

$$\left|\sum_{t' \le t} A_t\right| \le M^{2/3} \text{ and } \left|\sum_{t' \le t} G_{t'}\right| \le \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{t}. \tag{6.15}$$

By Lemmas 6.2.14 and 6.2.15, the fraction of inputs that do not satisfy this is $o(M^{-2})$. Thus, it suffices to show that for every input that does satisfy these inequalities, $X_t = 0$ for some $t \le \delta^{1/18}M$. For each such input, consider the value of X'_T where $T := \delta^{1/18}M$. Express X'_T as

$$\begin{aligned} X'_T &= X'_0 + \sum_{i=1}^{T}(d(w_i) - 2) \\ &= X'_0 + \sum_{t=1}^{T}\mathbb{E}_{t-1}\left[d(w_t) - 2\right] + \sum_{t=1}^{T}(d(w_t) - \mathbb{E}_{t-1}\left[d(w_t)\right]) \\ &= X'_0 + \sum_{t=1}^{T}\mathbb{E}_{t-1}\left[d(w_t) - 2\right] + \sum_{t=1}^{T}A_t. \end{aligned}$$

Lemma 6.2.11 states that $X'_0 \le 7\delta^{1/4}M$. By the assumption given in (6.15), we know that $\sum_{t=1}^{T}A_t \le M^{2/3}$, and also Lemma 6.2.13 implies that $\mathbb{E}_{t-1}\left[d(w_t) - 2\right] \le -ct/M_{t-1}$ for all $t \le T$. Thus, for all such inputs,

$$\begin{aligned} X'_T &\le 7\delta^{1/4}M + M^{2/3} - \frac{cT(T-1)}{2M} \\ &\le 7\delta^{1/4}M + M^{2/3} + \frac{1}{2M_T}(4\sqrt{T\delta^{-1}\log M})^2 - \frac{c\delta^{1/9}M^2}{3M} \\ &\le 7\delta^{1/4}M + M^{2/3} + 9\delta^{-17/18}\log M - \delta^{1/8}M \\ &< 0. \end{aligned}$$

Therefore, since $X_t \le X'_t$ for all t, there exists some $t \le T$ such that $X_t = 0$. Since the inequalities in (6.15) are both satisfied with probability $1 - o(M^{-2})$, this completes the proof. □

The proof of Theorem 6.1.1(a) follows quickly from Lemma 6.2.20.

Proof of Theorem 6.1.1(a). Recall that $V_0 = U \cup \{v\}$ for an arbitrary $v \in S$. Taking the union bound over all choices of $v \in S$ and applying Lemma 6.2.20 implies that, for every choice of $v \in S$,

a.a.s. there exists some $t \leq \delta^{1/18}M$ such that $X_t = 0$. This implies that a.a.s. all components containing vertices in V_0 are explored in at most $\delta^{1/18}M$ steps of the exploration process. It also implies that the total degree of all vertices in every component of $P[S]$ explored at time $\delta^{1/18}M$ is at most $2\delta^{1/18}M$, as the fact that $X'_{\delta^{1/18}M} < 0$ implies that

$$d(V_0) + \sum_{t=1}^{\delta^{1/18}M} d(w_t) < 2\delta^{1/18}M.$$

This means that a.a.s. every component of $P[S]$ contains at most $2\delta^{1/18}M$ edges. This completes the proof. □

6.3 From size to order

In this section we prove Theorem 6.1.2. First we prove that if $(\boldsymbol{d}, S)$ satisfies the conditions of Theorem 6.1.2 and $R_S = o(M)$, then $M = \Theta(|S|)$. The proof method is borrowed from a similar result by Joos et al. [83].

Lemma 6.3.1. If $d(S) = \Theta(M)$ and $R_S = o(M)$, then $M = \Theta(|S|)$.

Proof. Suppose for contradiction that $d(S) = \Theta(M)$, $R_S = o(M)$, but $M = |S|f(|S|)$ for some $f = \omega(1)$. By assumption, there exists some $\gamma = \Theta(1)$ such that $d(S) = \gamma M$ and $d(\overline{S}) = (1-\gamma)M$. So the average degree in G of vertices in S is $\gamma|S|f(|S|)/|S| = \gamma f(|S|)$. The Cauchy-Schwarz inequality then implies that

$$\sum_{i \in S} d(i)\,(d(i)-2) \geq |S|\gamma f(|S|)\,(\gamma f(|S|) - 2) = \Omega(|S|f(|S|)^2). \tag{6.16}$$

Let (U_i, W_i) denote a pair of subsets where U_i contains the i lowest-degree vertices in S, and $W_i = \{i\} \cup S \backslash U_i$. There must be some $k \leq |S|$ such that $d(U_k) = \Theta(M)$ and $d(W_k) = \Theta(M)$. This implies that U_k and W_k have average degree $\Omega(nf(n)/|U_k|)$ and $\Omega(nf(n)/|W_k|)$ respectively. Then analogously to Equation (6.16),

$$\sum_{i \in U_k} d(i)\,(d(i)-2) = \Omega\left(nf(n)^2\right) = \omega(d(\overline{S})).$$

Thus, $j_S < k$. Therefore, $R_S > d(W_k) = \Theta(M)$. This is a contradiction. The claim of the lemma follows. □

Since $|S| \leq n \leq M$, this also implies that $|S| = \Theta(n)$ under these assumptions.

Corollary 6.3.2. If $d(S) = \Theta(M)$ and $R_S = o(M)$, then a.a.s. $P[S]$ does not contain a component with $\Theta(|S|)$ vertices.

Proof. Lemma 6.2.20 states that a.a.s. there exists some $t \leq \delta^{1/18}M$ such that $X_t = 0$. Thus, a.a.s. every component of $P[S]$ contains at most $\delta^{1/18}M + |V_0|$ vertices. Since $V_0 = U \cup \{v\}$, for a uniformly random $v \in S$ and the preprocessing set U defined in (6.7), it follows that a.a.s. every component of $P[S]$ contains $o(M)$ vertices. By Lemma 6.3.1, $M = \Theta(|S|)$, and thus a.a.s. every component contains $o(|S|)$ vertices. □

Lemma 6.3.3. If $d(S) = \Theta(M)$, and $R_S = \Theta(M)$, then a.a.s. $P[S]$ contains a component with $\Theta(M)$ edges and $\Theta(|S|)$ vertices.

Proof. Note that if $|S| = O(1)$, the result is trivially true, so we assume that $|S| = \omega(1)$. Theorem 6.1.1 implies that a.a.s. $P[S]$ has a connected component with $\Theta(M)$ edges. Let $f(M)$ be the total degree of the largest component of $P[S]$. Suppose that (U, W) is a partition of S such that U has total degree $f(M)$ (specifically, Lemma 6.2.9 implies that $f(M) \geq \beta^2 M$), and $W = S \backslash U$. Thus, $d(W) = d(S) - f(M)$. Lemma 6.2.2 states that the probability that no edge exists between some vertex in U and some vertex in W in a uniformly random pairing P is at most

$$\exp\left(-\frac{1}{2M} d(U) d(W)\right).$$

Since $U \subset S$, there are $\binom{|S|}{k}$ ways to choose a k-subset of vertices in S. Then note that if U contains at most $\gamma|S|$ vertices, for some $\gamma < 1$, then W contains at least $(1-\gamma)|S|$ vertices. Since each of these vertices has degree at least 1 in G, the lower bound $|W| \geq (1-\gamma)|S|$ implies that $d(W) \geq (1-\gamma)|S|$. Then performing a union bound over all possible sets U where $|U| \leq \gamma|S|$ gives

$$\begin{aligned}\mathbb{P}\left(|U| \leq \gamma|S|\right) &\leq \sum_{k=1}^{\gamma|S|} \exp\left(-\frac{1}{2M} f(M)(1-\gamma)|S|\right) \binom{|S|}{k} \\ &< \exp\left(-\frac{1}{2M} f(M)(1-\gamma)|S|\right) \sum_{k=1}^{\gamma|S|} \binom{|S|}{k}.\end{aligned}$$

Since $\sum_{i=0}^{x} \binom{n}{i} \leq \left(\frac{ne}{x}\right)^x$, the summation in this expression is at most $(e/\gamma)^{\gamma|S|}$, and thus

$$\mathbb{P}\left(|U| \leq \gamma|S|\right) < \exp\left(-\frac{1}{2M} f(M)(1-\gamma)|S| + \gamma|S|(1 - \log\gamma)\right). \tag{6.17}$$

Note that $f(M) = \Theta(M)$ and $\gamma(1 - \log\gamma) \to 0$ as $\gamma \to 0$. Thus, there exists some small constant γ such that this bound is at most $\exp\left(-\frac{1}{3}\gamma|S|\right)$. This implies that the probability that there exists a partition (U, W) of S that induces an empty cut satisfying $d(U) = \Theta(M)$ and $|U| < \gamma|S|$ is at most $\exp\left(-\frac{1}{3}\gamma|S|\right)$, which is $o(1)$ since we can assume that $|S| = \omega(1)$.

Thus, a.a.s. $P[S]$ contains a component with $\Theta(M)$ edges, and P contains no partition (U, W) of S with an empty cut such that $d(U) = \Theta(M)$ and $|U| \leq \gamma|S|$. This implies that if $P[S]$ contains a component with $\Theta(M)$ edges, then a.a.s. this component has at least $\gamma|S|$ vertices. Therefore, a.a.s. $P[S]$ contains a component with $\Theta(M)$ edges and $\Theta(|S|)$ vertices. □

Chapter 7

Exploration approach in the random graph model

In this chapter we adapt the results from Chapter 6 to give analogous results about giant components of induced subgraphs of uniformly random graphs with a given degree sequence. We give sufficient conditions on $\boldsymbol{d}$ and S such that the thresholds given in Theorems 6.1.1 and 6.1.2 also apply to induced subgraphs of random graphs with degree sequence $\boldsymbol{d}$. This forms the second part of the analysis of the exploration method as outlined in Chapter 1. The model $\mathcal{G}(\boldsymbol{d})$ is considerably more nuanced than the configuration model, as the existence of one edge is dependent on the existence or non-existence of every other edge in the graph. One of the main challenges and limitations in adapting these results is determining the various adjacency probabilities that the configuration model gave us for free, such as Observation 6.2.1. We use the switching method extensively throughout this chapter to determine upper and lower bounds on probabilities of two vertices being adjacent. In the context of an exploration process on $\mathcal{G}(\boldsymbol{d})$, we also use the switching method to determine bounds on the probability of exploring a particular vertex at a given step of the process.

We cannot hope to recover the threshold for all sequences $\boldsymbol{d}$ and all subsets S that we cover in Theorems 6.1.1 and 6.1.2. In part, this is due to the limitations of the switching method, particularly in the presence of high-degree vertices. However, in part this is due to different underlying behaviour between the two models. The following example shows that Theorems 6.1.1 and 6.1.2 do not carry over to the random graph model $\mathcal{G}(\boldsymbol{d})$ in their complete generality.

Example 7.0.1. Consider the sequence $\boldsymbol{d} = (4, 4, \dots, 4, n-1, n-1)$, and consider the set $S = \{1, \dots, n-2\}$; that is, S contains exactly the degree 4 vertices. In the random graph model, $G[S]$ is a uniformly random 2-regular graph, and thus by Theorem 2.2.15 the order (and thus size) of the largest component is not almost surely linear or almost surely sublinear. However, by Theorem 6.1.1, if P is a uniformly random pairing with this degree sequence, then $P[S]$ a.a.s. contains a component with a linear number of edges and vertices.

This is because the neighbourhoods of high-degree vertices behave very differently in the random graph model to in the configuration model. This becomes exceptionally problematic when these vertices are in $\overline{S}$, as the way that they affect the degrees of vertices in S is not easily predictable using the switching method. One notable difference in this chapter is that all of the results in the random graph model have extra conditions on the total degree of the high-degree

vertices in G, and in particular in $\overline{S}$, the set which is deleted. In the configuration model, this was not an issue: a deleted degree n vertex is, in some sense, equivalent to n deleted degree 1 vertices, and so we could assume everything in $\overline{S}$ had degree 1. This is absolutely not the case in the random graph model, and thus we impose extra conditions to ensure that the important parts of the configuration model proof translate to this model. These conditions are likely not all necessary, and later in this chapter, as well as in the following chapter, we discuss some of the possible relaxations that could be true and possible to prove. However, it is worth noting from the outset that the threshold in $\mathcal{G}(\boldsymbol{d})$ cannot be the exact same as the threshold in $\mathcal{C}(\boldsymbol{d})$ for all pairs $(\boldsymbol{d}, S)$ for which Theorems 6.1.1 and 6.1.2 apply.

In the following section, we give our main results, describing for which pairs $(\boldsymbol{d}, S)$ threshold in $\mathcal{C}(\boldsymbol{d})$ carries over to $\mathcal{G}(\boldsymbol{d})$. We then give the proof for the case where $M/n \to \infty$, where M is the total degree of the sequence $\boldsymbol{d}$. This part of the proof does not rely on the exploration process, and allows us to focus the exploration process on the case where $M = \Theta(n)$. In Section 7.3, we define an exploration process which is equivalent to the one defined in Chapter 6 for the random graph model. Next, we use this to determine the threshold for the a.a.s. (non-)existence of a component with $\Theta(M)$ edges. We then determine, under slightly stricter conditions on $(\boldsymbol{d}, S)$, the threshold for the a.a.s. (non-)existence of a component with $\Theta(|S|)$ vertices. At the end of this chapter, we also give some examples of sequences $\boldsymbol{d}$ and subsets S which highlight the need for these stricter conditions.

7.1 Main results

Let $\boldsymbol{d}$ be an n-element graphical sequence. We recall some notation from the start of Chapter 6, in particular 6.1. Let $S \subset [n]$ and $\overline{S} = [n] \backslash S$, and recall from Equation (6.1) the definition of $\mathsf{d}_S(\cdot)$, where

$$\mathsf{d}_S(w) = \begin{cases} d(w) & w \in S, \\ 1 & w \notin S. \end{cases}$$

Again, without loss of generality we assume that the sequence $\boldsymbol{d}$ is ordered in non-decreasing order with respect to $d(w)\,(\mathsf{d}_S(w) - 2)$. For such a sequence $\boldsymbol{d}$ and set S, recall from (6.2) the definitions of $j_S(\boldsymbol{d})$ and $R_S(\boldsymbol{d})$, as well as the total degree $M(\boldsymbol{d})$:

$$\begin{aligned} j_S(\boldsymbol{d}) &= \min\left(\left\{j : j \in [n] \text{ and } \sum_{i \le j} d(i)\,(\mathsf{d}_S(i) - 2) > 0\right\} \cup \{n\}\right), \\ R_S(\boldsymbol{d}) &= \sum_{i=j_S}^{n} d(i), \\ M(\boldsymbol{d}) &= \sum_{i=1}^{n} d(i). \end{aligned}$$

We abbreviate these by omitting the $\boldsymbol{d}$ in most cases.

Define

$$H^+ = \left\{ v \in [n] : d(v) > \delta\sqrt{M} \right\}. \tag{7.1}$$

for some $\delta \to 0$ arbitrarily slowly. Define $S^+ := S \cap H^+$, and $\overline{S}^+ := \overline{S} \cap H^+$. Specifically, we choose δ such that it satisfies certain conditions given later (specifically in Remark 7.3.1). For the purposes of our proof, we always assume that $\delta = \omega(\log^{-1} M)$. This condition is an artefact of the proof, where in some cases it is not necessary or beneficial for δ to be minimal. However, since we do not optimize this proof for the specific size or order of the largest component, it does not meaningfully affect the results to assume this.

Theorem 7.1.1. Let $\boldsymbol{d}$ be an n-element graphical sequence and $S \subset [n]$. Let $\delta = o(1)$ be some function such that $d(\overline{S}^+) = o(M)$. Suppose that $d(S)$, $d(\overline{S}) = \Theta(M)$ and $R_S = o(M)$. Consider a uniformly random element $G \in \mathcal{G}(\boldsymbol{d})$.

(a) The probability that $G[S]$ contains a component of size $\Theta(M)$ is $o(1)$.

(b) The probability that $G[S]$ contains a component of order $\Theta(|S|)$ is $o(1)$.

In the supercritical case, it is not always true that a component with $\Theta(M)$ edges a.a.s. has $\Theta(|S|)$ vertices. As we discuss later, the problem arises from the presence of high-degree vertices in $\overline{S}$. By placing a tighter restriction on $d(\overline{S}^+)$, we can show the a.a.s. existence of a component with a positive fraction of the vertices of $G[S]$.

Theorem 7.1.2. Let $\boldsymbol{d}$ be an n-element graphical sequence and $S \subset [n]$. Let $\delta = o(1)$ be some function such that $d(\overline{S}^+) = o(M)$. Suppose that $d(\overline{S}) = \Theta(M)$ and $d(S \backslash S^+) = \Theta(M)$. Suppose that $R_S = \Theta(M)$. Consider a uniformly random element $G \in \mathcal{G}(\boldsymbol{d})$.

(a) The probability that $G[S]$ contains a component of size $\Theta(M)$ is $1 - o(1)$.

(b) Suppose further that $d(\overline{S}^+) = o(n)$. Then the probability that $G[S]$ contains a component of size $\Theta(M)$ and order $\Theta(|S|)$ is $1 - o(1)$.

Again, we refer to the case in Theorem 7.1.1 as the subcritical case (or regime), and the case in Theorem 7.1.2 as the supercritical case. The requirement that $d(S \backslash S^+) = \Theta(M)$ is somewhat artificial, and can most likely be relaxed. If $d(S) = \Theta(M)$ and $d(S \backslash S^+) = o(M)$, then S^+ has total degree $\Theta(M)$. The vertices in S^+ are intuitively likely to be adjacent to each other and form a giant component. However, this component may a.a.s. have $o(M)$ edges, or even have $o(M)$ edges always if $|S^+| = o(\sqrt{M})$. To see this, consider the following example.

Example 7.1.3. Consider a pair $(\boldsymbol{d}, S)$ such that $M = 100n$, $S \backslash S^+$ is empty, and S^+ contains $n^{1/3}$ vertices of degree $n^{2/3}$. Then it is not hard to show that $R_S(\boldsymbol{d}) \geq n(1 - o(1))$. However, the graph $G[S]$ contains $O(n^{2/3})$ edges.

Our results do not consider cases like this. By enforcing that the total degree of the "small" vertices in S is not a vanishingly small fraction of $d(S)$, we remove cases where $G[S]$ has $o(M)$ edges always. We note that the condition that $|S \backslash S^+| = \Theta(|S|)$ is not equivalent to the condition that $d(S \backslash S^+) = \Theta(M)$, as the vertices in S^+ can have much higher degree; the condition we impose is more strict.

The extra assumption on $d(\overline{S}^+)$ in Theorem 7.1.2(b) is important. As we discuss later (in Section 7.8), there are $(\boldsymbol{d}, S)$ pairs where $R_S = \Theta(M)$, $d(\overline{S}^+) = \omega(n)$, and $d(\overline{S}^+) = o(M)$ where $G[S]$ a.a.s. does not contain a component with $\Theta(|S|)$ vertices. This highlights an important difference between this result and the result of Joos et al. [83], as well as this result and Theorems 6.1.1 and 6.1.2. In a uniformly random graph with degree sequence $\boldsymbol{d}$, the a.a.s. existence of a giant component with $\Theta(M)$ edges also guaranteed the a.a.s. existence of a component with $\Theta(n)$ vertices. In the induced subgraph case, there are sequences $\boldsymbol{d}$ and subsets S such that $G[S]$ a.a.s. has a component with a linear number of edges, but no component with a linear (in $|S|$) number of vertices. With this distinction in mind, we sometimes refer to edge-giant components or vertex-giant components, which are components with $\Theta(M)$ edges or $\Theta(|S|)$ vertices respectively.

One important note is that the proof for the case where $M = \omega(n)$ differs substantially from the configuration model. In the case that $M = \omega(n)$, under the restrictions that we impose on $(\boldsymbol{d}, S)$ in the random graph model (that is, $d(S)$, $d(S \backslash S^+)$, $d(\overline{S}) = \Theta(M)$, and $d(\overline{S}^+) = o(M)$), we show that the induced graph $G[S]$ a.a.s. has a component with $\Theta(M)$ edges. Specifically, we prove the following result.

Lemma 7.1.4. Suppose that $d(S \backslash S^+) = \Theta(M)$, $d(\overline{S}) = \Theta(M)$, $d(\overline{S}^+) = o(M)$, and $M = \omega(|S|)$. Then $R_S = \Theta(M)$ and a.a.s. $G[S]$ contains a component with $\Theta(M)$ edges.

This reduces proving Theorem 7.1.1 and Theorem 7.1.2(a) to the case where $M = \Theta(n)$. In this case, we adapt the configuration model proof and analyse a similar exploration process. This extra assumption on M is vital to the analysis of the exploration process, particularly because it implies a stronger assumption about $d(\overline{S}^+)$: we can suppose that $d(\overline{S}^+) = o(n)$. This makes the switching analysis much more tractable despite the presence of high-degree vertices in $\overline{S}$. This is useful, and in fact necessary for our methods, when analysing the probability of exploring a particular vertex at time t.

7.2 The case where $M = \omega(n)$

Suppose that $d(S \backslash S^+) = \Theta(M)$, $d(\overline{S}) = \Theta(M)$, and $d(\overline{S}^+) = o(M)$. Lemma 6.3.1, which states that $M = \Theta(|S|)$ if $R_S = o(M)$ and $d(S) = \Theta(M)$, is model-agnostic and applies to the random graph model without modification. Thus, the contrapositive of Lemma 6.3.1 implies that if $M = \omega(n)$ (and thus $M = \omega(|S|)$), then $R_S = \Theta(M)$. So we need to show that if $M = \omega(n)$, a.a.s. $G[S]$ contains a component with $\Theta(M)$ edges. Here we give the proof of this.

Below we give the edge-centric version of the giant component result of Joos et al. [83], since we also use this in the proof of Lemma 7.1.4. This result is (an immediate corollary of) Theorem 9 of their paper, and is an intermediate result used to prove their main results.

Theorem 7.2.1. ([83], Theorem 9) Let $\boldsymbol{d}$, $j_{\boldsymbol{d}}$, $R(\boldsymbol{d})$, and $\widetilde{M}(\boldsymbol{d})$ be as defined in Theorem 2.2.15. Then for any positive constant $\varepsilon > 0$, there is a $\gamma > 0$ such that if $\boldsymbol{d}$ is a well-behaved graphical sequence with $R(\boldsymbol{d}) \geq \varepsilon \widetilde{M}(\boldsymbol{d})$, then the probability that a random graph with degree sequence $\boldsymbol{d}$ has a component of size at least $\gamma \widetilde{M}(\boldsymbol{d})$ is $1 - o(1)$.

Recall from Proposition 3.1.1 that conditional on the induced degree of each vertex in S (that is, the degree sequence of $G[S]$), the graph $G[S]$ is distributed as a uniformly random graph with that degree sequence. We show that, under the assumptions of Lemma 7.1.4, the graph $G[S]$ a.a.s.

has total degree $\omega(n)$. Then by applying Claim 2.2.17 and Theorem 7.2.1, it follows that a.a.s. $G[S]$ has an edge-giant component. Now we give the details. Let $E_G(A, B)$ be the set of edges between A and B in the graph G, for disjoint sets A and B.

Lemma 7.2.2. Suppose that $d(S\backslash S^+) = aM$ for some $a = \Theta(1)$, $d(\overline{S}) = \Theta(M)$, and $d(\overline{S}^+) = o(M)$. Then a.a.s. $G[S]$ contains at least $10^{-6}a^2M$ edges with at least one end incident to $S\backslash S^+$.

Proof. We prove the lemma by partitioning the probability space $\mathcal{G}(\boldsymbol{d})$ into disjoint parts and analysing each of them in turn. First consider the set of graphs G such that $|E_G(S^+, S\backslash S^+)| \geq \frac{1}{100}aM$. Trivially it follows that $G[S]$ contains at least $10^{-6}a^2M$ edges with one end incident to $S\backslash S^+$. Now consider the graphs G such that $|E_G(S^+, S\backslash S^+)| < \frac{1}{100}aM$. Let A be the set of all such $G \in \mathcal{G}(\boldsymbol{d})$. By the assumptions on G and on $d(\overline{S}^+)$, each $G \in A$ contains at least $aM - \frac{1}{100}aM - d(\overline{S}^+)$ edges with one end in $S\backslash S^+$ and the other end in $[n]\backslash H^+$, which means at least $\frac{49}{50}aM$ such edges exist. We define a switching from A to A to show that a.a.s. at least $10^{-6}a^2M$ of these edges have both ends in $S\backslash S^+$ (note that $a^2 < a < 1$).

Let $A_i \subset A$ be the set of $G \in A$ such that $|E(G[S\backslash S^+])| = i$. We define a switching between A_i and A_{i+1} as follows. Suppose $G \in A_i$. Choose an edge $ux \in E(G)$ such that $u \in S\backslash S^+$ and $x \in \overline{S}\backslash\overline{S}^+$. Choose a second edge $vy \in E(G)$ such that $v \in S\backslash S^+$, $y \in \overline{S}\backslash\overline{S}^+$, $u \neq v$, $x \neq y$, and $uv, xy \notin E(G)$. Then a switching deletes edges ux and vy and replaces them with uv and xy, creating a new graph $G' \in A_{i+1}$. Note that a reverse switching corresponds to choosing two ordered pairs (u, v) and (x, y) such that $uv, xy \in E(G)$, $u, v \in S\backslash S^+$, $x, y \in \overline{S}\backslash\overline{S}^+$, and $ux, vy \notin E(G)$.

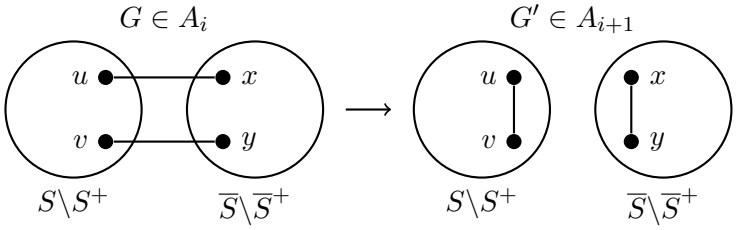

Figure 7.1: A diagram of the main switching used to prove Lemma 7.1.4, which sends $G \in A_i$ to $G' \in A_{i+1}$.

First we give a lower bound on the number of valid forward switchings that can be applied to each $G \in A_i$. By the definition of A, there are at least $\frac{49}{50}aM - 2i$ choices for an edge ux such that $u \in S\backslash S^+$ and $x \in \overline{S}\backslash\overline{S}^+$, and at least $\frac{49}{50}aM - 2i - 1$ choices for a different edge vy such that $v \in S\backslash S^+$ and $y \in \overline{S}\backslash\overline{S}^+$. This choice does not correspond to a valid switching if and only if

(a) $uv \in E(G)$ or $xy \in E(G)$, or

(b) $u = v$ or $x = y$.

First we bound the number of choices described above that satisfy condition (a). Given a choice for the edge ux, there are at most $d(u)$ choices for a vertex $v \in S\backslash S^+$ such that $uv \in E(G)$, and given such a choice of v there are at most $d(v)$ choices for a vertex y such that $vy \in E(G)$. Thus, there are at most $\delta^2 M$ choices for an edge vy such that $uv \in E(G)$, since $\{u, v, x, y\} \subset [n]\backslash H^+$. Analogously, there are at most $\delta^2 M$ choices for an edge vy such that $xy \in E(G)$. Therefore, the number of choices for $\{u, v, x, y\}$ such that $uv \in E(G)$ or $xy \in E(G)$ is at most $2\delta^2 M$. Case (b) is similar:

given a particular choice for the edge ux, there are at most $2\delta\sqrt{M}$ choices for an edge vy such that $u = v$ or $x = y$. Thus, for a given choice of edge ux, there are most $3\delta^2 M$ choices for the edge vy that do not correspond to a valid forward switching. Therefore, the number of valid forward switchings that can be applied to each $G \in A_i$ is at least $\left(\frac{49}{50}aM - 2i\right)\left(\frac{49}{50}aM - 2i - 3\delta^2 M\right)$. For an upper bound on the number of reverse switchings that can be applied to each $G' \in A_{i+1}$, note that there are $2(i+1)$ choices for an ordered pair (u, v) such that $uv \in E(G')$ and $u, v \in S \backslash S^+$, and at most $d(\overline{S})$ choices for the ordered pair (x, y) such that $xy \in E(G')$ and $x, y \in \overline{S} \backslash \overline{S}^+$. Thus, for each $G' \in A_{i+1}$, there are at most $2(i+1)d(\overline{S})$ choices for $\{u, v, x, y\}$ that correspond to a reverse switching. Therefore, it follows that

$$\frac{|A_i|}{|A_{i+1}|} \leq \frac{2(i+1)d(\overline{S})}{\left(\frac{49}{50}aM - 2i\right)\left(\frac{49}{50}aM - 2i - 3\delta^2 M\right)}.$$

Define $\varepsilon = 10^{-3}a^2$. If $i \leq \varepsilon M$, then this upper bound on $|A_i|/|A_{i+1}|$ is at most 10^{-2}. Thus, if $\varepsilon' = 10^{-3}\varepsilon$, then for all $i \leq \varepsilon' M$,

$$\frac{|A_i|}{|A_{\varepsilon M}|} \leq (10^{-2})^{(\varepsilon - \varepsilon')M} = \exp\left(-2(\varepsilon - \varepsilon')\log 10\right) < e^{-\varepsilon M}.$$

Since $|A_{\varepsilon M}| \leq |A|$ and the probability measure on $\mathcal{G}(\boldsymbol{d})$ is the uniform measure, this implies that $\mathbb{P}\left(|E(G[S \backslash S^+])| = i \,|\, G \in A\right) < e^{-\varepsilon M}$ for all $i \leq \varepsilon' M$. Therefore,

$$\mathbb{P}\left(|E(G[S \backslash S^+])| < \varepsilon' M \,\middle|\, G \in A\right) = \sum_{i=0}^{\varepsilon' M - 1} \mathbb{P}\left(|E(G[S \backslash S^+])| = i \,\middle|\, G \in A\right) < \sum_{i=0}^{\varepsilon' M - 1} e^{-\varepsilon M} = o(1). \tag{7.2}$$

Therefore, conditional on $G \in A$, a.a.s. $|E(G[S \backslash S^+])| > \varepsilon' M$. If $G \notin A$, then by definition $G[S]$ contains at least $\frac{1}{100}aM$ edges that are incident to $S \backslash S^+$. Since $E(G[S \backslash S^+]) \subseteq E(G[S])$, this completes the proof. □

Now we use this result to prove Lemma 7.1.4. Despite having little specific knowledge of the degree sequence of $G[S]$ beyond a lower bound on its total degree, the proof is a fairly straightforward application of Theorem 7.2.1 and Lemma 7.2.2. We do this by considering the degree sequence of the induced subgraph $G[S]$ after deleting all the isolated vertices, or equivalently after removing all elements equal to 0 from the degree sequence. We remove the zero elements because Theorem 7.2.1 assumes a minimum degree of at least 1. For the purposes of finding a component with $\Theta(M)$ edges, this is just a technicality, as deleting isolated vertices does not affect the number of edges in any component of $G[S]$.

Recall from Proposition 3.1.1 that, conditional on the event that $G[S]$ has degree sequence $\boldsymbol{k}$, the graph $G[S]$ is a uniformly random graph with degree sequence $\boldsymbol{k}$. We first prove the following claim.

Claim 7.2.3. For each $G \in \mathcal{G}(\boldsymbol{d})$, let $\boldsymbol{t} := \boldsymbol{t}(G)$ be the degree sequence of $G[S]$, ordered in non-decreasing order with isolated vertices excluded. If there exists some constant $\varepsilon > 0$ such that a.a.s. $R(\boldsymbol{t}) \geq \varepsilon M$, then $G[S]$ a.a.s. contains a component with $\Theta(M)$ edges.

Proof of Claim 7.2.3. Let $A(\boldsymbol{k})$ be the subset of $\mathcal{G}(\boldsymbol{d})$ such that $\boldsymbol{t} = \boldsymbol{k}$. Let $\gamma = \gamma(\varepsilon) > 0$ be as defined in Theorem 7.2.1 for this choice of ε. The law of total probability then implies that the

probability that $G[S]$ contains a component with at least $\gamma\widetilde{M}(\boldsymbol{t})$ edges is

$$\sum_{\boldsymbol{k}} \mathbb{P}\left(G \in \Gamma(\boldsymbol{k}) \middle| \boldsymbol{t} = \boldsymbol{k}\right) \mathbb{P}\left(\boldsymbol{t} = \boldsymbol{k}\right),$$

where the sum is over all possible sequences $\boldsymbol{k}$ such that $\boldsymbol{t}(G) = \boldsymbol{k}$. Let $\mathcal{K}$ be the set of sequences $\boldsymbol{k}$ such that there exists $G \in \mathcal{G}(\boldsymbol{d})$ where $\boldsymbol{t}(G) = \boldsymbol{k}$ and $R(\boldsymbol{k}) < \varepsilon M$. By assumption,

$$\sum_{\boldsymbol{k} \in \mathcal{K}} \mathbb{P}\left(\boldsymbol{t} = \boldsymbol{k}\right) = o(1).$$

Let $\Gamma(\boldsymbol{k})$ be the set of $G \in \mathcal{G}(\boldsymbol{d})$ such that $\boldsymbol{t} = \boldsymbol{k}$ and $G[S]$ contains a component with at least $\gamma\widetilde{M}(\boldsymbol{k})$ edges. Proposition 3.1.1 and Theorem 7.2.1 imply that $\mathbb{P}\left(G \in \Gamma(\boldsymbol{k}) \middle| \boldsymbol{t} = \boldsymbol{k}\right) = 1 - o(1)$ for all $k \notin \mathcal{K}$. Therefore,

$$\sum_{\boldsymbol{k}} \mathbb{P}\left(G \in \Gamma(\boldsymbol{k}) \middle| \boldsymbol{t} = \boldsymbol{k}\right) \mathbb{P}\left(\boldsymbol{t} = \boldsymbol{k}\right) = (1 - o(1))(1 - o(1)) + o(1) = 1 - o(1).$$

Finally, note that $\widetilde{M}(\boldsymbol{k}) \geq R(\boldsymbol{k}) \geq \varepsilon M$ for all $k \notin \mathcal{K}$. Therefore, every component with at least $\gamma\widetilde{M}(\boldsymbol{k})$ edges also has at least $\gamma\varepsilon M$ edges. Therefore, a.a.s. $G[S]$ contains a component with at least $\gamma\varepsilon M$ edges, which proves the claim. □

Now we apply this claim to prove Lemma 7.1.4. We also recall Proposition 2.2.16, which states that if $\boldsymbol{d}$ is an n-element graphical sequence such that $M(\boldsymbol{d}) = \omega(n)$, then $R(\boldsymbol{d}) = M(\boldsymbol{d})(1 - o(1))$.

Proof of Lemma 7.1.4. For each $G \in \mathcal{G}(\boldsymbol{d})$, define $\boldsymbol{t} := \boldsymbol{t}(G)$ to be the degree sequence of $G[S]$, ordered in non-decreasing order with isolated vertices excluded. Lemma 7.2.2 implies that, under the conditions on Lemma 7.1.4, a.a.s. $|E(G[S])| \geq 10^{-6}a^2M$, where $a = d(S \backslash S^+)/M = \Theta(1)$. Since $M = \omega(n)$, this implies that a.a.s. $M(\boldsymbol{t}) = \omega(n)$. For all such $\boldsymbol{t}$, Proposition 2.2.16 implies that $R(\boldsymbol{t}) \geq M(\boldsymbol{t})(1 - o(1))$. Thus, there exists a constant $\varepsilon > 0$ (for example, 10^7a^2) such that a.a.s. $R(\boldsymbol{t}) \geq \varepsilon M$. Therefore, Claim 7.2.3 implies that $G[S]$ a.a.s. contains a component with $\Theta(M)$ edges. □

Note that we did not use the assumption that $M = \omega(|S|)$ in the proof of Lemma 7.2.2. This non-vanishing fraction of edges in $G[S]$ with at least one endpoint in $S \backslash S^+$ is a consequence of our assumption that $d(S \backslash S^+) = \Theta(M)$. However, such a result does not immediately tell us whether the induced graph a.a.s. contains a giant component or not, as the non-vanishing fraction could still be incredibly small with respect to M. In order to make such claims, we would need to know more about the distribution of the induced degree sequence, or at the very least have a much stronger lower bound on the number of edges in $E(G[S \backslash S^+])$.

Remark 7.2.4. One might wish to try a similar idea to the proof of Lemma 7.1.4 to show that $G[S]$ a.a.s. contains a component with $\Theta(|S|)$ vertices. Theorem 2.2.15 also applies under the same conditions as Theorem 7.2.1, and so a similar argument can be applied. This looks promising, but there is an issue: these theorems only apply to degree sequences with minimum degree at least 1, and the graph $G[S]$ may contain many isolated vertices with non-trivial probability.

Let $N_0(G)$ be the set of vertices in S that have no neighbour in S, that is, the set of vertices with degree 0 in $G[S]$. Suppose $M = \omega(|S|)$, and $G \in A(\boldsymbol{k})$ as defined in the previous proof. A

proof based on Lemma 10 in the paper of Joos et al. [83] (replacing $\log\log M$ with an arbitrary function $f(M)$ such that $f(M) = \omega(1)$) implies that a uniformly random $G \in A(\boldsymbol{k})$ a.a.s. has a component of order $(|S| - |N_0(G)|)(1 - o(1))$, as well as a component containing $\Theta(M)$ edges. Thus, if there exists some constant $c > 0$ such that a.a.s. $|N_0| \leq (1-c)|S \backslash S^+|$ for a uniformly random $G \in \mathcal{G}(\boldsymbol{d})$, this immediately implies the existence of a giant component in $G[S]$ with probability $1 - o(1)$. However, as we discuss later (specifically in Section 7.7), $|N_0(G)|$ is a difficult value to control. There are many pairs $(\boldsymbol{d}, S)$ which satisfy the conditions of Theorem 7.1.2 for which $|S| - |N_0(G)| = o(|S|)$ with probability bounded away from 0.

7.3 The random graph exploration process

Now we analyse the case where $M = \Theta(n)$. We analyse this in the same way that we analysed the configuration model, using a similar breadth-first exploration process and the method of deferred decisions. Due to the similarities in the description of the exploration process, we carry over the language of the configuration model and continue to work on a restricted subset of $\mathcal{C}(\boldsymbol{d})$ designed to mirror the probability space $\mathcal{G}(\boldsymbol{d})$. Previously, a pairing P was sampled uniformly at random. In this model, a pairing is sampled uniformly at random conditional on the event that the corresponding graph G is simple. Let $\Phi \subset \mathcal{C}(\boldsymbol{d})$ be the subset of pairings that have no multiple edges or loops. We call these pairings *simple*, as they are the pairings that correspond to simple graphs. For each graph $G \in \mathcal{G}(\boldsymbol{d})$, there are exactly $\prod_{i=1}^{n} d(i)!$ pairings in Φ that correspond to G. Thus, the distributions of the size and order of the largest component in $\mathcal{G}(\boldsymbol{d})$ are the same as their distributions in Φ.

Let $G^* \in \Phi$ be a simple pairing. We denote the elements of Φ by G^* to emphasise the both its similarity to and distinction from a simple graph. Recall that the half edges in G^* are labelled with $[M] = \{1, \ldots, M\}$. For a given $G^* \in \Phi$ and $i \in [M]$, define $m(i)$ to be the mate of half edge i in the pairing G^*. Let the parent vertex of i be denoted u_i. Let $V(G^*)$ be the set of 'bins' or vertices of G^*, and let $B(u)$ be the set of half edges that are in the bin corresponding to vertex i. Let $E(G^*)$ be the set of edges of G^*, which are unordered pairs of half edges in $[M]$. For a given $G^* \in \Phi$, these can also be uniquely represented by the unordered pair of their parent vertices $\{u_i, u_{m(i)}\}$ (also written $u_i u_{m(i)}$ for brevity).

We define an exploration process that is, in effect, the same as the configuration model version. At each step we have a partial pairing $T_{t-1} = (V_{t-1}, E_{t-1}, \mathcal{X}_{t-1})$, and we explore a specific open edge from this partial pairing to another vertex w_t. The difference here is that if $w_t \notin S$, we do not expose any other edges between w_t and V_{t-1}. This makes the algorithm behave much more like exploring a graph where the vertices in $\overline{S}$ have been "exploded", as they were in the proof of the configuration model result. We avoid the notion of exploding vertices in this chapter for intuition reasons: as the pairings we consider here are simple, exploding the vertices in $\overline{S}$ gives the wrong intuition for vertex adjacency - a vertex in S could not be adjacent to two different degree 1 vertices that correspond to the same exploded vertex in $\overline{S}$. Thus, we alter the definition of the exploration process to mimic the configuration model process without having to handle any explosives.

However, this change to how vertices in $\overline{S}$ are dealt with makes it harder to define some of the variables that we wish to consider. For example, in the exploration process on the configuration model, M_{t-1} was defined as the sum of the degrees of unexplored vertices in the pairing. If we do

not explode vertices in $\overline{S}$, there may be vertices $u \in \overline{S}$ that have been "explored" at some time t in the sense that $w_t = u$, but it is still possible that $w_{t'} = u$ at some future time step $t' > t$. This is because the other edges between u and V_{t-1} (which we called back edges) were not revealed when u was explored at time t, as $u \in \overline{S}$.

More generally, we want notation to discuss the "available degree" of a given vertex. This is designed to capture the notion of how many half edges can be matched with the open edge being paired at some time step, or equivalently how much that vertex contributes to M_t. In line with the configuration model intuition, this corresponds to the number of half edges in the corresponding bin that are not paired or open in T_t. To this end, we define the available degree of each vertex $w \in V$ at time t as follows:

$$d^{(t)}(w) = \begin{cases} 0, & w \in S \cap V_t \\ d(w) - X & \text{otherwise} \end{cases}, \tag{7.3}$$

where $X = |\{v \in V \mid vw \in E_t\}|$, the number of edges incident to w that are in the partial pairing T_t. This definition is very useful when bounding the probability of discovering vertices at each step, and thus for analysing the expected number of open edges gained at each step t. For $w \in S$, this definition is simple: $d^{(t)}(w) = d(w)$ if w has not been explored yet (that is, if $w \notin V_{t-1}$), and $d^{(t)}(w) = 0$ if it has (that is, if $w \in V_{t-1}$). For some $w \notin S$, some edges between T_{t-1} and w may have already been revealed at previous steps. These half edges in the bin corresponding to w cannot be paired with the open edge being explored at time t, as they are already paired in T_{t-1}. This naturally affects the probability that $w = w_t$. This is formalised in Section 7.4.

Now we describe the exploration process explicitly. We consider the (graphical) sequence $\boldsymbol{d}$ and the set $S \subset [n]$ as fixed. The input for the process is a simple pairing $G^* \in \Phi$. At time $t = 0$ of the exploration process, initialise $T_0 = (V_0, E_0, \mathfrak{X}_0)$. The set V_0, called the *preprocessing set*, is a subset of V defined depending on the size of $R_S(\boldsymbol{d})$ (that is, whether $R_S(\boldsymbol{d}) = o(M)$ or $R_S(\boldsymbol{d}) = \Theta(M)$) and is defined specifically in the relevant later sections. The corresponding edge set E_0 also depends on the size of $R_S(\boldsymbol{d})$ and is defined later. As with the exploration process on $\mathcal{C}(\boldsymbol{d})$, each step of the exploration process produces a partial pairing T_t such that

$$T_0 \subset T_1 \subset T_2 \subset \cdots \subset (V(G^*), E(G^*), \emptyset),$$

where $T_t = (V_t, E_t, \mathfrak{X}_t)$ and the subset notation means that $V_{t-1} \subseteq V_t$ and $E_{t-1} \subset E_t$. The sets $V_t \subset V(G^*)$ and $E_t \subset E(G^*)$ correspond to the vertices and edges of the partial pairing, and are constructed iteratively in a way defined below. The set $\mathfrak{X}_t \subset [M]$ is a set of half edges i such that the parent vertex u_i is in $V_t \cap S$ and i is unpaired in T_t. These half edges are called *open* edges. As in the configuration model, we call T_t the *partially explored sub-pairing at time t*, or *partial pairing at time t* for brevity.

We now describe the iterative steps of the exploration process for all $t \geq 1$. If $X_{t-1} = 0$, then the partial pairing T_{t-1} contains no open edges. Then let i be a uniformly random half edge in G^* such that the parent vertex u_i is in $S \backslash V_t$. Define $w_t := u_i$. Then $V_t = V_{t-1} \cup \{w_t\}$, $E_t = E_{t-1}$, and $\mathfrak{X}_t = \bigcup_{j \in w_t} \{j\}$.

On the other hand, if $X_{t-1} > 0$, then there exists some open edge in T_{t-1}, that is, $\mathfrak{X}_{t-1} \neq \emptyset$. Let i be the lowest-indexed half edge in $\mathfrak{X}_{t-1}$. Denote its parent vertex v_t, and define w_t to be the

parent vertex of its mate $m(i)$. Then $V_t = V_{t-1} \cup \{w_t\}$. If $w_t \notin S$, then $E_t = E_{t-1} \cup \{v_t w_t\}$ and $\mathcal{X}_t = \mathcal{X}_{t-1} \backslash \{i\}$. If $w_t \in S$, then $E_t = E_{t-1} \cup \bigcup_{i \in \mathcal{B}_t} \{u_i, u_{m(i)}\}$ and $\mathcal{X}_t = \mathcal{X}_{t-1} \cup B(w_t) \backslash \mathcal{B}_t$, where

$$\mathcal{B}_t := \{i \mid i \in B(w_t),\ m(i) \in \mathcal{X}_t\}$$

and $B(w_t)$ is the set of half edges associated with vertex w_t. The edges added to E_t at this step, other than $v_t w_t$, are called *back edges*. The number of back edges added to E_t at time t is denoted $d'_{S,t}(w_t)$, which is equal to $|\mathcal{B}_t| - 1$, since we do not consider the edge $v_t w_t$ a back edge. Note that if $w_t \notin S$, then it is possible that $w_t = w_s$ for some $s < t$. On the other hand, if $w_t \in S$, then $w_t \neq w_s$ for all $s \neq t$.

Much like when studying $\mathcal{C}(\boldsymbol{d})$, we use the method of deferred decisions to analyse this process on a random element of Φ. This way the pairing is generated as the algorithm progresses - instead of exploring a predetermined uniformly random simple pairing, the deferred decisions process can be thought of as constructing a uniformly random simple pairing in a stepwise manner. However, since G^* is not a uniformly random element of $\mathcal{C}(\boldsymbol{d})$, the probability that $w_t = w$ for any given vertex w is non-trivial to determine. For a given partial pairing T, we can consider the set of all simple pairings $G^* \in \Phi$ such that G^* *agrees with* T *at time* $t-1$, that is, such that G^* contains all pairs in T and does not contain any edges between vertices in $V_{t-1} \cap S$ that are not present in T. This corresponds exactly to the set of G^* such that there exists a sequences of choices for half edges at each step $t' < t$ such that $T_{t-1}(G^*) = T$, that is, the set of simple pairings $G^* \in \Phi$ that could possibly have partial pairing equal to T at time $t-1$. We denote this set $\Phi(T)$, though it is implicitly dependent on t as well.

We say that vertices in V_t and edges in E_t have been *discovered* or *revealed* at time t, and w_t is the vertex *explored* at time t. If $w_t \in S$, back edges between w_t and V_{t-1} are said to have been *revealed* at time t. We also say that the (connected) component of T_t containing the parent vertices of the open edges in $\mathcal{X}_t$ is called the *active component*. Again we define $\mathbb{P}_t(\cdot) := \mathbb{P}(\cdot \,|\, T_t)$, where we condition on the deferred decision process up to time t; we define $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot \,|\, T_t]$ to be the corresponding expectation. This conditional expectation is over all $G^* \in \Phi$ such that the partial pairing at time t is equal to T_t, which is precisely the set $\Phi(T_t)$.

Recall the definition of $d^{(t)}(w)$, the available degree of $w \in V$ at time t, given in (7.3). We define M_t by

$$M_t := \sum_{w \in V} d^{(t)}(w).$$

This is analogous to the definition of M_t for the process on $\mathcal{C}(\boldsymbol{d})$, representing the number of "available" half edges for a given open edge to be paired with at time t of the deferred decision process. We also define $S_t = S \backslash V_t$. This represents the set of vertices in S that have not yet been "discovered" by the exploration process at time t. These vertices are available to be explored at time t, and thus contribute to the expected number of open edges gained or lost at each step of the process. Thus, $d(S_t)$ is the total degree of the vertices in $S \backslash V_t$. Equivalently, this can be defined as

$$d(S_t) = \sum_{w \in S} d^{(t)}(w).$$

Defining the analogous set for $\overline{S}$ is more complicated, as half edges i where $u_i \in \overline{S} \cap V_t$ can still be paired with other half edges in $\mathcal{X}_t$. Define $\overline{S}_t$ to be the set of vertices $v \in \overline{S}$ with at most $d(v) - 1$ neighbours in T_t. Then we define

$$d(\overline{S}_t) := \sum_{w \in \overline{S}} d^{(t)}(w).$$

It follows from these definitions that $M_t = d(S_t) + d(\overline{S}_t)$. Conditional on an arbitrary sequence of partial pairings $(T_s)_{s \leq t}$, if $d^{(t)}(v) = 0$ then $\mathbb{P}_t(w_{t+1} = v) = 0$. This is because if $v \in S$, then E_t contains all edges between v and $V_t \cap S$, and thus no open edge in $\mathcal{X}_t$ can be paired with any unpaired half edge in v. On the other hand, if $v \notin S$, then $d^{(t)}(v) = 0$ means that E_t contains all edges incident to v in the simple pairing G^*.

At each step of the exploration process, all of these aforementioned quantities change based on w_t. Recall the definition of $\mathtt{d}_S(\cdot)$ from (6.1), that $\mathtt{d}_S(w) = d(w)$ if $w \in S$ and $\mathtt{d}_S(w) = 1$ if $w \notin S$. At each step $t \geq 1$, $M_t = M_{t-1} - \mathtt{d}_S(w_t)$. If $w_t \in S$, then $d(S_t) = d(S_{t-1}) - d(w_t)$, and $d(\overline{S}_t) = d(\overline{S}_{t-1})$. If $w_t \notin S$, then $d(S_t) = d(S_{t-1})$, and $d(\overline{S}_t) = d(\overline{S}_{t-1}) - 1$. If $w_t \in S$ and $X_{t-1} = 0$, then $X_t = d(w_t)$. If $w_t \in S$ in $X_{t-1} > 0$, then $X_t = X_{t-1} + d(w_t) - 2 - 2d'_{S,t}(w_t)$, where we recall that $d'_{S,t}(w_t)$ is the number of back edges between w_t and $S \cap V_{t-1}$. If $w_t \notin S$, then $X_t = X_{t-1} - 1$. Equivalently, these two cases where $X_{t-1} > 0$ can be expressed by the single equation

$$X_t = X_{t-1} + \mathtt{d}_S(w_t) - 2 - 2d'_{S,t}(w_t).$$

These relationships form the crux of our analysis.

Remark 7.3.1. In order to simplify some of the computations in the analysis of the exploration process, we amalgamate various different $o(1)$ functions such that δ can take the place of all of them. Specifically, in the case that $M = \Theta(n)$ we assume that δ satisfies the following conditions:

(a) $\delta \to 0$ as $n \to \infty$,

(b) $d(\overline{S}^+) \leq \delta^2 n$,

(c) $\delta = \omega(\log^{-1} M)$.

In Section 7.5 where we prove Theorem 7.1.1, the subcritical result, we also impose that $R_S \leq \delta M$. Since Theorem 7.1.1 assumes that $R_S = o(M)$, and increasing δ monotonically excludes more elements from H^+, these conditions can all be satisfied simultaneously by taking the largest $\delta \to 0$ that satisfies the latter three conditions. Importantly, since we assume that $M = \Theta(n)$ and we assume that $d(\overline{S}^+) = o(M)$, we can choose δ such that $d(\overline{S}^+) \leq \delta^2 n$.

7.4 The probability of discovering a given vertex at each step

In the configuration model, Observation 6.2.1 implied that, conditional on the process up to time $t - 1$, the probability of discovering a particular vertex w at time t was exactly $d(w)/M_{t-1}$. This incredibly important piece of the analysis was an immediate consequence of the definition of the configuration model. In this section we determine analogous upper and lower bounds on

the probability of exploring a given vertex $w \in V$ at time t in the deferred decision exploration process. We prove these results using switchings. To do this, we assume that the partial pairing at time $t-1$ satisfies certain conditions that allow the switching proofs to work. Recall the definition of v_t, the parent vertex of the half edge being paired by the exploration process at time t. We call a partial pairing $T = (V_{t-1}, E_{t-1}, \mathcal{X}_{t-1})$ *predictable* if it satisfies the following conditions.

(P1) $M_{t-1} = \Theta(M)$,

(P2) $d(v_t) \leq \delta\sqrt{M}$,

(P3) S_{t-1} has maximum degree at most $\delta\sqrt{M}$,

(P4) $X_{t-1} \leq \alpha n$ for some $\alpha < 10^{-3}$,

(P5) $|V_{t-1}| \leq \frac{1}{3}n$.

Implicitly, a partial pairing at time $t-1$ is predictable with respect to the degree sequence of the underlying simple pairing G^*. If a partial pairing is predictable with respect to G^*, it is also predictable with respect to all other simple pairings $G^* \in \Phi$ that agree with T, since all pairings in Φ have the same degree sequence. Thus, we just say that the partial pairing itself is predictable.

We call such a partial pairing predictable because we show that under these conditions, the probability of discovering a given vertex at each step is (up to some small caveats) approximately equal to its value in the configuration model. These conditions are sufficient but not necessary — for example, we do not optimise the constants in Items (P4) and (P5) at all. Since we assume that $M = \Theta(n)$ throughout this part of the proof, the penultimate condition is equivalent to the condition that $X_{t-1} \leq \alpha' M$ for $\alpha' = \alpha n/M = \Theta(1)$. We show in the respective sections that, as a consequence of the preprocessing step, T_{t-1} satisfies these conditions for some large number of time steps $t \geq 1$ in both the subcritical and supercritical regimes.

There are two main steps to obtaining the vertex discovery probabilities for predictable partial pairings. Firstly, we show that the probability that $w_t \in \overline{S}^+$ for each t is very small. Specifically, we show that this probability is bounded above by δ, where $\delta := \delta(n) \to 0$ is a function such that $d(\overline{S}^+) \leq \delta^2 n$. This means that the presence of a small number of high-degree vertices in $\overline{S}$ does not meaningfully affect the probability of exploring a given vertex at time t. Then we show that, conditional on the event that $w_t \notin \overline{S}^+$, the probability that $w_t = w$ is close to what it is in the configuration model. The reason for dealing with these two cases separately, rather than just having a single lemma for all vertices, is that the switchings used to prove the bounds used for the lower-degree vertices rely on being able to assume that particular vertices have maximum degree $\delta\sqrt{M}$. However, the probability of discovering a vertex in $\overline{S}^+$ is still unlikely at each step, since we imposed that $d(\overline{S}^+) \leq \delta^2 n$. Thus, we deal with this as a separate case, and use the law of total probability to recombine these cases in the later proofs where these bounds are used.

Recall from (7.3) the definition of $d^{(t)}(w)$ for a vertex w, the number of unpaired half edges incident to w that have not been revealed at time t. We call this the *available* degree of w at time t. The appearance of $d^{(t-1)}(w)$ in the following proofs is quite natural: for a vertex $w \notin S$, there may exist edges $wv \in E_{t-1}$, as it may have been discovered by the process previously. The function $d^{(t-1)}(w)$ essentially counts the number of "available" half edges in $B(w)$, the set of half edges corresponding to w, that can be explored at time t. This is why many bounds in the following proofs are given in terms of $d^{(t-1)}(w)$, as we condition on the partial pairing at time $t-1$

and thus exclude switchings that use edges that are in E_{t-1}. Recall that $M_{t-1} = \sum_{w \in V} d^{(t-1)}(w)$, and that $\mathbb{P}_{t-1}(\cdot)$ is the probability of an event conditional on the partial pairing at time $t-1$.

Lemma 7.4.1. Suppose that $M = \Theta(n)$ and $d(\overline{S}^+) \leq \delta^2 n$ for some $\delta \to 0$. Suppose that $T_{t-1} = (V_{t-1}, E_{t-1}, \mathcal{X}_{t-1})$, the partial pairing at time $t-1$, is predictable. Then $\mathbb{P}_{t-1}\left(w_t \in \overline{S}^+\right) \leq \delta$.

Proof. Let $T = (V_{t-1}, E_{t-1}, \mathcal{X}_{t-1})$ be a predictable partial pairing, and let i_t be the lowest-indexed open edge in $\mathcal{X}_{t-1}$. Let $\Phi(T)$ be the set of $G^* \in \Phi$ that agree with T, that is, the set of simple pairings such that there exists a sequence of choices for half edges at each step $t' < t$ such that $T_{t-1} = T$. Let A be the set of simple pairings $G^* \in \Phi(T)$ such that $w_t \in \overline{S}^+$, and let B be its complement (the set of simple pairings such that $w_t \notin \overline{S}^+$). We define a switching mapping simple pairings from A to B. Suppose $G^* \in A$, and define w to be the vertex that contains the half edge paired with i_t, that is, $w_t = w$. Then let (x, y) be an ordered pair of vertices such that $x \notin \overline{S}^+$ and $xy \in E(G^*)$. A switching takes edges $v_t w$ and xy, deletes them, and replaces them with edges $v_t x$ and wy, creating a new pairing G'^* where $w_t = x$. This switching is valid as long as G'^* is a simple pairing in B.

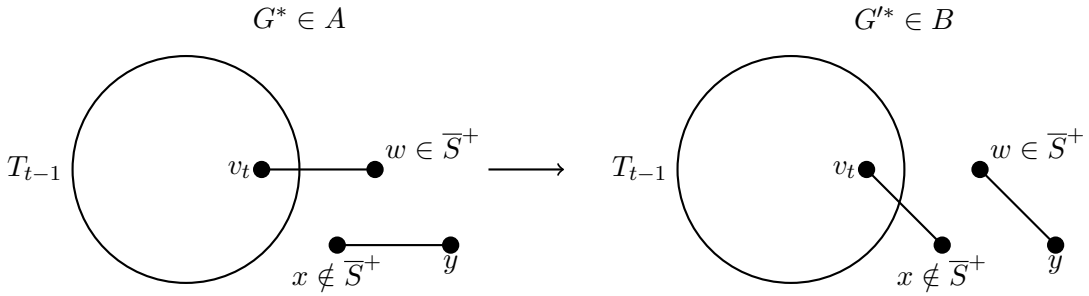

Figure 7.2: A diagram of the switching used in the proof of Lemma 7.4.1, which sends $G^* \in A$ to $G'^* \in B$.

First we determine a lower bound on the number of valid forward switchings that can be applied to each $G^* \in A$. The number of choices for a vertex y such that $d^{(t-1)}(y) \geq 1$ is at least $\frac{2}{3}n$, since the number of vertices y such that $d^{(t-1)}(y) = 0$ is at most $|V_{t-1}|$, which is at most $\frac{1}{3}n$ since T is predictable and thus satisfies (P5). Given y, there is at least one choice for x such that $xy \in E(G^*)$ and $xy \notin E_{t-1}$. Such a choice of $\{v_t, x, w, y\}$ corresponds to a valid switching if and only if

(a) $x \notin \overline{S}^+$ and $x \notin V_{t-1} \cap S$,

(b) $wy \notin E(G^*)$,

(c) $v_t x \notin E(G^*)$,

(d) $\{v_t, x, w, y\}$ are all distinct.

We bound from below the number of choices for $\{v_t, x, w, y\}$ such that

(a′) $y \notin \overline{S}^+$ and y has no neighbour in $\overline{S}^+$,

(b′) $x \notin V_{t-1} \cap S$,

(c′) $v_t x \notin E(G^*)$,

(d′) $\{v_t, x, w, y\}$ are all distinct.

Every choice that satisfies (a′) – (d′) also satisfies (a) – (d). Let $Z(G^*)$ be the number of choices for (x, y) such that one of (a′) – (d′) are not satisfied. Then the number of switchings that can be applied to each $G^* \in A$ is at least $\frac{2}{3}n - Z(G^*)$. We bound $Z(G^*)$ from above for all $G^* \in A$. For case (a′), there are at most $d(\overline{S}^+) + d(\overline{S}^+)/(\delta\sqrt{M})$ choices for y such that $y \in \overline{S}^+$ or y has a neighbour in $\overline{S}^+$. For case (b′), there are X_{t-1} choices for (x, y) such that $d^{(t-1)}(y) \geq 1$ and $x \in V_{t-1} \cap S$. For case (c′), there are at most $d(v_t)$ choices for a vertex x such that $x \notin \overline{S}^+$ and $v_t x \in E(G^*)$, and given x there are at most $d(x)$ choices for y. Since we assume that T is predictable (and thus satisfies (P2) and (P3)) and $x \notin \overline{S}^+$, it follows that there are at most $\delta^2 M$ choices for $\{v_t, x, w, y\}$ that do not satisfy (c′). Finally, given the way the vertices are chosen, the only way that these vertices can be non-distinct is if $v_t = x$. However, this is counted in part (b′), since this implies that $x \in V_{t-1} \cap S$. Therefore,

$$Z(G^*) \leq d(\overline{S}^+) + d(\overline{S})/(\delta\sqrt{M}) + X_{t-1} + \delta^2 M.$$

Since we assume that $M = \Theta(n)$ and $d(\overline{S}^+) \leq \delta^2 n$, this implies that the number of valid switchings from A to B is at least $\frac{2}{3}n - o(n) - X_{t-1}$, which is at least $\frac{1}{2}n$ since $X_{t-1} < 10^{-3}n$ by condition (P4). The number of backward switchings is at most $d(\overline{S}^+)$, the number of choices for a vertex y such that $wy \in E(G^*)$ for some $w \in \overline{S}^+$, since the vertices v_t and x are fixed by the partial pairing T. This implies that

$$\frac{|A|}{|B|} \leq \frac{d(\overline{S}^+)}{\frac{1}{2}n} \leq 2\delta^2.$$

Thus, it follows that

$$\mathbb{P}_{t-1}\left(w_t \in \overline{S}^+\right) = \frac{|A|}{|A| + |B|} = \frac{|A|/|B|}{|A|/|B| + 1} \leq 2\delta^2.$$

This completes the proof. □

Now we show that, conditional on the event that $w_t \notin \overline{S}^+$, the probability that $w_t = w$ for some $w \notin \overline{S}^+$ is approximately proportional to its available degree, $d^{(t-1)}(w)$. This aligns with the intuition from the configuration model, and allows us to employ similar methods to show that $\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t) - 2\right]$ evolves similarly to the analogous quantity (which was simply $\mathbb{E}_{t-1}\left[d(w_t) - 2\right]$) in the configuration model. By conditioning on the case where $w_t \notin \overline{S}^+$, we can study this case using a similar switching analysis to that done by Joos et al. [83].

In the following lemma, we also have to exclude vertices $w \in \overline{S}$ that are already adjacent to v_t in the partial pairing T_{t-1}. Since the partial pairing is simple, if $v_t w \in E_{t-1}$, then $\mathbb{P}_{t-1}(w_t = w) = 0$. However, if we consider the case where T is predictable, then $d(v_t) \leq \delta\sqrt{M}$. This allows us to show in the later sections that this caveat does not significantly affect how the process evolves.

Recall that $S_{t-1} := S \backslash V_{t-1}$, and $\overline{S}_{t-1}$ is the set of all vertices $w \in \overline{S}$ such that $d^{(t-1)}(w) \geq 1$.

Lemma 7.4.2. Suppose that $M = \Theta(n)$ and $d(\overline{S}^+) \leq \delta^2 n$ for some $\delta \to 0$. Suppose that $T_{t-1} = (V_{t-1}, E_{t-1}, \mathcal{X}_{t-1})$, the partial pairing at time $t-1$, is predictable. Let $w \notin \overline{S}^+$ be a vertex

such that $d^{(t-1)}(w) \geq 1$ and $v_t w \notin E_{t-1}$. Then

$$\mathbb{P}_{t-1}\left(w_t = w \middle| w_t \notin \overline{S}^+\right) = \frac{d^{(t-1)}(w)}{M_{t-1}}\left(1 \pm O(\delta)\right).$$

Proof. Let $T = (V_{t-1}, E_{t-1}, \mathfrak{X}_{t-1})$ be a predictable partial pairing. Let $\Phi(T)$ be the subset of simple pairings $G^* \in \Phi$ that agree with T. Let A_w be the set of $G^* \in \Phi(T)$ such that $w_t = w$, and let B_w be the set of $G^* \in \Phi(T)$ such that $w_t \neq w$ and $w_t \notin \overline{S}^+$. We define a switching between A_w and B_w. Let $G^* \in A_w$ be a simple pairing. To perform a switching, choose an ordered pair of vertices (x, y) such that $xy \in E(G^*)$. Then the switching deletes edges $\{v_t w, xy\}$ and adds in edges $\{v_t x, wy\}$ to create a new G'^*. This switching is considered valid if and only if $G'^* \in B_w$, which occurs if and only if

(a) $x \in S_{t-1} \cup \overline{S}_{t-1}$ and $x \notin \overline{S}^+$,

(b) $v_t x, wy \notin E(G^*)$,

(c) the vertices $\{v_t, w, x, y\}$ are all distinct,

(d) $xy \notin E_{t-1}$.

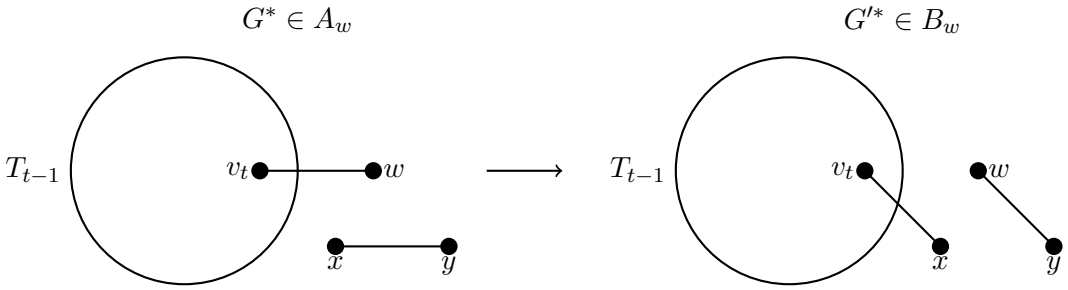

Figure 7.3: A diagram of the main switching used to prove Lemma 7.4.2, which sends $G^* \in A_w$ to $G'^* \in B_w$.

The switching is invalid if $x \in \overline{S}^+$ because $w_t = x$ in G'^* (since every simple pairing in $\Phi(T)$ is pairing the same half edge at time t), and we are conditioning on the event that $w_t \notin \overline{S}^+$.

Now we obtain upper and lower bounds on the number of valid switchings that can be applied to each $G^* \in A_w$. By definition, there are M_{t-1} choices for an ordered pair of adjacent vertices (x, y) such that $d^{(t-1)}(x) \geq 1$ (which is equivalent to the condition that $x \in S_{t-1} \cup \overline{S}_{t-1}$) and $xy \notin E_{t-1}$. This means that every such choice satisfies (d). Let $W(G^*)$ be the number of these M_{t-1} choices for (x, y) such that at least one of (a), (b), or (c) is not satisfied. Then the number of valid switchings that can be applied to G^* is at least $M_{t-1} - W(G^*)$. Let $Z(G^*)$ be the number of these M_{t-1} choices such that at least one of the following statements is true:

(a′) $x \in \overline{S}^+$ or $y \in \overline{S}^+$,

(b′) $x, y \notin \overline{S}^+$ and $v_t x$ or $wy \in E(G^*)$,

(c′) $x, y \notin \overline{S}^+$ and the vertices $\{v_t, w, x, y\}$ are not all distinct.

It follows that $M_{t-1} - Z(G^*)$ is a lower bound on $M_{t-1} - W(G^*)$ for all G^*. We give an upper bound on Z for all $G^* \in A_w$.

For case (a′), there are at most $d(\overline{S}^+)$ choices for (x, y) such that $x \in \overline{S}^+$ and $xy \in E(G^*)$, and similarly at most $d(\overline{S}^+)$ choices for the case where $y \in \overline{S}^+$. Since $d(\overline{S}^+) \leq \delta^2 n$, there are at most $2\delta^2 n$ choices for (x, y) such that (a′) is true.

For case (b′), we first give an upper bound on the number of choices for (x, y) such that $v_t x$ is an edge. There are at most $d(v_t)$ choices for x such that $v_t x \in E(G^*)$ and $x \notin \overline{S}^+$. Then, given x, there are at most $d(x)$ choices for y such that $xy \in E(G^*)$. If T is predictable, then T satisfies (P2), and thus $d(v_t) \leq \delta\sqrt{M}$. Since T also satisfies (P3) and $x \notin \overline{S}^+$, it also follows that $d(x) \leq \delta\sqrt{M}$. By the same reasoning, the number of choices for (x, y) such that $x, y \notin \overline{S}^+$ and $wy \in E(G^*)$ is at most $d(w)\delta\sqrt{M}$, which is at most $\delta^2 M$ as $w \notin \overline{S}^+$. Thus, the number of choices for (x, y) such that (b′) is true is at most $2\delta^2 M$.

For case (c′), note that $v_t \neq x$, since $v_t \in V_{t-1} \cap S$ and thus $d^{(t-1)}(v_t) = 0$. So the possible cases for non-distinct vertices are $v_t = y$, $w = y$, and $w = x$. At most $d(v_t)$ choices for (x, y) correspond to the case where $v_t = y$, and at most $2d(w)$ correspond to the cases where either $w = y$ or $w = x$. By the same logic to case (b′), this implies that the number of choices for (x, y) that satisfy (c′) is at most $3\delta\sqrt{M}$.

Altogether, this gives the bound

$$Z \leq 2\delta^2 n + 2\delta^2 M + 3\delta\sqrt{M},$$

which is asymptotically at most δM since $M \geq n$ and $\delta = o(1)$. Thus, the number of switchings that can be applied to each $G^* \in A_w$ is at least $M_{t-1} - \delta M$ and at most M_{t-1}. Recall that since T is predictable, condition (P1) implies that $M_{t-1} = \Theta(M)$.

Now we consider the analogous upper and lower bounds on the number of $G^* \in A_w$ that can be mapped to a given $G'^* \in B_w$ via a switching, or equivalently, the number of reverse switchings that map an element $G'^* \in B_w$ to an element $G^* \in A_w$. A reverse switching applied to some $G'^* \in B_w$ is equivalent to the following: choose an edge $wy \notin E_{t-1}$ (noting that $w_t \neq w$), then delete edges $\{v_t w_t, wy\}$ and add in edges $\{v_t w, w_t y\}$, hence creating some $G^* \in A_w$. Given some $G'^* \in B_w$, there are at most $d^{(t-1)}(w)$ choices for a neighbour y such that $wy \notin E_{t-1}$. Such a choice corresponds to a valid reverse switching if and only if the resulting G^* is an element of A_w, which occurs if and only if

(i) neither $v_t w \in E(G'^*)$ nor $w_t y \in E(G'^*)$,

(ii) the vertices $\{v_t, w, w_t, y\}$ are all distinct.

Note that since $G'^* \in B_w$ and $B_w \subseteq \Phi(T)$, it follows that $w_t \notin \overline{S}^+$ for all such simple pairings. We use extra, separate switching arguments to show the following claims. We defer their proofs for the moment.

Claim 7.4.3. Conditional on the event that $G'^* \in B_w$, the probability that $v_t w \notin E(G'^*)$ is $1 - O(\delta^2)$.

Claim 7.4.4. Conditional on the event that $G'^* \in B_w$, the probability that $w_t w \notin E(G'^*)$ is $1 - O(\delta^2)$.

Claim 7.4.5. Conditional on the event that $G'^* \in B_w$, the probability that there are at most $\delta d^{(t-1)}(w)$ edges not in E_{t-1} between w and the neighbours of w_t is $1 - O(\delta)$.

We call a vertex $v \in N_G(w)$ (the neighbourhood of w in G) an "undiscovered neighbour of w" if $vw \notin E_{t-1}$. Assuming these claims, we now complete the proof of the lemma. These claims imply that the proportion of elements $G'^* \in B_w$ such that neither $v_t w$ nor $w_t w$ are in $E(G'^*)$ and at most $\delta d^{(t-1)}(w)$ undiscovered neighbours of w are also neighbours of w_t is at least $1 - O(\delta)$. For each such element of B_w, we can obtain a strong lower bound on the number of valid choices for wy and thus the number of possible reverse switchings. By design, there are at most $\delta d^{(t-1)}(w)$ choices that do not satisfy (i). For case (ii), we consider the different possible ways that $\{v_t, w, w_t, y\}$ can be non-distinct. The vertices v_t and y must be distinct, as $v_t w \notin E(G'^*)$ for all such G'^*. By assumption, v_t is distinct from w_t and w, and w_t and w are distinct from each other. Finally, w and y are distinct, since they are adjacent. The only remaining case is if $w_t = y$. However, since $wy \in E(G'^*)$, the condition $w_t = y$ implies that $w_t w \in E(G'^*)$, which is a contradiction of our choice of G'^*. Thus, the fraction of elements of B_w for which there are at most $\delta d^{(t-1)}(w)$ choices for y that do not satisfy (i) and (ii) is $1-O(\delta)$. Therefore, the average number of reverse switchings that can be applied to each $G'^* \in B_w$ is at least $d^{(t-1)}(w)(1-\delta)(1-O(\delta)) \geq d^{(t-1)}(w)(1-O(\delta))$. Thus, it follows that

$$\frac{d^{(t-1)}(w)(1-O(\delta))}{M_{t-1}} \leq \frac{|A_w|}{|B_w|} \leq \frac{d^{(t-1)}(w)}{M_{t-1} - \delta M}.$$

Therefore, since $d^{(t-1)}(w) \leq d(w) \leq \delta\sqrt{M}$ and $M_{t-1} = \Theta(M)$,

$$\mathbb{P}_{t-1}\left(w_t = w \middle| w_t \in \overline{S}^+\right) = \frac{|A_w|}{|A_w| + |B_w|} = \frac{d^{(t-1)}(w)(1 \pm O(\delta))}{M_{t-1} + d^{(t-1)}(w)(1 \pm O(\delta))} = \frac{d^{(t-1)}(w)}{M_{t-1}}(1 \pm O(\delta)).$$

This completes the proof, except for the proofs of the claims. We now give these proofs.

Proof of Claim 7.4.3. Let C_w be the set of $G^* \in B_w$ such that $v_t w \in E(G^*)$. We define a switching between C_w and $B_w \backslash C_w$. Suppose $G^* \in C_w$. Let (a, b) be an ordered pair of adjacent vertices such that $ab \notin E_{t-1}$. Then the switching deletes edges $v_t w$ and ab, replacing them with edges $v_t a$ and wb, hence creating G'^*. The choice of (a, b) is considered valid if and only if $G'^* \in B_w \backslash C_w$.

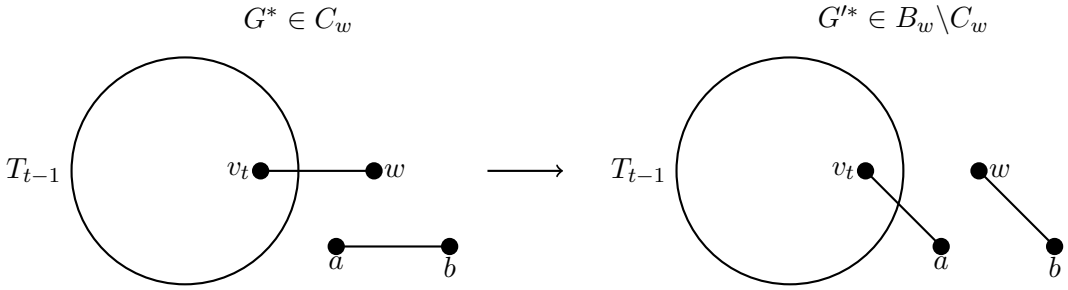

Figure 7.4: A diagram of the switching used in the proof of Claim 7.4.3, which sends $G^* \in C_w$ to $G'^* \in B_w \backslash C_w$.

There are M_{t-1} choices for an ordered pair of adjacent vertices (a, b) such that $d^{(t-1)}(a) \geq 1$ and $ab \notin E_{t-1}$. Then for such a choice of (a, b), the set $\{v_t, w, a, b\}$ corresponds to a valid switching if and only if

(a) $v_t a \notin E(G^*)$ and $wb \notin E(G^*)$,

(b) $\{a, b\} \cap \{v_t, w\} = \emptyset$.

Let $W(G^*)$ be the number of choices for (a, b) such that one of the above conditions is not satisfied. Then the number of switchings that can be applied to G^* is $M_{t-1} - W(G^*)$. Let $Z(G^*)$ be the number of choices for (a, b) such that

(a′) a or $b \in \overline{S}^+$,

(b′) $a, b \notin \overline{S}^+$, $v_t a \in E(G^*)$ or $wb \in E(G^*)$,

(c′) $a, b \notin \overline{S}^+$, $\{a, b\} \cap \{v_t, w\} \neq \emptyset$.

By similar counting arguments to the previous switching, it follows that $Z \leq 2\delta^2 n + 2\delta^2 M + 3\delta\sqrt{M} \leq \delta M$. Now we count reverse switchings. Note that a reverse switching corresponds to choosing a neighbour a of vertex v_t (where $a \neq w_t$ and $v_t a \notin E_{t-1}$) and a neighbour b of w, deleting the edges $v_t a$ and wb, and replacing them with $v_t w$ and ab to create some $G^* \in C_w$. Thus, there are at most $d(v_t)d(w)$ reverse switchings that create each G^*. Therefore,

$$|C_w| \leq |B_w \backslash C_w| \frac{2d(w)d(v_t)}{M_{t-1}}.$$

Condition (P2) implies that $d(v_t) \leq \delta\sqrt{M}$. Since we assume that $w \notin \overline{S}^+$, we also know that $d(w) \leq \delta\sqrt{M}$. Recalling that $M_{t-1} = \Theta(M)$ from condition (P1), this implies that there exists some constant $C > 0$ such that

$$\frac{|C_w|}{|B_w \backslash C_w|} \leq C\delta^2.$$

Therefore,

$$\mathbb{P}_{t-1}\left(v_t w \in E(G^*) \middle| G^* \in B_w\right) = \frac{|C_w|}{|B_w|} \leq C\delta^2.$$

This completes the proof of the claim. □

Proof of Claim 7.4.4. Let F_w be the set of all graphs in B_w such that $w_t w \in E(G^*)$, and let F'_w be its complement in B_w. We define a switching from F_w to F'_w to prove the claim. Let $G^* \in F_w$. Choose an ordered pair of vertices (a, b) such that

(a) $ab \in E(G^*)$ and $ab \notin E_{t-1}$,

(b) $aw_t, bw \notin E(G^*)$, and

(c) $\{a, b, w_t, w, v_t\}$ are all distinct.

The switching then deletes edges $w_t w$ and ab from $E(G^*)$ and adds in edges aw_t and bw, creating a new $G'^* \in F'_w$. A reverse switching is equivalent to taking edges aw_t and bw, deleting these edges, and replacing them with the edges $w_t w$ and ab. Thus, there are at most $d(w)d(w_t)$ such switchings that map an element of F_w to a given $G'^* \in F'_w$, or equivalently at most $d(w)d(w_t)$ reverse switchings can be applied to each $G'^* \in F'_w$.

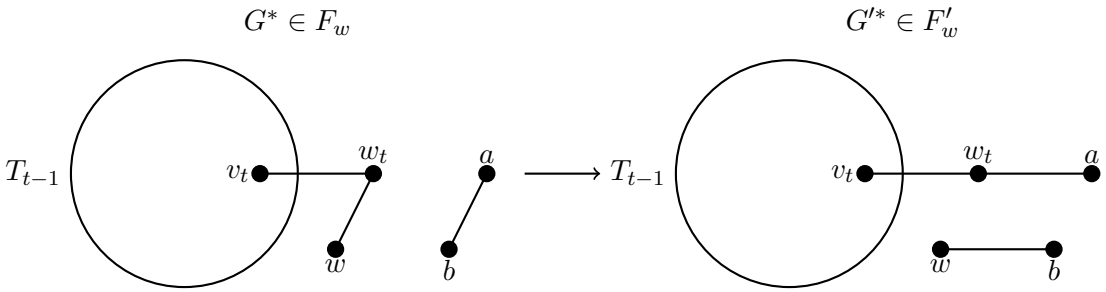

Figure 7.5: A diagram of the switching used in the proof of Claim 7.4.4, which sends $G^* \in F_w$ to $G'^* \in F'_w$.

Now we give a lower bound on the number of switchings that can be applied to each $G^* \in F_w$. There are M_{t-1} choices for an ordered pair of adjacent vertices (a, b) such that $d^{(t-1)}(a) \geq 1$ and $ab \notin E_{t-1}$. The number of such choices where $a \in \overline{S}^+$ or $b \in \overline{S}^+$ is at most $2\delta^2 n$. Of the remaining $M_{t-1} - 2\delta^2 n$ choices, at most $\delta\sqrt{M}(d(w_t) + d(w))$ do not satisfy (b), and at most $\delta\sqrt{M} + 2(d(w_t) + d(w))$ do not satisfy (c). Recall that $d(w), d(w_t) \leq \delta\sqrt{M}$, since $w, w_t \notin \overline{S}^+$ by assumption. Thus, there are at least $M_{t-1} - 5\delta^2 M$ switchings that can be applied to each $G^* \in F_w$. Altogether, this implies that

$$\frac{|F_w|}{|F'_w|} \leq \frac{d(w)d(w_t)}{M_{t-1} - 5\delta^2 M} = O(\delta^2),$$

where the last inequality follows from the facts that $d(w), d(w_t) \leq \delta\sqrt{M}$ and $M_{t-1} = \Theta(M)$ (as T is predictable). This completes the proof. □

Proof of Claim 7.4.5. This proof is in two parts. Recall the definition of C_w from the proof of Claim 7.4.3, the subset of B_w such that $v_t w \in E(G^*)$ for all $G^* \in C_w$. Firstly we show that, conditional on $G^* \in B_w \backslash C_w$, with probability at least $1 - O(\delta^2)$ at most $\frac{1}{2}\delta d^{(t-1)}(w)$ undiscovered neighbours of w in $S_{t-1} \cup \overline{S}_{t-1} \backslash \overline{S}^+$ are also neighbours of w_t. Applying Claim 7.4.3 gives us the desired result for $G^* \in B_w$. Secondly we show that, conditional on $G^* \in B_w$, with probability $1 - O(\delta)$ the vertex w has at most $\frac{1}{2}\delta d^{(t-1)}(w)$ undiscovered neighbours in $\overline{S}^+$ (and thus at most this many are also neighbours of w_t). Each proof is done with its own switching. These results combine to prove the claim by the union bound for $G^* \in B_w$.

Let $D^i_w \subset B_w \backslash C_w$ be the set of $G^* \in B_w \backslash C_w$ such that there are exactly i edges between w and the neighbours z of w_t such that z is not in $\overline{S}^+$ and $wz \notin E_{t-1}$. We define a switching from D^{i+1}_w to D^i_w. Suppose $G^* \in D^{i+1}_w$, and let $z \in S_{t-1} \cup \overline{S}_{t-1} \backslash \overline{S}^+$ be a vertex such that $w_t z w$ is a path in G^* and $wz \notin E_{t-1}$. Let (a, b) be an ordered pair of adjacent vertices such that $d^{(t-1)}(a) \geq 1$ and $ab \notin E_{t-1}$. Then a switching takes the edges wz and ab, deletes both of these edges, and replaces them with aw and bz, creating a new pairing G'^*. This switching is valid if and only if $G'^* \in D^i_w$, which occurs if and only if

(a) $aw, aw_t, bz \notin E(G^*)$,

(b) the vertices $\{v_t, w_t, w, z, a, b\}$ are all distinct.

We forbid the edge aw_t because if $aw_t \in E(G^*)$, then the output of the switching G'^* would contain the path waw_t and thus would still be in D^{i+1}_w, not D^i_w. Now we determine a lower bound

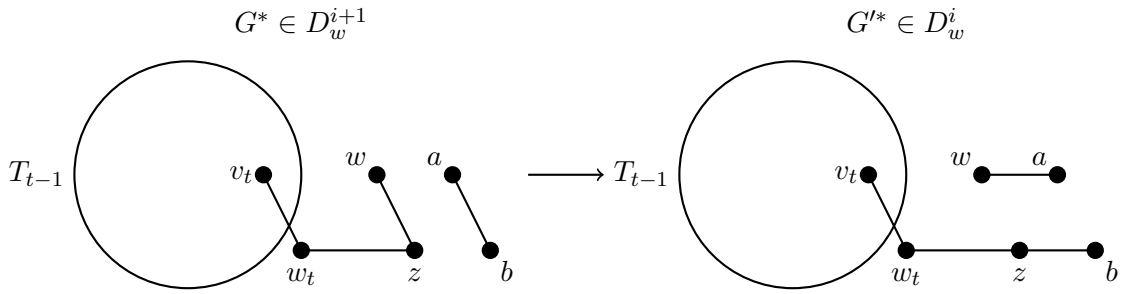

Figure 7.6: A diagram of one of the switchings used to prove Claim 7.4.5, which sends $G^* \in D_w^{i+1}$ to $G'^* \in D_w^i$.

on the number of valid switchings that can be applied to each $G^* \in D_w^{i+1}$. There are $i+1$ choices for the vertex z by definition, and for each z there are M_{t-1} choices for an ordered pair of vertices (a, b) such that $ab \in E(G^*) \backslash E_{t-1}$. As is standard, we bound from above the number of choices for (a, b) such that one the following occur:

(a′) $a, b \in \overline{S}^+$,

(b′) $a, b \notin \overline{S}^+$ and $aw \in E(G^*)$ or $aw_t \in E(G^*)$ or $bz \in E(G^*)$,

(c′) $a, b \notin \overline{S}^+$ and the vertices $\{v_t, w_t, w, z, a, b\}$ are not all distinct.

There are at most $2\delta^2 n$ choices that do not satisfy (a′). For case (b′), since $w, w_t, z \notin \overline{S}^+$, we know that $d(a), d(b), d(w), d(w_t), d(z) \leq \delta\sqrt{M}$. Thus, for each choice of z there are at most $3\delta^2 M$ choices for (a, b) such that case (b′) occurs. Finally, for case (c′), we consider the possible ways that the vertices $\{v_t, w_t, w, z, a, b\}$ can be non-distinct. The vertices v_t and z are distinct, as $D_w^{i+1} \subseteq B_w \backslash C_w$ implies that $v_t w \notin E(G^*)$. Thus, the vertices $\{v_t, w, w_t, z\}$ are all necessarily distinct. The vertices a and b are distinct from each other, as they are adjacent. Since $d^{(t-1)}(a) \geq 1$, the vertices v_t and a are distinct. Thus, the remaining possibilities are if $b = v_t$, or if $\{a, b\} \cap \{w, w_t, z\} \neq \emptyset$. At most $\delta\sqrt{M}$ choices for (a, b) satisfy each such combination, and thus there are at most $7\delta\sqrt{M}$ choices that satisfy (c′). Therefore, the number of switchings that can be applied to each $G^* \in D_w^{i+1}$ is at least $(i+1)(M_{t-1} - 4\delta^2 M)$.

Now we count the number of reverse switchings that can be applied to each $G'^* \in D_w^i$. Note that a reverse switching is equivalent to the following: choose a vertex $z \in S_{t-1} \cup \overline{S}_{t-1} \backslash \overline{S}^+$ which is a neighbour of w_t such that $wz \notin E_{t-1}$ and z is not a neighbour of w. Then choose one neighbour a of w and one neighbour b of z. A reverse switching then deletes $\{aw, bz\}$ from $E(G'^*)$ and adds in edges $\{ab, wz\}$ to create some G^*. With this in mind, there are at most $d^{(t-1)}(w)\delta^2 M$ reverse switchings that can be applied to each $G'^* \in D_w^i$, since $d(w_t), d(z) \leq \delta\sqrt{M}$ by assumption and there are at most $d^{(t-1)}(w)$ choices for such a vertex a.

Thus, for all $i \leq d^{(t-1)}(w)$,

$$\frac{|D_w^{i+1}|}{|D_w^i|} \leq \frac{2d^{(t-1)}(w)\delta^2 M}{(i+1)M_{t-1}} = \frac{d^{(t-1)}(w)}{(i+1)} O(\delta^2), \tag{7.4}$$

since $M_{t-1} = \Theta(M)$. Let $D_w^* = \cup_{i \geq i_0+1} D_w^i$, where $i_0 = \frac{1}{2}\delta d^{(t-1)}(w)$. Applying (7.4) iteratively

gives that

$$|D_w^*| = \sum_{i \geq i_0+1} |D_w^i| \leq |D_w^{i_0}| \sum_{i \geq i_0+1} \left(\frac{d^{(t-1)}(w)}{(i+1)} O(\delta^2) \right)^i \leq |D_w^{i_0}| \sum_{i \geq i_0+1} (O(\delta))^i = |D_w^{i_0}| O(\delta),$$

since $\delta = o(1)$. Thus, $\mathbb{P}(D_w^* | B_w \backslash C_w) = O(\delta)$. Therefore, conditional on $B_w \backslash C_w$, with probability $1 - O(\delta)$ the vertices w and w_t have at most $\delta d^{(t-1)}(w)$ common neighbours z such that $z \notin \overline{S}^+$ and $wz \notin E_{t-1}$. Since Claim 7.4.3 implies that $|C_w|/|B_w| = O(\delta^2)$, this proves the first half of the claim.

Now we prove the second part of the claim. Note that due to our choice of w, E_{t-1} contains no edges between w and $\overline{S}^+$, that is, all neighbours of w in $\overline{S}^+$ are undiscovered in T. For the second part of the claim, we show that, conditional on $G^* \in B_w$, the probability that the vertex w has less than $\frac{1}{2}\delta d^{(t-1)}(w)$ neighbours in $\overline{S}^+$ is $1 - O(\delta)$. This immediately implies that w_t and w have at most this many common neighbours in $\overline{S}^+$. Let $\mathcal{D}_w^i \subset B_w$ such that w has exactly i neighbours in $\overline{S}^+$ for $i \in \{0, \dots, d^{(t-1)}(w)\}$. We define a switching between $\mathcal{D}_w^{i+1}$ and $\mathcal{D}_w^i$ as follows. Let $G^* \in \mathcal{D}_w^{i+1}$. Let wy be an edge such that $y \in \overline{S}^+$ and $wy \notin E_{t-1}$. Let (a, b) be an ordered pair of vertices such that $ab \in E(G^*)$, where $a \notin \overline{S}^+$ and $ab \notin E_{t-1}$. Then the switching deletes the edges ab and wy and adds in aw and by, creating a new pairing $G'^* \in \mathcal{D}_w^i$.

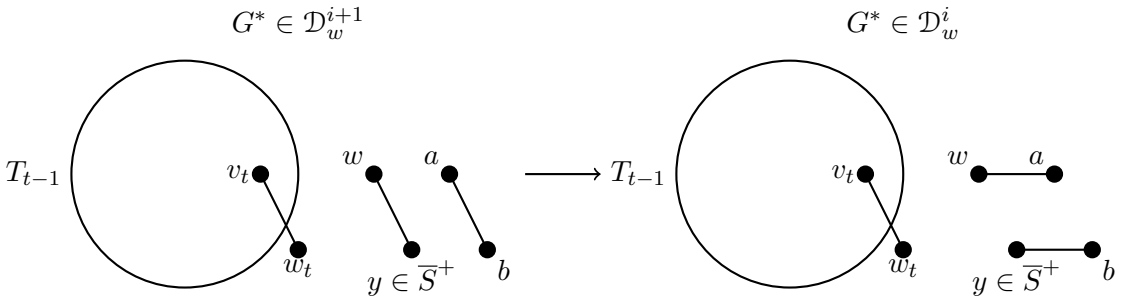

Figure 7.7: A diagram of the second switching used to prove Claim 7.4.5, which sends $G^* \in \mathcal{D}_w^{i+1}$ to $G'^* \in \mathcal{D}_w^i$.

The number of switchings that can create each $G'^* \in \mathcal{D}_w^i$ is at most $d^{(t-1)}(w) d(\overline{S}^+)$, which is the number of ways of choosing a neighbour a of w where $aw \notin E_{t-1}$ and an edge by such that $y \in \overline{S}^+$. Now we determine a lower bound on the number of switchings that can be applied to each $G^* \in \mathcal{D}_w^{i+1}$. By definition, there are exactly $i+1$ choices for the edge wy, since w is fixed. If (a, b) is an ordered pair of adjacent vertices, then $\{w, y, a, b\}$ corresponds to a valid switching if the output is in $\mathcal{D}_w^i$, which occurs if

(a) $ab \notin E_{t-1}$,

(b) $a \notin \overline{S}^+$,

(c) $aw, by \notin E(G^*)$,

(d) $\{a, b, w, y\}$ are all distinct.

Condition (P5) implies that there are at least $\frac{2}{3}n - 2$ choices for some vertex $b \notin V_{t-1} \cup \{w, y\}$. Since $d(\overline{S}^+) \leq \delta^2 n$, it follows that $o(n)$ of these choices for b are in $\overline{S}^+$ or have a neighbour in

$\overline{S}^+$. Of the remaining choices for b, at most $\delta^2 M$ of these are adjacent to w or have a common neighbour with w (since b has no neighbours in $\overline{S}^+$). For each of the remaining choices for b, there exists at least one choice for the vertex a such that $ab \in E(G^*)$, $a \notin \overline{S}^+$, $a \neq w$, and $aw \notin E(G^*)$. Thus, for each choice of an edge wy there are at least $\frac{2}{3}n - o(n)$ choices for (a, b) such that $\{w, y, a, b\}$ corresponds to a valid switching. Thus,

$$\frac{|\mathcal{D}_w^{i+1}|}{|\mathcal{D}_w^i|} \leq \frac{d^{(t-1)}(w)d(\overline{S}^+)}{\frac{2}{3}(i+1)n}(1+o(1)) \leq \frac{3\delta^2 d^{(t-1)}(w)}{2(i+1)}(1+o(1)),$$

since $d(\overline{S}^+) \leq \delta^2 n$. If $i+1 \geq \frac{1}{2}\delta d^{(t-1)}(w)$, then the above ratio is at most 4δ. Therefore,

$$\mathbb{P}\left(\cup_{i \geq \frac{1}{2}\delta d^{(t-1)}(w)} \mathcal{D}_w^i \,\middle|\, B_w\right) \leq \sum_{i \geq 1} (4\delta)^i = O(\delta).$$

This completes the second part of the proof of the claim, and the full statement of the claim follows from the union bound. □

With these claims proved, the proof of the lemma is complete. □

Remark 7.4.6. Suppose that $T_{t-1} = (V_{t-1}, E_{t-1}, \mathcal{X}_{t-1})$ is a partial pairing. Let $\mathcal{F}_{t-1} \subset \overline{S}_{t-1}$ be the set of $w \in \overline{S}_{t-1}$ such that $v_t w \in E_{t-1}$ and $w \notin \overline{S}^+$. If $w \in \mathcal{F}_{t-1}$, then $\mathbb{P}_{t-1}(w_t = w) = 0$. This is because the pairings in Φ are simple, and if $w_t = w$ then there would be a double edge between v_t and w. We say that such $w \in \mathcal{F}_{t-1}$ are *forbidden* vertices (at time $t-1$), as they cannot be explored by the process at time t. If T_{t-1} is predictable, then (P2) states that $d(v_t) \leq \delta\sqrt{M}$ for all $t > 0$, and thus $d(\mathcal{F}_{t-1}) \leq \delta^2 M$. This bound is useful for analysing $\mathbb{E}_{t-1}[\mathsf{d}_S(w_t) - 2]$, particularly in the supercritical regime. An analogous definition could be constructed for $w \in \overline{S}^+$ too, however this is not necessary as $\mathbb{P}_{t-1}\left(w_t \in \overline{S}^+\right)$ is small if T_{t-1} is predictable.

The framework for the analysis in the supercritical and subcritical regimes is similar. In either case, we define the initial sets V_0 and E_0 such that T_{t-1} is predictable for a large window of time $t \geq 1$. Then we adapt the proof from the corresponding case in the configuration model. We show that, in spite of the switching errors, the exploration process and the value of X_{t-1} (recall $X_{t-1} = |\mathcal{X}_{t-1}|$) evolve similarly to their configuration model equivalents.

7.5 Analysis in the subcritical regime

In this section we prove Theorem 7.1.1. The assumptions in the theorem that $d(S) = \Theta(M)$ and $R_S = o(M)$ imply that $d(S \backslash \overline{S}^+) = \Theta(M)$. Furthermore, Lemma 6.3.1 implies that if $R_S = o(M)$, then $M = \Theta(|S|)$. Recall from Remark 7.3.1 the various conditions that we impose on δ, specifically that $\delta \to 0$, $R_S \leq \delta M$, $d(\overline{S}^+) \leq \delta^2 n$, and $\delta = \omega(\log^{-1} M)$. We assume these conditions, often implicitly, throughout this section.

7.5.1 Preprocessing and setup

As in the configuration model case, we start by defining the preprocessing step $t = 0$. We now define the sets V_0 and E_0 for the exploration process in the subcritical regime. They are similar

to their configuration model analogues defined in Section 6.2.4, but include some extra vertices and edges to help simplify the later analysis.

Let U be the preprocessing set as defined in (6.7), that is, the smallest set of the highest degree vertices in S with total degree greater than $5\delta^{1/4}M$. Let $U^* \subset U$ be the set of vertices with degree greater than $\delta\sqrt{M}$. Let W be the set of neighbours of U^* in $S\backslash U$. Let W' be the set of neighbours of U^* in $\overline{S}$. Define $U_1 = U \cup W$, and then define $V_0 := U_1 \cup W' \cup \{v\}$ for some arbitrary $v \in S$. Define $E_0 := E(G^*[U_1]) \cup E(U^*, W')$, where $E(U^*, W')$ denotes the set of edges $uw \in E(G^*)$ such that $u \in U^*$ and $w \in W'$ and $E(G^*[U_1])$ is the set of edges in G^* with both ends in U_1 (equivalently, the set of edges of the induced pairing $G^*[U_1]$).

Importantly, Lemma 6.2.11 from the configuration model proof still applies in the random graph case (with the same U definition as in (6.7) and an identical proof). This means that the maximum degree of a vertex in $S\backslash U$ is $\delta^{-1/4}$. This implies that $U^* = S^+$, and thus E_0 contains every edge that is incident to a vertex in S^+. We use this to prove that the new preprocessing set U_1 is only negligibly larger than U, and thus all the important bounds on $d(U)$ used in the configuration model immediately carry over to U_1. From here, the exploration process evolves as defined in Section 7.3.

Lemma 7.5.1. Suppose that $R_S \leq \delta M$. Under the assumptions of Theorem 7.1.1, $d(U^*) \leq 2\delta M$ and $d(W) \leq 2\delta^{3/4}M$.

Proof. We first show that $d(U^*) \leq 2\delta M$. Suppose for contradiction that $d(U^*) > 2\delta M$. It follows that

$$\sum_{u \in U^*} d(u)\left(\mathtt{d}_S(u) - 2\right) > \left(\delta\sqrt{M} - 2\right) \sum_{u \in U^*} d(u) > \delta^2 M^{3/2} = \omega(M).$$

Note that if U^* contains a vertex of degree greater than δM, then immediately we get a contradiction that $R_S > \delta M$. So we suppose that U^* contains no vertex of degree greater than δM. Let $Q \subset U^*$ be the smallest set of the lowest degree vertices in U^* with total degree at least δM, and let q be the highest degree vertex in Q. Then

$$\sum_{u \in Q} d(u)\left(\mathtt{d}_S(u) - 2\right) > \left(\delta\sqrt{M} - 2\right) \sum_{u \in Q} d(u) > \frac{1}{2}\delta^2 M^{3/2}.$$

Since $\sum_{i=1}^{j} d(i)(\mathtt{d}_S(i) - 2) \geq -M$ for all $j \in [n]$, this immediately implies that $j_S \leq q$, and thus

$$R_S \geq d(q) + \sum_{u \in U^*\backslash Q} d(u) > \delta M.$$

This is a contradiction since $R_S \leq \delta M$ by the assumption of Theorem 7.1.1. Thus, it must hold that $d(U^*) \leq 2\delta M$. Then the total degree of W is at most $\delta^{-1/4} d(U^*)$, since $W \subset S\backslash U$ and the maximum degree in S outside of U is $\delta^{-1/4}$. Therefore, it immediately follows that $d(W) \leq 2\delta^{3/4}M$. □

The following corollary contains many useful bounds on the maximum degree of various vertices and sets of vertices. These bounds will be used extensively throughout the following proofs, and thus we compile them all here.

Corollary 7.5.2. For a pair $(\boldsymbol{d}, S)$ satisfying the conditions of Theorem 7.1.1 such that $R_S \leq \delta M$, the following statements are true:

(a) $d(U_1) \leq 6\delta^{1/4}M$,

(b) the maximum degree of a vertex in $S\backslash U_1$ is at most $\delta^{-1/4}$,

(c) For all $t > 0$, $d(v_t) \leq \delta\sqrt{M}$.

Proof. Part (a) simply follows from Lemma 7.5.1, since it implies that

$$d(U_1) = d(U) + d(W) \leq 5\delta^{1/4}M + 2\delta^{3/4}M \leq 6\delta^{1/4}M.$$

Part (b) follows immediately since $U \subset U_1$, and the maximum degree of any vertex in $S\backslash U$ is at most $\delta^{-1/4}$. Part (c) then follows from the fact $d^{(t)}(u) = 0$ for all $u \in U^*$ and $t \geq 0$, by the definition of V_0 and E_0. Thus, $v_t \in S\backslash S^+$ for all $t > 0$, which immediately implies the desired claim. □

Thus, all the bounds in the configuration model that were based on the size of U apply with minimal alteration if we swap out U with U_1. Intuitively, this means that also placing W and U^* into V_0 does not affect the evolution of the exploration process by too much. However, now we have a tighter control on $\mathtt{d}_S(w_t)$, as well as $d(v_t)$ at each step in the exploration process. This helps with the switching arguments and also helps to show that X_t, the number of open edges at some time t, is concentrated. These points are formalised later.

Now we show that, for sufficiently many time steps $t > 0$, the partial pairing T_{t-1} is predictable, that is, satisfies (P1) to (P5). Note that an analogous statement to Observation 6.2.10 still applies to the exploration process on $G^* \in \Phi$. The caveat is that we must now also take into account the edges that are revealed in the preprocessing step.

Observation 7.5.3. For all $t \geq 1$, $M_t \geq d(\overline{S}_{t-1}) \geq d(\overline{S}) - t - |W^*| \geq d(\overline{S}) - t - 6\delta^{1/4}M$.

This is because we do not reveal back edges when $w_t \in \overline{S}$, and thus for each time $t \geq 1$ there are two cases: if $w_t \in S$, then $d(\overline{S}_t) = d(\overline{S}_{t-1})$, and if $w_t \in \overline{S}$, then $d(\overline{S}_t) = d(\overline{S}_{t-1}) - 1$. When $t = o(M)$, this implies that $M_{t-1} = \Theta(M)$ and $d(\overline{S}_{t-1}) = \Theta(M)$, since $d(\overline{S}) = \Theta(M)$ by assumption.

Proposition 7.5.4. Suppose that $M = \Theta(n)$, $d(\overline{S}^+) \leq \delta^2 n$. If $t = o(M)$ and $X_{t-1} \leq \alpha n$ for some $\alpha < 10^{-3}$, then T_{t-1} is predictable for all $t > 0$.

Proof. Under these conditions, Observation 7.5.3 immediately implies that (P1) is satisfied. Corollary 7.5.2 implies that (P2) and (P3) are satisfied, as well as implying that $|V_0| \leq 6\delta^{1/4}M$. Since $M = \Theta(n)$ and $|V_{t-1}| \leq |V_0| + t = o(M)$, this implies that (P5) is satisfied. Finally, (P4) is satisfied by assumption. □

The next two lemmas are thus an immediate consequence of this proposition and Lemmas 7.4.1 and 7.4.2 respectively. Recall that $\mathbb{P}_{t-1}(\cdot)$ is the probability of an event conditional on the partial pairing at time $t-1$, and recall from (7.3) the definition of $d^{(t)}(w)$, the available degree of a vertex w at time t.

Lemma 7.5.5. Suppose that $M = \Theta(n)$, $d(\overline{S}^+) \leq \delta^2 n$. If $t = o(M)$ and $X_{t-1} \leq \alpha n$ for some $\alpha < 10^{-3}$, then $\mathbb{P}_{t-1}\left(w_t \in \overline{S}^+\right) \leq \delta$.

Lemma 7.5.6. Suppose that the conditions of Theorem 7.1.1 are satisfied and $R_S \leq \delta M$. Suppose that $t = o(M)$ and $X_{t-1} \leq \alpha n$ for some $\alpha < 10^{-3}$. Let $w \notin \overline{S}^+$ be a vertex such that $d^{(t-1)}(w) \geq 1$ and $v_t w \notin E_{t-1}$. Then

$$\mathbb{P}_{t-1}\left(w_t = w \middle| \, w_t \notin \overline{S}^+\right) = \frac{d^{(t-1)}(w)}{M_{t-1}}\left(1 \pm O(\delta)\right).$$

7.5.2 Process evolution in the subcritical regime

In this section we study how X_t, the number of open edges in T_t, evolves over time throughout the deferred decision process in the subcritical regime on the set of simple pairings Φ. Specifically, we show that it behaves very similarly to the corresponding process in the configuration model $\mathcal{C}(\boldsymbol{d})$. As in the configuration model, we bound X_t from above for all $t \geq 0$ by

$$X'_t = X'_0 + \sum_{i=1}^{t} (\mathsf{d}_S(w_t) - 2)$$

where $X'_0 = \sum_{u \in U_1 \cup \{v\}} d(u)$. This is an upper bound on X_t since it does not consider half edges that are consumed by revealing back edges or edges between vertices in V_0. So if $X'_t \leq 0$, then there exists some time $t' \leq t$ such that $X_{t'} = 0$. In this section, we show that $X'_{\delta^{1/18}M} \leq 0$ with probability $1 - o(M^{-1})$. We then use this to prove Theorem 7.1.1.

Recall that $S_{t-1} = S \backslash V_{t-1}$, and write $V = V(G^*)$ for brevity. Define

$$f_t := -1 + \sum_{w \in S_{t-1}} \frac{d(w)}{M_{t-1}}(d(w) - 1) = \sum_{w \in V} \frac{d^{(t-1)}(w)}{M_{t-1}}\left(\mathsf{d}_S(w) - 2\right). \tag{7.5}$$

This is equivalent to the definition of f_t in the configuration model case; the second form accounts for the fact that vertices in $\overline{S}$ are no longer "exploded" into degree 1 vertices. In contrast to the configuration model process, it is not true that $f_t = \mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t) - 2\right]$. This is because $\mathbb{P}_{t-1}\left(w_t = w\right)$ differs between the two probability spaces $\mathcal{C}(\boldsymbol{d})$ and Φ. However, Lemma 7.5.6 argues that for $w \notin \overline{S}^+$ and $w \notin \mathcal{F}_{t-1}$ the probability of discovering w at time $t-1$ in the configuration model is a good approximation of the probability of the same event in Φ. So intuitively, one might assume that f_t behaves similarly in both models. To account for $w \in \overline{S}^+$ or $w \in \mathcal{F}_{t-1}$, Lemma 7.5.5 implies that the probability of discovering a vertex in $\overline{S}^+$ is small, and as discussed in Remark 7.4.6 the set of forbidden vertices at each time step is both a subset of $\overline{S}$ and also has small total degree. This intuitively means that these vertices do not significantly affect the value of $\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t) - 2\right]$. First we use Lemma 7.5.6 to show that if $t = o(M)$ and $X_{t-1} \leq \alpha n$ for some $\alpha < 10^{-3}$, then $\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t) - 2\right]$ is within a small approximation error of f_t.

Lemma 7.5.7. If $t = o(M)$ and $X_{t-1} \leq \alpha n$ for some $\alpha < 10^{-3}$, then $\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t) - 2\right] = f_t(1 \pm O(\delta)) + O(\delta)$.

Proof. Since $\mathsf{d}_S(w) = 1$ for all $w \notin S$ and $\mathbb{P}_{t-1}\left(w_t = w\right) = 0$ if $d^{(t-1)}(w) = 0$, we can express

$\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t) - 2\right]$ as

$$\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t) - 2\right] = -1 + \sum_{w \in S_{t-1}} \mathbb{P}_{t-1}\left(w_t = w\right)\left(d(w) - 1\right).$$

By assumption, Lemma 7.5.6 applies. Since $d^{(t-1)}(w) = d(w)$ for all $w \in S_{t-1}$,

$$\begin{aligned}
\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t) - 2\right] &= -1 + \sum_{w \in S_{t-1}} \frac{d(w)}{M_{t-1}}(d(w) - 1)(1 \pm O(\delta)) \\
&= -1 + (1 \pm O(\delta)) \sum_{w \in S_{t-1}} \frac{d(w)}{M_{t-1}}\left(d(w) - 1\right) \\
&= -1 + \sum_{w \in S_{t-1}} \frac{d(w)}{M_{t-1}}\left(d(w) - 1\right) \pm O(\delta) \sum_{w \in S_{t-1}} \frac{d(w)}{M_{t-1}}(d(w) - 1) \\
&= f_t \pm O(\delta)\left(f_t + 1\right).
\end{aligned}$$

This completes the proof. □

If f_t is sufficiently negative (specifically, $\omega(\delta)$ in magnitude) and the conditions on t and X_{t-1} are satisfied, then the above lemma implies that $\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t) - 2\right] = f_t(1 + o(1))$. We show that a.a.s. this is the case for all $t \leq \delta^{1/19}M$. This is done in a way that mimics the configuration model results about f_t. However, there is slightly more work to be done, as we have to also keep track of X_{t-1} throughout the process and make sure it does not grow too large. First we show that f_1 is bounded above by some function that is $-\omega(\delta)$.

Lemma 7.5.8. If $R_S \leq \delta M$ for $\delta \to 0$, $f_1 \leq -4\delta^{1/4}$.

Proof. It is useful to express f_1 with some redundant terms. Recall the set W, the set of neighbours of S^+ which are in S, and W', the set of neighbours of S^+ which are in $\overline{S}$. Then the definitions of V_0 and E_0 imply that

$$\begin{aligned}
f_1 = \frac{1}{M_0}\sum_{i=1}^{u-1} d(i)(\mathsf{d}_S(i) - 2) - \frac{d(v)(d(v)-2)}{M_0}\mathbb{1}_{\{v \leq u-1\}} \\
- \frac{1}{M_0}\sum_{w \in W} d(w)\left(\mathsf{d}_S(w) - 2\right) + \frac{|E(U^*, W')|}{M_0}.
\end{aligned} \tag{7.6}$$

Lemma 7.5.1 states that $d(U^*) \leq 2\delta M$ and $d(W) \leq 2\delta^{3/4}M$. This implies that $|E(U^*, W')| \leq d(U^*) \leq 2\delta M$. This combined with Corollary 7.5.2 implies that $M_0 = M(1 - O(\delta^{1/4}))$.

By definition, V_0 contains U_1 (and thus also contains U) and so has total degree at least $5\delta^{1/4}M$. Recall that u is defined to be the lowest-indexed vertex in U. The definition of U implies that $u < j_S$ and $U \subset S$. First suppose that U contains only vertices of degree at least 3. Then

$$\sum_{i=u}^{j_S-1} d(i)(\mathsf{d}_S(i) - 2) = \sum_{i=u}^{j_S-1} d(i)(d(i) - 2) \geq \sum_{i=u}^{j_S-1} d(i).$$

Since $\sum_{i=1}^{j_S-1} d(i)(\mathsf{d}_S(i)-2) \leq 0$ by the definition of j_S, this implies that

$$\begin{aligned}\sum_{i=1}^{u-1} d(i)(\mathsf{d}_S(i)-2) &= \sum_{i=1}^{j_S-1} d(i)(\mathsf{d}_S(i)-2) - \sum_{i=u}^{j_S-1} d(i)(\mathsf{d}_S(i)-2)\\ &\leq 0 - \sum_{i=u}^{j_S-1} d(i)\\ &= \sum_{i=j_S}^{n} d(i) - \sum_{i=u}^{n} d(i)\\ &\leq -5\delta^{1/4}M + \delta M.\end{aligned}$$

Thus, it follows from Equation (7.6) that

$$\begin{aligned}f_1 &\leq \frac{1}{M_0}\sum_{i=1}^{u-1} d(i)(\mathsf{d}_S(i)-2) - \frac{d(v)}{M_0}(d(v)-2)\mathbb{1}_{\{v\leq u-1\}} + 3\delta\\ &\leq -\frac{9}{2}\delta^{1/4} - \frac{d(v)}{M_0}(d(v)-2)\mathbb{1}_{\{v\leq u-1\}}\\ &\leq -4\delta^{1/4}.\end{aligned}$$

This proves the lemma in the case that U only contains vertices of degree at least 3. Now suppose that U contains a vertex of degree 2 or lower. If this is the case, then every vertex in $S\backslash U$ must have degree at most 2. Then Equation (7.6) implies that

$$\begin{aligned}f_1 &\leq \frac{1}{M_0}\sum_{i=1}^{u-1} d(i)(\mathsf{d}_S(i)-2) - \frac{d(v)(d(v)-2)}{M_0}\mathbb{1}_{\{v\leq u-1\}} - \frac{1}{M_0}\sum_{w\in W} d(w)\left(\mathsf{d}_S(w)-2\right) + 3\delta\\ &\leq -\frac{d(\overline{S}_0)}{M_0} - \frac{d(v)(d(v)-2)}{M_0}\mathbb{1}_{\{v\leq u-1\}} - \frac{1}{M_0}\sum_{w\in W} d(w)\left(\mathsf{d}_S(w)-2\right) + 3\delta\\ &= -\Theta(1).\end{aligned}$$

Since $\delta \to 0$, this concludes the proof. □

We next show that $\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t)-2\right]$ decreases at a similar rate to its configuration model counterpart. Specifically, we prove an analogue to Lemma 6.2.13. This is given in the next lemma. Define a sequence of random variables $(A_t)_{t\geq 1}$, where $A_t := \mathsf{d}_S(w_t) - \mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t)\right]$. Recall the definition of the sequence of random variables $(G_t)_{t\geq 1}$ from the proof of the configuration model result, where $G_t = f_{t+1} - f_t - \mathbb{E}_{t-1}\left[f_{t+1}-f_t\right]$. We defer the proof of this lemma for now.

Lemma 7.5.9. Let $(T_t)_{t\geq 1}$ be a sequence of partial pairings such that for all t, $\left|\sum_{t'\leq t} A_{t'}\right| > M^{2/3}$ and $\left|\sum_{t'\leq t} G_{t'}\right| \leq \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{t}$. Then there exists a constant $c>0$ such that for all $t \leq \delta^{1/19}M$, $\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t)-2\right] \leq -\frac{ct}{M_{t-1}}$.

The extra condition on $\sum_{t'\leq t} A_t$ will be used to ensure that $X_t = o(M)$ for all t, and thus T_t is predictable for all $t \leq \delta^{1/19}M$. This allows us to use Lemmas 7.5.5 and 7.5.6 to analyse how $\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t)-2\right]$ and f_t change over time.

The following lemmas show that f_t decreases in a similar fashion to its corresponding term in the configuration model. Combining these results with Lemma 7.5.7 then gives the proof of

Lemma 7.5.9. In some cases the proofs are identical to their configuration model counterparts and thus we simply refer to the proofs in that section. However, when the probabilities obtained in Lemma 7.5.6 are needed, the proofs are more involved, as we need to be careful with the error terms.

One significant complication, when compared to the configuration model, comes from analysing the contribution of vertices in $\overline{S}^+$. However, since Lemma 7.5.5 states that $\mathbb{P}_{t-1}\left(w_t \in \overline{S}^+\right)$ is low, and $\mathsf{d}_S(w) = 1$ for all $w \in \overline{S}^+$, this issue is more technical than problematic. Another complication comes from needing to consider forbidden vertices $\mathcal{F}_{t-1}$ at each step, however as discussed in Remark 7.4.6 this set is small and thus also not a major issue.

Define $\lambda := 10^{-6} d(\overline{S}) \frac{n}{M} \delta^{1/4}$ and $\mathcal{I}_x = \{t \in \mathbb{Z} \mid (x-1)\lambda < t \leq x\lambda\}$. These are similar to the definitions of λ and $\mathcal{I}_x$ from Chapter 6, but in this definition λ is smaller by a constant factor since $M = \Theta(n)$. This will allow us to show that X_t stays sufficiently small throughout the process. Note that $\lambda = \Theta(\delta^{1/4} M)$. Recall that $d(w) \leq \delta^{-1/4}$ for all $w \in S_t$ for all $t \geq 0$ from Corollary 7.5.2. Thus, if $|t - s| \leq \lambda$, then $|X'_t - X'_s| < 10^{-6} \frac{d(\overline{S})}{M} n$. Consequently, this implies that $X'_{t-1} < 10^{-6} n$ for all $t \in \mathcal{I}_1$, and thus Lemmas 7.5.5 and 7.5.6 apply for all $t \in \mathcal{I}_1$. The first result we give is the analogue of Lemma 6.2.16.

Lemma 7.5.10. For all $t \in \mathcal{I}_1$, $f_{t+1} - f_t \leq \frac{1}{M_{t-1}}$ and $f_t < -2\delta^{1/4}$.

Proof. The proof is almost identical to the analogous result in the configuration model, which is Lemma 6.2.16. The only thing that needs to be changed is the base case, which follows from Lemma 7.5.8. Much like Lemma 6.2.16, we prove this lemma by induction on t. Lemma 7.5.8 shows that $f_1 \leq -4\delta^{1/4}$. We can express f_{t+1} as

$$\begin{aligned} f_{t+1} &= \sum_{w \in V} \frac{d^{(t)}(w)}{M_t} (\mathsf{d}_S(w) - 2) = \frac{M_{t-1}}{M_t} \left(\sum_{w \in V} \frac{d^{(t-1)}}{M_{t-1}} (\mathsf{d}_S(w_t) - 2) - \frac{\mathsf{d}_S(w_t)}{M_{t-1}} (\mathsf{d}_S(w_t) - 2) \right) \\ &= \left(1 + \frac{\mathsf{d}_S(w_t)}{M_t}\right) \left(f_t - \frac{\mathsf{d}_S(w_t)}{M_{t-1}} (\mathsf{d}_S(w_t) - 2) \right). \end{aligned}$$

Rearranging this gives that

$$f_{t+1} - f_t = -\frac{\mathsf{d}_S(w_t)}{M_{t-1}} (\mathsf{d}_S(w_t) - 2) + \frac{\mathsf{d}_S(w_t)}{M_t} f_t - \frac{\mathsf{d}_S(w_t)^2}{M_t M_{t-1}} (\mathsf{d}_S(w_t) - 2). \tag{7.7}$$

This implies that if $f_t \leq -2\delta^{1/4}$, then $f_{t+1} - f_t \leq \frac{1}{M_{t-1}}$. This concludes the base case. Now suppose that the lemma holds for all $t' < t$. Then Observation 7.5.3 implies that, for all $t \in \mathcal{I}_1$,

$$f_t = f_1 + \sum_{i=1}^{t-1} (f_{i+1} - f_i) < -4\delta^{1/4} + \frac{t-1}{M_{t-1}} < -4\delta^{1/4} + \frac{\lambda}{d(\overline{S})(1 - o(1))} < -2\delta^{1/4}.$$

Therefore, by induction $f_t < -2\delta^{1/4}$ and $f_{t+1} - f_t \leq \frac{1}{M_{t-1}}$ for all $t \in \mathcal{I}_1$. □

Importantly, combining this with Lemma 7.5.7 implies that $\mathbb{E}_{t-1}[\mathsf{d}_S(w_t) - 2] = f_t(1 + o(1))$ for all $t \in \mathcal{I}_1$. Next we give the analogue to Lemma 6.2.13, bounding $\mathbb{E}_{t-1}[f_{t+1} - f_t]$ for all t such that f_t is sufficiently small, as in the configuration model case. The proof is similar, but we need to be careful to manage the errors arising from bounding $\mathbb{P}_{t-1}(w_t = w)$ using the switching method. Since we invoke Lemmas 7.5.5 and 7.5.6, we also assume that $t = o(M)$ and $X_{t-1} < \alpha M$ to ensure that these lemmas apply.

Lemma 7.5.11. Suppose $t = o(M)$, $X_{t-1} < \alpha M$ for some $\alpha < 10^{-3}$, and $f_t \leq -2\delta^{1/4}$. Then there exists a constant $c \in \mathbb{R}^+$ such that either $\mathbb{E}_{t-1}[f_{t+1} - f_t] \leq -\frac{C}{M_{t-1}}$ or $f_t < -C$.

Proof. First we recall the definition of f_t from (7.5) and write $\mathbb{E}_{t-1}[f_{t+1} - f_t]$ as

$$\sum_{w \in V} \mathbb{P}_{t-1}(w_t = w) \left(\sum_{u \in S_{t-1} \setminus \{w\}} \frac{d(u)}{M_{t-1} - \mathsf{d}_S(w)}(d(u) - 1) - \sum_{u \in S_{t-1}} \frac{d(u)}{M_{t-1}}(d(u) - 1) \right).$$

For brevity, we denote the bracketed term of the summand $A(w)$. Note that this expression for $f_{t+1} - f_t$ is equal to the expression given in (7.7), as $M_t = M_{t-1} + \mathsf{d}_S(w_t)$. Note that if $w \notin S_{t-1} \cup \overline{S}_{t-1}$, then $\mathbb{P}_{t-1}(w_t = w) = 0$. Thus,

$$\begin{aligned} \mathbb{E}_{t-1}[f_{t+1} - f_t] = &\sum_{w \in S_{t-1}} \mathbb{P}_{t-1}(w_t = w) A(w) \\ &+ \sum_{w \in \overline{S}_{t-1} \setminus \overline{S}^+} \mathbb{P}_{t-1}(w_t = w) A(w) \\ &+ \sum_{w \in \overline{S}^+} \mathbb{P}_{t-1}(w_t = w) A(w). \end{aligned} \tag{7.8}$$

The first summation on the right hand side can be expressed as

$$\begin{aligned} \sum_{w \in S_{t-1}} \mathbb{P}_{t-1}(w_t = w) A(w) &= \sum_{w \in S_{t-1}} \mathbb{P}_{t-1}\left(w_t = w \middle| w_t \notin \overline{S}^+\right) \mathbb{P}_{t-1}\left(w_t \notin \overline{S}^+\right) A(w) \\ &\quad + \sum_{w \in S_{t-1}} \mathbb{P}_{t-1}\left(w_t = w \middle| w_t \in \overline{S}^+\right) \mathbb{P}_{t-1}\left(w_t \in \overline{S}^+\right) A(w) \\ &= \sum_{w \in S_{t-1}} \mathbb{P}_{t-1}\left(w_t = w \middle| w_t \notin \overline{S}^+\right) \mathbb{P}_{t-1}\left(w_t \notin \overline{S}^+\right) A(w), \end{aligned}$$

since $\mathbb{P}_{t-1}\left(w_t = w \middle| w_t \in \overline{S}^+\right) = 0$ for all $w \in S_{t-1}$. Thus, Lemmas 7.5.5 and 7.5.6 imply that

$$\sum_{w \in S_{t-1}} \mathbb{P}_{t-1}(w_t = w) A(w) = \sum_{w \in S_{t-1}} \frac{d(w)}{M_{t-1}} (1 \pm O(\delta)) A(w).$$

We use the same idea to rewrite the second summation in Equation (7.8). Recall from Remark 7.4.6 the definition of $\mathcal{F}_{t-1}$, the set of vertices in $\overline{S}_{t-1}$ that are adjacent to v_t in the partial pairing T_{t-1}. The second summation can be expressed as

$$\sum_{w \in \overline{S}_{t-1} \setminus \overline{S}^+} \mathbb{P}_{t-1}(w_t = w) A(w) = \sum_{w \in \overline{S}_{t-1} \setminus \left(\overline{S}^+ \cup \mathcal{F}_{t-1}\right)} \frac{d^{(t-1)}(w)}{M_{t-1}} (1 \pm O(\delta)) A(w).$$

Finally, the third summation in Equation (7.8) can be expressed as

$$\sum_{w \in \overline{S}^+} \mathbb{P}_{t-1}(w_t = w) A(w) = \mathbb{P}_{t-1}\left(w_t \in \overline{S}^+\right) \sum_{w \in \overline{S}^+} \mathbb{P}_{t-1}\left(w_t = w \middle| w_t \in \overline{S}^+\right) A(w).$$

Note that if $w_t \notin S_{t-1}$, then

$$A(w) = \sum_{u \in S_{t-1}} \frac{d(u)}{M_{t-1}-1}(d(u)-1) - \sum_{u \in S_{t-1}} \frac{d(u)}{M_{t-1}}(d(u)-1) = \frac{f_t+1}{M_{t-1}-1}.$$

Since $f_t \geq 0$ for all t and we assume that $f_t \leq -2\delta^{1/4}$, it thus follows that $A(w) = (1 - \Omega(\delta^{1/4}))/M_{t-1}$ and $A(w) \geq 0$ for $w \in \overline{S}$. Therefore,

$$\sum_{w \in \overline{S}^+} \mathbb{P}_{t-1}(w_t = w) A(w) = \mathbb{P}\left(w_t \in \overline{S}^+\right) \sum_{w \in \overline{S}^+} \mathbb{P}_{t-1}\left(w_t = w | \, w_t \in \overline{S}^+\right) A(w) = O\left(\frac{\delta}{M_{t-1}}\right).$$

Since $d(\mathcal{F}_{t-1}) \leq \delta^2 M$ (by Corollary 7.5.2 and Remark 7.4.6), this is also implies that

$$0 \leq \sum_{w \in \mathcal{F}_{t-1}} \frac{d^{(t-1)}(w)}{M_{t-1}}(1 \pm O(\delta)) A(w) \leq \frac{2\delta^2 M}{M_{t-1}^2} = O\left(\frac{\delta^2}{M_{t-1}}\right).$$

On the other hand, if $w_t \in S$, then Equation (7.7) implies that

$$A(w) \leq -\frac{d(w)}{M_{t-1}}(d(w)-2),$$

since we assume that $f_t \leq -2\delta^{1/4}$. We apply all of these bounds to Equation (7.8) to obtain

$$\begin{aligned}
\mathbb{E}_{t-1}[f_{t+1} - f_t] &\leq \sum_{w \in S_{t-1}} \frac{d(w)}{M_{t-1}}(1 \pm O(\delta)) \left(-\frac{d(w)}{M_{t-1}}(d(w)-2)\right) \\
&\quad + \sum_{w \in \overline{S}_{t-1} \setminus \overline{S}^+} \frac{d^{(t-1)}(w)}{M_{t-1}}(1 \pm O(\delta)) \frac{f_t+1}{M_{t-1}-1} + O\left(\frac{\delta}{M_{t-1}}\right) \\
&< \sum_{w \in S_{t-1}} \frac{-d(w)^2}{M_{t-1}^2}(d(w)-2)(1 \pm O(\delta)) \\
&\quad + \sum_{w \in \overline{S}_{t-1} \setminus \overline{S}^+} \frac{d^{(t-1)}(w)}{M_{t-1}}(1 + O(\delta)) \frac{1}{M_{t-1}} + O\left(\frac{\delta}{M_{t-1}}\right) \\
&= (1 - O(\delta)) \sum_{w \in S_{t-1}} \frac{-d(w)^2}{M_{t-1}^2}(d(w)-2) + O(\delta) \sum_{w \in S_{t-1}, d(w)=1} \frac{d(w)^2}{M_{t-1}^2} \\
&\quad + \sum_{w \in \overline{S}_{t-1}} \frac{d^{(t-1)}(w)}{M_{t-1}^2} + O(\delta) \sum_{w \in \overline{S}_{t-1}} \frac{d^{(t-1)}(w)}{M_{t-1}^2} + O\left(\frac{\delta}{M_{t-1}}\right) \\
&= (1 - O(\delta)) \left(\sum_{w \in S_{t-1}} \frac{-d(w)^2}{M_{t-1}^2}(d(w)-2) + \frac{d(\overline{S}_{t-1})}{M_{t-1}^2} \right) + O\left(\frac{\delta}{M_{t-1}}\right).
\end{aligned}$$

Now note that the bracketed expression is identical to that which was bounded in Lemma 6.2.18. Thus, we use very similar analysis: if $\sum_{w \in S_{t-1}} d(w)^2(d(w)-2) \geq \frac{5}{4} d(\overline{S}_{t-1})$, we conclude that there exists some constant $C' > 0$ such that

$$\mathbb{E}_{t-1}[f_{t+1} - f_t] \leq \frac{-C'}{M_{t-1}}(1 - O(\delta)) \leq \frac{-C'}{2M_{t-1}}.$$

On the other hand, suppose $\sum_{w\in S_{t-1}} d(w)^2(d(w)-2) < \frac{5}{4}d(\overline{S}_{t-1})$. Lemma 7.5.7 states that $\mathbb{E}_{t-1}[\mathsf{d}_S(w_t)-2] = f_t + O(\delta)(1+f_t)$. Since we assume that $f_t \le -2\delta^{1/4}$, it follows that

$$\mathbb{E}_{t-1}[\mathsf{d}_S(w_t)-2] = f_t(1+o(1)) = \frac{1}{M_{t-1}}\left(\sum_{w\in S_{t-1}} d(w)(d(w)-2) - d(\overline{S}_{t-1})\right)(1+o(1)).$$

Observation 6.2.17 then implies that $\sum_{w\in S_{t-1}} d(w)(d(w)-2) \le \frac{2}{3}\sum_{w\in S_{t-1}} d(w)^2(d(w)-2) \le \frac{5}{6}d(\overline{S}_{t-1})$. Therefore,

$$\begin{aligned}\mathbb{E}_{t-1}[\mathsf{d}_S(w_t)-2] &\le \frac{1}{M_{t-1}}\left(\frac{2}{3}\sum_{w\in S_{t-1}} d(w)^2(d(w)-2) - d(\overline{S}_{t-1})\right)(1+o(1))\\ &\le \frac{1}{M_{t-1}}\left(\frac{5}{6}d(\overline{S}_{t-1}) - d(\overline{S}_{t-1})\right)(1+o(1))\\ &\le -C\end{aligned}$$

for some constant $C>0$, since $d(\overline{S}_{t-1}) = \Theta(M)$ for all $t = o(M)$ by Observation 7.5.3. This completes the proof. □

Now we can give the proof of Lemma 7.5.9. The remaining part of the argument is very similar to the proofs of Lemmas 6.2.13 and 6.2.19. We show that X_t stays small for all necessary values of t under the assumptions of Lemma 7.5.9. Thus, we show that f_t decreases as t increases, much like it did in the configuration model. Then we simply apply Lemma 7.5.7 to prove the desired bound on $\mathbb{E}_{t-1}[\mathsf{d}_S(w_t)-2]$ for all $t \le \delta^{1/19}M$. Recall the definitions of $\lambda := 10^{-6}d(\overline{S})\frac{n}{M}\delta^{1/4}$ and $\mathcal{I}_x = \{t\in\mathbb{Z} \mid (x-1)\lambda < t \le x\lambda\}$.

Proof of Lemma 7.5.9. The first step is showing that an analogous result to Lemma 6.2.19 holds for all the necessary values of t. The proof method is effectively the same, except we also need to check the extra condition that $X_t = o(M)$ for all $t \in \mathcal{I}_x$ for $x \le \delta^{-1/5}$. Let $(T_t)_{t\ge 1}$ be a sequence of partial pairings that satisfies the conditions of Lemma 7.5.9. We prove the following statement for all $x \in \{1,\dots,\delta^{-1/5}\}$ by induction on x: if $(T_t)_{t\ge1}$ is such that $\left|\sum_{t'\le t} A_{t'}\right| \le M^{2/3}$ and $\left|\sum_{t'\le t} G_{t'}\right| \le \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{t}$ for all $t\ge 1$, then $f_{t+1}-f_t \le \frac{1}{M_{t-1}}$, $f_t \le -2\delta^{1/4}$, and $X_t \le \delta^{1/5}$ for all $t\in\mathcal{I}_x$. Since $\delta^{1/19}M = o(\delta^{-1/5}\lambda)$, this implies that the statement holds for all $t \le \delta^{1/19}M$.

We first prove the base case where $x=1$. The definitions of λ and $\mathcal{I}_1$ imply that $X_t < 10^{-6}n$ for all $t\in\mathcal{I}_1$. Lemma 7.5.10 states that $f_t \le -2\delta^{1/4}$ for all $t\in\mathcal{I}_1$. Thus, Lemma 7.5.7 implies that $\mathbb{E}_{t-1}[\mathsf{d}_S(w_t)-2] = f_t(1+o(1))$ for all $t\in\mathcal{I}_1$. Combining this with the assumption on $\sum_{t'\le t} A_{t'}$ implies that, for all $t\in\mathcal{I}_1$,

$$X_t \le X_0' + \sum_{i=1}^{t}\mathbb{E}_{i-1}[\mathsf{d}_S(w_i)-2] + \sum_{i=1}^{t} A_i \le 7\delta^{1/4}M + \sum_{i=1}^{t} f_i(1+o(1)) + M^{2/3} \le 8\delta^{1/4}M. \quad (7.9)$$

This proves the base case. The inductive step follows similarly: suppose that the claims hold for $x' \le x$. Then the bound given in (7.9) that $X_t \le 8\delta^{1/4}M$ for all $t\in\mathcal{I}_{x'}$, for all $x'\le x$. Thus, the definition of λ implies that $X_t < 10^{-6}n$ for all $t\in\mathcal{I}_{x+1}$. Then Lemma 7.5.11 applies for all $t\in\mathcal{I}_{x'}$, for all $x'\le x$. Thus, by identical reasoning to Lemma 6.2.19 (specifically the calculation given in (6.14)), it follows that $f_{\lfloor x\lambda\rfloor+1} \le -4\delta^{1/4}$. Then applying Equation (7.7) gives

that $f_t \le -2\delta^{1/4}$ and $f_{t+1} - f_t \le 1/M_{t-1}$ for all $t \in \mathcal{I}_{x+1}$. Thus, Lemma 7.5.7 implies that $\mathbb{E}_{t-1}[\mathsf{d}_S(w_t) - 2] = f_t(1 + o(1))$ for all $t \in \mathcal{I}_{x+1}$. Therefore, the inequalities given in (7.9) apply for all $t \in \mathcal{I}_{x+1}$. This completes the inductive proof.

Now suppose that $(T_t)_{t\ge1}$ is a sequence of partial pairings such that $\left|\sum_{t'\le t} A_t\right| \le M^{2/3}$ and $\left|\sum_{t'\le t} G_{t'}\right| \le \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{t}$ for all $t \ge 1$. Then the proof of Lemma 6.2.13 carries over exactly under this restriction on t. That is, if there exists some $s \le t$ such that $f_s < -C$ for some constant $C > 0$, then for all $t \le \delta^{1/19}M$ it follows that

$$f_t = f_s + \sum_{i=s}^{t-1}(f_{i+1} - f_i) < -C + \frac{t}{M_{t-1}} = -C(1 + o(1)).$$

Otherwise, $\mathbb{E}_{s-1}[f_{s+1} - f_s] \le -\frac{C}{M_{s-1}}$ for all $s \le t$, and the assumed bound on $\left|\sum_{t'\le t} G_{t'}\right|$ implies that $f_t \le \frac{-Ct}{M_{t-1}}$.

This implies that for all $t \le \delta^{1/19}M$, we know that $f_t \le -Ct/M_{t-1}$ for some constant $C > 0$, and also that $f_t \le -2\delta^{1/4}$ for all $t \le \delta^{1/19}M$. This means that for all such t, Lemma 7.5.7 implies that $\mathbb{E}_{t-1}[\mathsf{d}_S(w_t) - 2] = f_t(1 + o(1))$. Therefore, for all $t \le \delta^{1/19}M$, there exists some constant c, where $c \le C$, such that $\mathbb{E}_{t-1}[\mathsf{d}_S(w_t) - 2] \le -ct/M_{t-1}$. This completes the proof. □

Now we show that $\sum_{t'\le t} A_{t'}$ and $\sum_{t'\le t} G_{t'}$ are sufficiently concentrated to apply Lemma 7.5.9. The proof of this is analogous to Lemmas 6.2.14 and 6.2.15, and the results carry over almost immediately. Both concentration results are stated again below for convenience. Recall $A_t := \mathsf{d}_S(w_t) - \mathbb{E}_{t-1}[\mathsf{d}_S(w_t)]$ and recall $G_t := f_{t+1} - f_t - \mathbb{E}_{t-1}[f_{t+1} - f_t]$. We give a brief sketch of why the respective configuration model result translates to this model, and refer to the original proofs for the full details.

Lemma 7.5.12. The probability that $\left|\sum_{t'\le t} A_{t'}\right| > M^{2/3}$ for some $t \le M$ is at most $2e^{-M^{1/4}}$. The probability that $\left|\sum_{t'\le t} G_{t'}\right| > \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{t}$ for some $t \le M$ is $o(M^{-2})$.

Proof sketch. Corollary 7.5.2(b) states that S_t has maximum degree $\delta^{-1/4}$ for all $t \ge 0$. Thus, $|A_t| \le \delta^{-1/4}$ for all $t \ge 1$. Thus, an identical proof to that given in Lemma 6.2.15 shows that $\left(\sum_{t'\le t} A_t\right)_{t\ge1}$ is a martingale and satisfies the desired concentration inequality. To bound $|G_t|$, note that $f_{t+1} - f_t$ can be expressed as in Equation (7.7), and thus

$$\begin{aligned}
|f_{t+1} - f_t| &= \left|-\frac{\mathsf{d}_S(w_t)}{M_{t-1}}(\mathsf{d}_S(w_t) - 2) + \frac{\mathsf{d}_S(w_t)}{M_t}f_t - \frac{\mathsf{d}_S(w_t)^2}{M_tM_{t-1}}(\mathsf{d}_S(w_t) - 2)\right| \\
&\le \frac{\delta^{-1/4}(\delta^{-1/4} - 2)}{M_{t-1}} + \frac{\mathsf{d}_S(w_t)}{M_t}\left|f_t - \frac{\mathsf{d}_S(w_t)}{M_{t-1}}(\mathsf{d}_S(w_t) - 2)\right| \\
&\le \frac{2\delta^{-1/2}}{M_t}.
\end{aligned}$$

Therefore, the proof of Lemma 6.2.14 carries over exactly. □

We now apply Lemmas 7.5.9 and 7.5.12 to show that in the deferred decision process, starting from $V_0 = U_1 \cup W' \cup \{v\}$ for some arbitrary $v \in S$, with probability $1 - o(M^{-2})$ there exists some time $t \le \delta^{1/19}M$ such that $X_t = 0$. By design, the proof of this is almost identical to Lemma 6.2.20.

Lemma 7.5.13. With probability $1 - o(M^{-2})$, there exists some time $t \leq \delta^{1/18}M$ such that $X_t = 0$ in the deferred decision process.

Proof. Suppose that $(T_t)_{t\geq 0}$ is a sequence of partial pairings such that, for all $t \leq \delta^{1/18}M$,

$$\left|\sum_{t'\leq t} A_{t'}\right| \leq M^{2/3} \text{ and } \left|\sum_{t'\leq t} G_{t'}\right| \leq \frac{4\sqrt{\log M}}{M\delta^{1/2}}\sqrt{t}. \tag{7.10}$$

Lemma 7.5.12 states that this occurs with probability $1 - o(M^{-2})$. Thus, it suffices to show that for each such sequence of partial pairings, there exists some $t \leq \delta^{1/18}M$ such that $X_t = 0$. Since we assume $(T_t)_{t\geq 1}$ satisfies the inequalities given in (7.10), it follows that Lemma 7.5.9 applies, and thus there exists some constant $c > 0$ such that $\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t) - 2\right] \leq -ct/M_{t-1}$ for all $t \leq \delta^{1/19}M$. Consider the value of X'_T at time $T := \delta^{1/18}M$:

$$\begin{aligned}
X'_T &\leq 7\delta^{1/4}M + \sum_{t=1}^{T}(\mathsf{d}_S(w_t) - 2) \\
&\leq 7\delta^{1/4}M + M^{2/3} - \sum_{t=1}^{T}\frac{ct}{M_{t-1}} \\
&\leq 7\delta^{1/4}M + M^{2/3} + \frac{cT(T-1)}{2M} \\
&\leq 7\delta^{1/4}M + M^{2/3} - \delta^{1/8}M \\
&< 0.
\end{aligned}$$

Therefore, since $X_t \leq X'_t$ for all t, there exists some $t \leq T$ such that $X_t = 0$. Since this holds for all sequences of partial pairings such that the inequalities in (7.10) holds, and those inequalities hold for all $t \geq 0$ with probability $1 - o(M^{-2})$, this completes the proof. □

As in the configuration model, we use Lemma 7.5.13 to show that every component in $G^*[S]$ a.a.s. contains $o(|S|)$ edges and vertices, which completes the proof of Theorem 7.1.1. This proof of the a.a.s. non-existence of an induced giant component in the subcritical case is almost identical to the proof in the configuration model. The only difference is in the exact definition of V_0, but this difference is purely nominal. As mentioned at the start of this section, Lemma 6.3.1 implies that $M = \Theta(|S|)$, since $R_S = o(M)$ and $d(S) = \Theta(M)$.

Proof of Theorem 7.1.1. Lemma 6.3.1 states that if $R_S = o(M)$ and $d(S) = \Theta(M)$ then $M = \Theta(|S|)$. Suppose that $V_0 = U_1 \cup W' \cup \{v\}$ for an arbitrary $v \in S$. By Lemma 7.5.13, with probability $1 - o(M^{-2})$ there is some time $t \leq \delta^{1/18}M$ such that $X'_t \leq 0$, and thus $X_t = 0$. This implies that, with probability $1 - o(M^{-2})$, the total degree of all vertices in S in every component explored by the process starting at v at this time t is at most $d(U_1) + \sum_{i=1}^{t}\mathsf{d}_S(w_i) < 6\delta^{1/4}M + 2t < 3\delta^{1/18}M$ (where the bound on $d(U_1)$ is from Corollary 7.5.2(a)).

Taking the union bound over all (at most n) choices for the vertex v, a.a.s. every component of $G^*[S]$ has at most $|V_0 \cap S| + \delta^{1/18}M$ vertices and at most $3\delta^{1/18}M$ edges. Since $|V_0 \cap S| = |U_1 \cup \{v\}| \leq 6\delta^{1/4}M + 1$ and $M = \Theta(|S|)$, this implies that the largest component in $G^*[S]$ a.a.s. has less than $2\delta^{1/18}M = o(|S|)$ vertices and $o(M)$ edges. Recall that $G^* \in \Phi$ is a simple pairing corresponding to the graph G. Since the largest component in $G^*[S]$ has the same distribution as the largest component in $G[S]$ (where $G \sim \mathcal{G}(\boldsymbol{d})$), this completes the proof. □

7.6 Analysis in the supercritical regime

In this section we prove Theorem 7.1.2(a) in the case that $M = \Theta(n)$. In combination with Lemma 7.1.4, this completes the proof of Theorem 7.1.2(a). Recall that

$$H^+ = \left\{v \in [n] : d(v) > \delta\sqrt{M}\right\} \tag{7.11}$$

for some $\delta \to 0$ satisfying conditions (a) – (c) of Remark 7.3.1, and recall that $S^+ = H^+ \cap S$ and $\overline{S}^+ = H^+ \cap \overline{S}$. Much like in the configuration model proof, we split the proof of the supercritical case into two parts: one for pairs $(\boldsymbol{d}, S)$ such that $d(S^+) = o(M)$ and one where $d(S^+) = \Theta(M)$. The first case is analysed via the exploration process in much the same way as the configuration model proof, using Lemmas 7.4.1 and 7.4.2. Similarly to the configuration model proof, the case where $d(S^+) = \Theta(M)$ is not shown via the exploration process. However, the proof we use here is quite different to the configuration model result, due to the complexities of $\mathcal{G}(\boldsymbol{d})$.

7.6.1 The case where $d(S^+) = o(M)$

In this section we again analyse the exploration process. The idea in this section is the same as in the previous section: proving that the growth of X_t throughout the deferred decision process is similar to its growth in the configuration model. Much like in the configuration model, the proof of the supercritical case is more straightforward than the proof of the subcritical case that we gave in the previous section. The first step is to define a preprocessing step such that Lemmas 7.4.1 and 7.4.2 apply to the process.

Preprocessing and setup

Define $\varepsilon > 0$ to be a constant such that $R_S \geq \varepsilon M$. This implies that there exists $\varepsilon' := \varepsilon\frac{M}{n}$ such that $R_S \geq \varepsilon' n$ and $\varepsilon' = \Theta(1)$. Note that since $n \leq M$, we know that $\varepsilon \leq \varepsilon'$. For convenience, we choose ε such that $\varepsilon' \leq 1$. Define τ similarly to the corresponding variable in the configuration model: τ is the smallest time t such that either $X_\tau > \beta M$, where $\beta := 10^{-6}\varepsilon^2$, or $M_t \leq \left(1 - \frac{\varepsilon}{4}\right) M$. Since $M_{t-1} - M_t \geq 1$ for all t, this immediately implies that $\tau \leq \frac{\varepsilon}{4}M = \frac{\varepsilon'}{4}n$.

If S^+ is non-empty and $d(S^+) = o(M)$, then in the preprocessing step we delete the set S^+, moving it into $\overline{S}$. If $S^+ = \emptyset$, then this preprocessing step does nothing. We show that $G[S\backslash S^+]$ a.a.s. contains an edge-giant component, and thus a.a.s. $G[S]$ also contains an edge-giant component. The important thing to check in this case is that moving the vertices from S^+ into $\overline{S}$ does not significantly affect the exploration process. Since $d(S^+)$ is small, this is indeed the case. This is similar to the idea used in the configuration model proof, but again we do not explode the deleted vertices into degree 1 vertices. For the preprocessing set, define $V_0 = \{v\}$ for an arbitrary vertex $v \in S\backslash S^+$, and $E_0 = \emptyset$.

Thus, we can think of this preprocessing step as setting $V_0 = \{v\}$, $S_0 = S\backslash(\{v\} \cup S^+)$, and $\overline{S}_0 = \overline{S} \cup S^+$. This implies that if $w \in S^+$, then $\mathtt{d}_S(w) = 1$ for the purposes of our analysis of the exploration process. In this sense, we do not consider the vertices in S^+ to actually be in S, implicitly redefining $S := S\backslash S^+$. As a result of this, when discussing the exploration process for $t \geq 0$ we often refer to H^+ simply as $\overline{S}^+$, using $\overline{S}^+$ to refer to every vertex with degree greater than $\delta\sqrt{M}$ in G. We still use S^+ to refer to the set of vertices in G with degree greater than $\delta\sqrt{M}$

that were originally in S, as there are times that we need to refer to this set specifically. Since $d(H^+) = o(M)$ and $M = \Theta(n)$, we can again define δ such that $d(H^+) \leq \delta^2 n$.

Recall $d^{(t)}(w)$, the number of "available" half edges in vertex w at time t, from Equation (7.3). We show, analogously to the subcritical case, that T_{t-1} is predictable for all $t \leq \tau$, meaning that it satisfies (P1) to (P5). These all follow from the definition of τ and the fact that S_t has maximum degree $\delta\sqrt{M}$ for all $t \geq 0$ by preprocessing.

Proposition 7.6.1. Suppose that $R_S \geq \varepsilon M$ for some constant $\varepsilon > 0$ such that $\varepsilon \leq \varepsilon' \leq 1$. Suppose that $M = \Theta(n)$, $d(H^+) \leq \delta^2 n$, and $t \leq \tau$. Then T_{t-1} is predictable.

Proof. The definition of τ implies that condition (P1) is satisfied. The preprocessing step implies that S_t has maximum degree $\delta\sqrt{M}$ for all $t \geq 0$. Since $v_t \in S$ for all $t > 0$, conditions (P2) and (P3) are satisfied. Recall that we chose $\varepsilon > 0$ such that $\varepsilon \leq \varepsilon' \leq 1$. The definition of τ implies that $X_{t-1} \leq 10^{-6}\varepsilon^2 M = 10^{-6}\varepsilon\varepsilon' n < 10^{-3}n$, and thus condition (P4) is satisfied. Finally, we know that $|V_{t-1}| \leq |V_0| + t - 1 = t$, and $t \leq \tau \leq \frac{\varepsilon'}{4}n \leq \frac{1}{4}n$, which means that condition (P5) is satisfied. □

Thus, the following two lemmas are an immediate consequence of this proposition and Lemmas 7.4.1 and 7.4.2 respectively.

Lemma 7.6.2. Suppose that $d(H^+) \leq \delta^2 n$ and $R_S \geq \varepsilon M$, where $M = \Theta(n)$ and $\varepsilon \leq \varepsilon' \leq 1$. For all $t \leq \tau$, $\mathbb{P}_{t-1}\left(w_t \in \overline{S}^+\right) \leq \delta$.

Lemma 7.6.3. Suppose that $d(H^+) \leq \delta^2 n$ and $R_S \geq \varepsilon M$, where $M = \Theta(n)$ and $\varepsilon \leq \varepsilon' \leq 1$. Furthermore, suppose that $t \leq \tau$. Let $w \notin \overline{S}^+$ be a vertex such that $d^{(t-1)}(w) \geq 1$ and $v_t w \notin E_{t-1}$. Then

$$\mathbb{P}_{t-1}\left(w_t = w \middle| w_t \notin \overline{S}^+\right) = \frac{d^{(t-1)}(w)}{M_{t-1}}\left(1 \pm O(\delta)\right).$$

Process evolution in the supercritical regime

Now we use Lemmas 7.6.2 and 7.6.3 to prove the case of Theorem 7.1.2 when the total degree of $d(H^+) = o(M)$ and $M = \Theta(n)$. Again we adapt the proof of the configuration model result. Since $R_S \geq \varepsilon M$, this implies that $\sum_{i \leq j_S} d(i)(\mathsf{d}_S(i) - 2) > 0$. From this we also know that $\mathsf{d}_S(j_S) = d(j_S) \geq 3$, otherwise this summation could not be positive.

In the following lemma we show that if S^+ has small total degree, then moving these vertices from S to $\overline{S}$ does not significantly alter f_t (recall that f_t, given in Equation (7.5), is intuitively the expectation of $\mathsf{d}_S(w_t) - 2$ if $\mathbb{P}_{t-1}(w_t = w) = \frac{d^{(t-1)}(w)}{M_{t-1}}$ for all $w \in V$). Furthermore, it also shows that as long as a sufficiently small (but positive) fraction of the graph has been explored, the value of f_t is guaranteed to be bounded away from 0. This proof is very similar to Lemma 6.2.4, but included for completeness as some of the functions considered are slightly different to the configuration model analogue (such as $\mathsf{d}_S(\cdot)$ rather than $d(\cdot)$).

Lemma 7.6.4. Suppose $d(S^+) \leq \frac{\varepsilon M}{100}$ and $R_S \geq \varepsilon M$ for fixed $\varepsilon > 0$. Suppose $U \subset S$ is a set containing all vertices in S with degree greater than $\delta\sqrt{M}$, and suppose $d(U) < \frac{\varepsilon}{4}M + d(S^+)$. Then

$$\sum_{i \in V \setminus U} d(i)\left(\mathsf{d}_S(i) - 2\right) \geq \frac{2}{3}\varepsilon M.$$

Proof. We first prove by contradiction that $d(j_S) \leq \delta\sqrt{M}$ under the assumptions of this lemma. If $d(j_S) > \delta\sqrt{M}$, then vertex j_S and all vertices of greater degree are in U, and thus $d(U) \geq R_S$ which violates the assumptions on U. We write

$$\sum_{i \in V \setminus U} d(i)\left(\mathsf{d}_S(i) - 2\right) = \sum_{i \leq j_S} d(i)(\mathsf{d}_S(i) - 2) - \sum_{i \in U, i \leq j_S} d(i)\left(\mathsf{d}_S(i) - 2\right) + \sum_{i \notin U,\ i > j_S} d(i)\left(\mathsf{d}_S(i) - 2\right).$$

To bound the first sum on the right hand side, first note that $\sum_{i \leq j_S} d(i)(\mathsf{d}_S(i) - 2) > 0$. Then for the latter two sums, note that $\mathsf{d}_S(i) = d(i) \geq 3$ for all $i \geq j_S$. This implies that $d(j_S) \in [3, \delta\sqrt{M}]$. Then the expression given above is bounded as follows.

$$\begin{aligned}
\sum_{i \in V \setminus U} d(i)\left(\mathsf{d}_S(i) - 2\right) &\geq -\sum_{i \in U, i \leq j_S} d(i)\left(\mathsf{d}_S(i) - 2\right) + \left(d(j_S) - 2\right) \sum_{i \notin U, i > j_S} d(i) \\
&\geq -\left(d(j_S) - 2\right) \sum_{i \in U, i \leq j_S} d(i) + \left(d(j_S) - 2\right) \left(\sum_{i > j_S} d(i) - \sum_{i \in U, i > j_S} d(i) \right) \\
&\geq \left(d(j_S) - 2\right)\left(R_S - d(j_S) - d(U)\right) \\
&\geq \left(d(j_S) - 2\right)\left(R_S - \delta\sqrt{M} - \frac{\varepsilon}{4}M - d(S^+)\right) \\
&\geq \frac{2}{3}\left(d(j_S) - 2\right) R_S.
\end{aligned}$$

Since $d(j_S) \geq 3$ and $R_S \geq \varepsilon M$, this proves the lemma. □

We now use this lemma to show that $\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t) - 2\right]$ is bounded away from 0 for all $t \leq \tau$.

Lemma 7.6.5. Suppose that $M = \Theta(n)$, $R_S \geq \varepsilon M$, and $d(S^+) \leq \frac{\varepsilon M}{100}$. For all $t \leq \tau$, $\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t) - 2\right] \geq \frac{1}{4}\varepsilon$.

Proof. We first give a quick outline of the proof. We use Lemma 7.6.4 to show that $f_t \geq \frac{1}{3}\varepsilon$. We then show that, since the switching errors given in Lemma 7.6.3 are small, the value of $\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t) - 2\right]$ is not too far from f_t. Firstly, f_t is bounded below as follows:

$$\begin{aligned}
\sum_{w \in V} \frac{d^{(t-1)}(w)}{M_{t-1}}\left(\mathsf{d}_S(w) - 2\right) &= \sum_{d^{(t-1)}(w) \geq 1, \mathsf{d}_S(w) \geq 2} \frac{d(w)}{M_{t-1}}\left(\mathsf{d}_S(w) - 2\right) - \sum_{d^{(t-1)}(w) \geq 1, \mathsf{d}_S(w) = 1} \frac{d^{(t-1)}(w)}{M_{t-1}} \\
&\geq \sum_{d^{(t-1)}(w) \geq 1, \mathsf{d}_S(w) \geq 2} \frac{d(w)}{M}\left(\mathsf{d}_S(w) - 2\right) - \sum_{d^{(t-1)}(w) \geq 1, \mathsf{d}_S(w) = 1} \frac{d^{(t-1)}(w)}{\left(1 - \frac{\varepsilon}{4}\right) M} \\
&= \sum_{w \in V} \frac{d^{(t-1)}(w)}{M}\left(\mathsf{d}_S(w) - 2\right) - \left(\frac{1}{1 - \frac{\varepsilon}{4}} - 1\right) \sum_{d^{(t-1)}(w) \geq 1, \mathsf{d}_S(w) = 1} \frac{d(w)}{M} \\
&= \sum_{w \in V} \frac{d^{(t-1)}(w)}{M}\left(\mathsf{d}_S(w) - 2\right) - \left(\frac{\varepsilon}{4} + O\left(\varepsilon^2\right)\right) \sum_{d^{(t-1)}(w) \geq 1, \mathsf{d}_S(w) = 1} \frac{d(w)}{M} \\
&\geq \sum_{w \in V} \frac{d^{(t-1)}(w)}{M}\left(\mathsf{d}_S(w) - 2\right) - \frac{\varepsilon}{4}.
\end{aligned}$$

For all $w \in S$, $d^{(t-1)}(w)$ is equal to either 0 if $w \in V_{t-1}$ or $d(w)$ otherwise. For all $w \in \overline{S}$, we use

the simple bound the $d^{(t-1)}(w) \leq d(w)$. It then follows that

$$\sum_{w \in V} \frac{d^{(t-1)}(w)}{M} (\mathsf{d}_S(w) - 2) \geq \sum_{w \notin S \cap V_{t-1}} \frac{d(w)}{M} (\mathsf{d}_S(w) - 2).$$

Let $U = (S \cap V_{t-1}) \cup S^+$. Note that since $t \leq \tau$, it follows that $d(U) < \frac{\varepsilon}{4} M + d(S^+)$. Then Lemma 7.6.4 applies, and it follows that $f_t \geq \left(\frac{2}{3} - \frac{1}{4}\right) \varepsilon \geq \frac{1}{3}\varepsilon$.

Now relating this to $\mathbb{E}_{t-1}[\mathsf{d}_S(w_t) - 2]$ via Lemma 7.6.3, it follows that

$$\begin{aligned}
\mathbb{E}_{t-1}[\mathsf{d}_S(w_t) - 2] &= \sum_{w \in V} \mathbb{P}_{t-1}(w_t = w)(\mathsf{d}_S(w) - 2) \\
&= \sum_{w \in V} \mathbb{P}_{t-1}\left(w_t = w \middle| w_t \notin \overline{S}^+\right)(\mathsf{d}_S(w) - 2)\mathbb{P}_{t-1}\left(w_t \notin \overline{S}^+\right) \\
&\quad + \sum_{w \in V} \mathbb{P}_{t-1}\left(w_t = w \middle| w_t \in \overline{S}^+\right)(\mathsf{d}_S(w) - 2)\mathbb{P}_{t-1}\left(w_t \in \overline{S}^+\right).
\end{aligned}$$

Note that $w_t \in S_{t-1}$, $w_t \in \overline{S}_{t-1} \backslash \overline{S}^+$, and $w_t \in \overline{S}^+$ are all mutually exclusive events. Thus,

$$\begin{aligned}
\mathbb{E}_{t-1}[\mathsf{d}_S(w_t) - 2] &= \sum_{w \in S_{t-1}} \mathbb{P}_{t-1}\left(w_t = w \middle| w_t \notin \overline{S}^+\right)(d(w) - 2)\mathbb{P}_{t-1}\left(w_t \notin \overline{S}^+\right) \\
&\quad + \sum_{w \in \overline{S}_{t-1} \backslash \overline{S}^+} \mathbb{P}_{t-1}\left(w_t = w \middle| w_t \notin \overline{S}^+\right)(-1)\mathbb{P}_{t-1}\left(w_t \notin \overline{S}^+\right) \\
&\quad + \sum_{w \in \overline{S}^+} \mathbb{P}_{t-1}\left(w_t = w \middle| w_t \in \overline{S}^+\right)(-1)\mathbb{P}_{t-1}\left(w_t \in \overline{S}^+\right).
\end{aligned} \tag{7.12}$$

We analyse each sum in turn. The last summation is

$$\begin{aligned}
&\sum_{w \in \overline{S}^+} \mathbb{P}_{t-1}\left(w_t = w \middle| w_t \in \overline{S}^+\right)(-1)\mathbb{P}_{t-1}\left(w_t \in \overline{S}^+\right) \\
&\quad = -\mathbb{P}_{t-1}\left(w_t \in \overline{S}^+\right) \sum_{w \in \overline{S}^+} \mathbb{P}_{t-1}\left(w_t = w \middle| w_t \in \overline{S}^+\right) \geq -\delta
\end{aligned} \tag{7.13}$$

by Lemma 7.6.2 and the law of total probability. We now apply Lemma 7.6.3 to bound the remaining two sums from below. Recall from Remark 7.4.6 the definition of $\mathcal{F}_{t-1}$, the set of vertices $w \in \overline{S}_{t-1} \backslash \overline{S}^+$ that are adjacent to v_t in the partial pairing T_{t-1}; this is the set of vertices such that $\mathbb{P}_{t-1}(w_t = w) = 0$ despite the fact that $d^{(t-1)}(w) \geq 1$. Since $d(v_t) \leq \delta\sqrt{M}$ for all $t \geq 0$, it follows that $d(\mathcal{F}_{t-1}) \leq \delta^2 M$. Thus, we can bound the second sum by

$$\begin{aligned}
-\mathbb{P}_{t-1}\left(w_t \notin \overline{S}^+\right) \sum_{w \in \overline{S}_{t-1} \backslash \overline{S}^+} \mathbb{P}_{t-1}\left(w_t = w \middle| w_t \notin \overline{S}^+\right) &\geq -\sum_{w \in \overline{S}_{t-1} \backslash \overline{S}^+} \mathbb{P}_{t-1}\left(w_t = w \middle| w_t \notin \overline{S}^+\right) \\
&\geq -\sum_{w \in \overline{S}_{t-1}} \frac{d^{(t-1)}(w)}{M_{t-1}}(1 + O(\delta)) \\
&= -\frac{d(\overline{S}_{t-1})}{M_{t-1}}(1 + O(\delta)).
\end{aligned} \tag{7.14}$$

To apply Lemma 7.6.3 to obtain a lower bound on the first sum, we first need to split the sum based on whether the value of the summand is negative or not. Doing so, we bound the first sum

from below by

$$\begin{aligned}
&(1-O(\delta))\left(\sum_{w\in S_{t-1},d(w)\geq 2}\frac{d(w)}{M_{t-1}}(d(w)-2)(1-O(\delta))-\sum_{w\in S_{t-1},d(w)=1}\frac{d(w)}{M_{t-1}}(1+O(\delta))\right)\\
=&(1-O(\delta))\sum_{w\in S_{t-1}}\frac{d(w)}{M_{t-1}}(d(w)-2)-O(\delta)\sum_{w\in S_{t-1},d(w)=1}\frac{d(w)}{M_{t-1}}\\
=&(1-O(\delta))\sum_{w\in S_{t-1}}\frac{d(w)}{M_{t-1}}(d(w)-2)-O(\delta), \qquad (7.15)
\end{aligned}$$

since $n_1/M_{t-1}=O(1)$. Therefore, combining the bounds given in (7.12) to (7.15) gives that

$$\begin{aligned}
\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t)-2\right] &\geq (1-O(\delta))\sum_{w\in S_{t-1}}\frac{d(w)}{M_{t-1}}(d(w)-2)-\frac{d(\overline{S}_{t-1})}{M_{t-1}}(1+O(\delta))-O(\delta)\\
&=(1-O(\delta))\sum_{w\in V}\frac{d(w)}{M_{t-1}}(\mathsf{d}_S(w)-2)-O(\delta)\\
&=(1-O(\delta))f_t-O(\delta).
\end{aligned}$$

Thus, it follows that $\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t)-2\right]\geq(1-O(\delta))\frac{1}{3}\varepsilon-O(\delta)\geq\frac{1}{4}\varepsilon$. Since Lemma 7.6.3 holds for all $t\leq\tau$, this completes the proof. □

Recall that $X_t=|\mathcal{X}_t|$, the number of open edges in the partial pairing at time t. The next ingredient we need to obtain a lower bound on X_t is a lower bound on the number of back edges that are revealed throughout the process. The definition of the exploration process means that we only reveal back edges between $V_{t-1}\cap S$ and $w_t\in S$. Intuitively, $\mathcal{G}(\boldsymbol{d})$ is locally tree-like in the sense that we expect very few short cycles. Thus, we expect that the number of back edges revealed at each step should usually be quite low. This indeed turns out to be the case, as we prove via the switching method in Lemma 7.6.6. This lemma implies that, on average, back edges do not consume all or almost all of the open edges gained at each step, and thus do not affect the average size of X_t beyond a constant factor. Recall that $d'_{S,t}(w_t)$ is the number of back edges revealed by the process at step t.

Lemma 7.6.6. Let $t\leq\tau$. Then conditional on the process up to time $t-1$,

$$\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right]\leq\frac{1}{12}\mathbb{E}_{t-1}\left[\mathsf{d}_S(w_t)-2\right].$$

Proof. Let $T=(V_{t-1},E_{t-1},\mathcal{X}_{t-1})$ be the partial pairing at time $t-1$. Let $\Phi_w(T)$ be the subset of simple pairings $G^*\in\Phi$ such that G^* agrees with T and $w_t=w$, for some vertex w. Note that if $w\in\overline{S}$ or $d(w)=1$, then $d'_{S,t}(w)=0$. Thus, we restrict our attention to the case where $w\in S_{t-1}$ and $d(w)\geq 2$. Let A_i be the subset of $\Phi_w(T)$ such that there are exactly i edges between w and $V_{t-1}\cap S\backslash\{v_t\}$, that is, there are exactly i back edges revealed at step t of the process. We define a switching between A_{i+1} and A_i, for all $i\geq 0$. Let $G^*\in A_{i+1}$, and let wz be a back edge, that is, an edge such that $z\in S\cap V_{t-1}$ and $wz\in E(G^*)$. Let (x,y) be an ordered pair of adjacent vertices such that $xy\notin E_{t-1}$. Then a switching deletes the edges wz and xy and adds in edges wy and xz, creating a new G'^*.

For each $G^*\in A_{i+1}$, by definition there are exactly $i+1$ choices for a back edge wz. There are

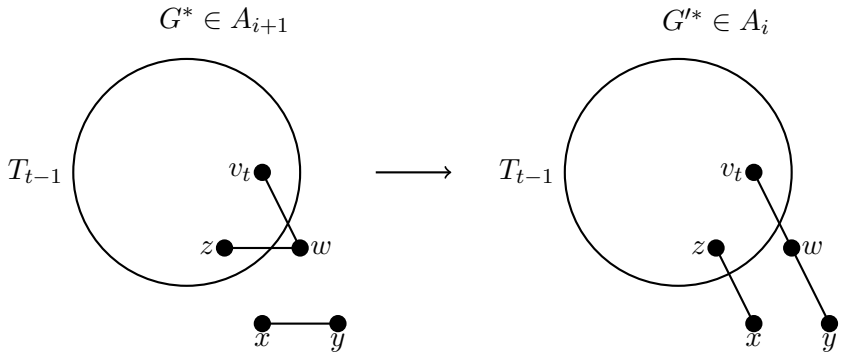

Figure 7.8: A diagram of the switching used to prove Lemma 7.6.6, which sends $G^* \in A_{i+1}$ to $G'^* \in A_i$.

M_{t-1} choices for an ordered pair of adjacent vertices (x, y) such that $xy \notin E_{t-1}$ and $d^{(t-1)}(y) \geq 1$. This implies that $y \notin V_{t-1} \cap S$, and thus wy is not a back edge in G'^*. This choice for the edge xy corresponds to a valid switching if and only if $G'^* \in A_i$, which occurs if

(a) $x, y \notin \overline{S}^+$,

(b) $x \notin V_{t-1} \cap S$,

(c) $xz, wy \notin E(G^*)$,

(d) the vertices $\{v_t, w, x, y, z\}$ are all distinct.

Of the M_{t-1} choices for (x, y) mentioned previously, at most $2\delta^2 n$ of these have either $x \in \overline{S}^+$ or $y \in \overline{S}^+$. We now bound from below the number of remaining choices for (x, y) that do not satisfy conditions (b) – (d). Suppose (b) is not satisfied. Then the edge xy corresponds to an open edge and its mate. Thus, there are at most X_{t-1} choices for (x, y) such that $x \in V_{t-1} \cap S$. By the standard counting arguments, since $\{v_t, w, x, y, z\}$ all have degree at most $\delta\sqrt{M}$ in G (as z must be in $V_{t-1} \cap S$ by definition and we have already counted the case where $x, y \in \overline{S}^+$), it follows that, for each choice of z, there are at most $2\delta^2 M$ choices for (x, y) such that either $xz \in E(G^*)$ or $wy \in E(G^*)$. Finally, there are $O(\delta\sqrt{M})$ choices for $\{v_t, w, x, y, z\}$ such that the vertices are not all distinct from each other. Therefore, for each $G^* \in A_{i+1}$, there are at least $(i+1)(M_{t-1} - X_{t-1} - O(\delta^2 M))$ choices for the vertices x, y, and z that correspond to a valid forward switching.

The number of reverse switchings that can be applied to each $G'^* \in A_i$ is at most the number of ways of choosing a neighbour y of w such that $y \notin V_{t-1} \cap S$ and choosing an edge xz where $z \in V_{t-1} \cap S$ and $x \notin V_{t-1} \cap S$. There are at most $d(w) - 1$ choices for y, as $v_t w \in E(G^*)$ by assumption. There are at most X_{t-1} choices for xz, since this edge corresponds to an open edge. For all $t \leq \tau$, we know that $X_{t-1} \leq \beta M$, where $\beta := 10^{-6}\varepsilon^2$. Thus, for all $i \geq 0$,

$$\frac{|A_{i+1}|}{|A_i|} \leq \frac{(d(w)-1)X_{t-1}}{(i+1)(M_{t-1} - X_{t-1})}(1 + o(1)) \leq \frac{(d(w)-1)\beta}{(i+1)(1 - 10^{-6}\varepsilon^2)\left(1 - \frac{\varepsilon}{4}\right)},$$

since $M_{t-1} \geq \left(1 - \frac{\varepsilon}{4}\right) M$ for all $t \leq \tau$. Now suppose that $i + 1 \geq \sqrt{\beta}(d(w) - 1)$. Then this ratio is at most $2\sqrt{\beta}$. Using this, we obtain the following bound on $\mathbb{E}_{t-1}\left[d'_{S,t}(w_t) \middle| w_t = w\right]$ for

all $w \in S_{t-1}$ with degree at least 2:

$$\begin{aligned}
\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\middle| w_t = w\right] &= \sum_{k\geq 0} k\mathbb{P}_{t-1}\left(d'_{S,t}(w_t) = k\middle| w_t = w\right) \\
&\leq \sqrt{\beta}(d(w_t)-1) + \sum_{k>\sqrt{\beta}(d(w_t)-1)} k\mathbb{P}_{t-1}\left(d'_{S,t}(w_t) = k\middle| w_t = w\right) \\
&\leq \sqrt{\beta}(d(w_t)-1) + \sum_{k>\sqrt{\beta}(d(w_t)-1)} \left(k + \sqrt{\beta}(d(w_t)-1)\right)(2\sqrt{\beta})^k \\
&= \sqrt{\beta}(d(w_t)-1)\left(1 + \sum_{i\geq 1}(2\sqrt{\beta})^i\right) + \sum_{i\geq 1} i(2\sqrt{\beta})^i \\
&< \sqrt{\beta}(d(w_t)-1)\left(1 + 2.5\sqrt{\beta}\right) + \frac{2\sqrt{\beta}}{(1-2\sqrt{\beta})^2} \\
&\leq \sqrt{\beta}(d(w_t)-1)\left(1 + 2.5\sqrt{\beta}\right) + \frac{2\sqrt{\beta}(d(w_t)-1)}{(1-2\sqrt{\beta})^2},
\end{aligned}$$

since we are conditioning on $w_t = w$ for some $w \in S_{t-1}$ such that $d(w) \geq 2$. Thus, for each such vertex w,

$$\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\middle| w_t = w\right] \leq 2\sqrt{\beta}(d(w_t)-1)\left(2 + 5\sqrt{\beta}\right) \leq 6\sqrt{\beta}(d(w_t)-1).$$

The law of total expectation gives that

$$\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right] = \sum_{w\in V} \mathbb{P}_{t-1}\left(w_t = w\right)\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\middle| w_t = w\right].$$

By definition, $\mathbb{P}_{t-1}\left(w_t = w\right) = 0$ if $w \notin S_{t-1} \cup \overline{S}_{t-1}$. Furthermore, $\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\middle| w_t = w\right] = 0$ if $w \in \overline{S}$. Combining this with the above bound in the case that $\mathtt{d}_S(w) \geq 2$, we obtain

$$\begin{aligned}
\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right] &\leq 6\sqrt{\beta}\sum_{w\in V}\mathbb{P}_{t-1}\left(w_t = w\right)\left(\mathtt{d}_S(w)-1\right) \\
&= 6\sqrt{\beta}\mathbb{E}_{t-1}\left[\mathtt{d}_S(w_t)-1\right] = 6\sqrt{\beta}\left(\mathbb{E}_{t-1}\left[\mathtt{d}_S(w_t)-2\right]+1\right).
\end{aligned}$$

By Lemma 7.6.5, $\mathbb{E}_{t-1}\left[\mathtt{d}_S(w_t)-2\right] \geq \varepsilon/4$ for all $t \leq \tau$. Since $\beta = 10^{-6}\varepsilon^2$, the result follows. □

The last ingredient we need is the concentration of X_t. Define $A_t = \mathtt{d}_S(w_t) - \mathbb{E}_{t-1}\left[\mathtt{d}_S(w_t)\right]$ and $B_t = d'_{S,t}(w_t) - \mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right]$. The proof of the concentration of these over all t is identical to the corresponding configuration model result (Lemma 6.2.7), as it still follows that $|A_t|,\ |B_t| \leq \delta\sqrt{M}$ (since the preprocessing step implies that $\mathtt{d}_S(w_t),\ d'_{S,t}(w_t) \leq \delta\sqrt{M}$ for all $t \geq 0$). Thus, we omit the proof here.

Lemma 7.6.7. With probability $1 - o(1)$, both $|\sum_{t'\leq t} A_{t'}|$ and $|\sum_{t'\leq t} B_{t'}|$ are strictly less than $\frac{M}{\log\log M}$ for all $t \leq \tau$.

Recall that τ is the minimum time at which either $X_\tau > \beta M$ or $M_\tau \leq \left(1 - \frac{\varepsilon}{4}\right)M$, where ε is the constant such that $R_S \geq \varepsilon M$ and $\beta = 10^{-6}\varepsilon^2$. We now show that a.a.s. $X_\tau > \beta M$. This proof uses many of the same ideas as the corresponding result in the configuration model (Lemma 6.2.8).

Lemma 7.6.8. With probability $1-o(1)$, at time τ of the deferred decision process we have that $X_\tau > \beta M$.

Proof. The definition of the exploration process implies that

$$X_\tau \geq X_0 + \sum_{t\leq\tau}[\mathsf{d}_S(w_t)-2] - 2\sum_{t\leq\tau} d'_{S,t}(w_t).$$

Recall that $A_t = \mathsf{d}_S(w_t) - \mathbb{E}_{t-1}[\mathsf{d}_S(w_t)]$ and $B_t = d'_{S,t}(w_t) - \mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right]$. Now consider all sequences of partial pairings $(T_t)_{t\geq 1}$ generated by the exploration process such that

$$\left|\sum_{t'\leq t} A_{t'}\right| \leq \frac{M}{\log\log M} \text{ and } \left|\sum_{t'\leq t} B_{t'}\right| \leq \frac{M}{\log\log M} \tag{7.16}$$

for all $t \leq \tau$. Then for each such sequence of partial pairings,

$$X_\tau \geq \sum_{t\leq\tau}\mathbb{E}_{t-1}[\mathsf{d}_S(w_t)-2] - 2\sum_{t\leq\tau}\mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right] - \frac{3M}{\log\log M}.$$

Applying Lemma 7.6.6 then gives that

$$X_\tau \geq \frac{5}{6}\frac{\varepsilon}{4}\tau - \frac{3M}{\log\log M}. \tag{7.17}$$

Suppose, for contradiction, that $X_\tau \leq \beta M$. Then the definition of τ implies that $M_\tau \leq \left(1-\frac{\varepsilon}{4}\right)M$, and thus $\sum_{t\leq\tau}\mathsf{d}_S(w_t) \geq \frac{\varepsilon}{4}M$. We consider two cases based on the value of τ. Firstly, if $\tau \geq \frac{3\varepsilon}{64}M$, then the inequality in (7.17) immediately gives the desired contradiction. On the other hand, if $\tau < \frac{3\varepsilon}{64}M$, then Lemma 7.6.6 and an identical computation to that given in (6.6) imply that

$$X_\tau \geq \frac{5}{6}\sum_{t\geq\tau}[\mathsf{d}_S(w_t)-2] + \frac{1}{6}\sum_{t\leq\tau}(A_t - 2B_t),$$

Since we are assuming that $\tau < \frac{3\varepsilon}{64}M$, $\sum_{t\leq\tau}\mathsf{d}_S(w_t) \geq \frac{\varepsilon}{4}M$, and the inequalities given in (7.16) hold for this sequence of partial pairings, then

$$X_\tau \geq \frac{5\varepsilon}{24}M - 2\tau - \frac{M}{2\log\log M} > \beta M.$$

Therefore, if the inequalities given in (7.16) are satisfied, $X_\tau > \beta M$. Lemma 7.6.7 states that a.a.s. these inequalities hold for all $t \leq M$. This implies that a.a.s. $X_\tau > \beta M$, which completes the proof. □

Lemma 7.6.8 shows that a.a.s. the partial pairing T_τ contains a large number of open edges. These open edges, by definition, must all be in the same connected component of T_τ. Therefore, all their parent vertices belong to the same component of $G[S]$. However, this does not imply that the component in $G[S]$ contains $\Theta(M)$ edges, as it is possible that all or almost all of these open edges pair with half edges in $\overline{S}$. We show that a.a.s. this is not the case, and that a positive fraction of these open edges pair with half edges in S_τ (recall $S_\tau = S\backslash(V_\tau)$). This implies that $G[S]$ a.a.s. contains a component with $\Theta(M)$ edges. Since the preprocessing step moves all vertices

from S^+ to $\overline{S}$, in the following proof we again refer to the set of all vertices with degree greater than $\delta\sqrt{M}$ simply as $\overline{S}^+$.

Lemma 7.6.9. Suppose $(\boldsymbol{d}, S)$ satisfies the conditions of Theorem 7.1.2 and $R_S \geq \varepsilon M$ for some constant $\varepsilon > 0$. Then a.a.s. $G[S]$ contains a component with at least $\frac{1}{5}\varepsilon\beta M$ edges.

Proof. Let T be a partial pairing. Let $\Phi(T)$ be the subset of simple pairings $G^* \in \Phi$ such that there exists a sequence of choices for half edges at each step of the exploration process such that $T_\tau(G^*) = T$. Suppose that T has X_τ open edges for some $X_\tau > \beta M$ and that $\Phi(T)$ is non-empty. We define a switching that maps elements of $\Phi(T)$ to $\Phi(T)$.

Let A_i be the set of simple pairings $G^* \in \Phi(T)$ such that exactly i of the X_τ open edges in $T_\tau(G^*)$ are paired with half edges in S_τ. We define a switching between A_{i+1} and A_i. Suppose $G^* \in A_{i+1}$. Let uv be an edge in G^* such that $u \in V_\tau \cap S$ and $v \in S_\tau$. Let (x, y) be an ordered pair of adjacent vertices such that $x \in \overline{S}_\tau$ (note that this means (x, y) and (y, x) are counted separately if $x, y \in \overline{S}_\tau$). A switching takes the edges uv and xy, deletes them, and replaces them with the edges ux and vy, creating a new pairing G'^*. This switching is valid as long as G'^* is a simple pairing in A_i. This occurs if and only if the vertices $\{u, v, x, y\}$ are distinct, neither of the edges ux or vy are present in G^*, and $y \notin V_\tau \cap S$. The number of forward switchings for all $G^* \in A_{i+1}$ is the number of choices for vertices $\{u, v, x, y\}$ that satisfy the above restrictions. There are exactly $i + 1$ choices for the edge uv so that $u \in V_\tau \cap S$ and $v \in S_\tau$. There are at most $d(\overline{S}_\tau)$ choices for vertex x and neighbour y, as this is the number of choices for a half edge with parent vertex in $\overline{S}_\tau$. Thus, the number of forward switchings is at most $(i + 1)d(\overline{S}_\tau)$.

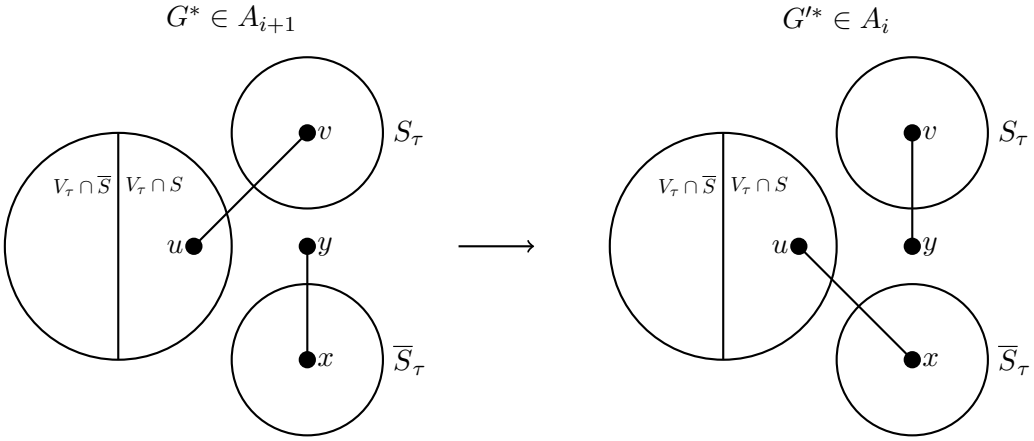

Figure 7.9: A diagram of the switching used in the proof of Lemma 7.6.9, which sends $G^* \in A_{i+1}$ to $G'^* \in A_i$. The position of vertex y in the diagram is to denote that $y \in S_\tau$ and $y \in \overline{S}_\tau$ are both allowed.

Now suppose $G'^* \in A_i$. We now look for a lower bound on the number of switchings that create G'^*, equivalently the number of backward switchings from some $G'^* \in A_i$ to $G^* \in A_{i+1}$. There are $X_\tau - i$ choices for the edge ux such that $u \in V_\tau \cap S$ and $x \in \overline{S}_\tau$. There are $d(S_\tau)$ choices for an ordered pair of adjacent vertices (v, y) such that $v \in S_\tau$, with no restrictions on y. One of these choices does not correspond to a valid switching if

(a) $uv \in E(G'^*)$,

(b) $xy \in E(G'^*)$,

(c) vertices $\{u, v, x, y\}$ are not distinct,

(d) $y \in V_\tau \cap S$.

Let $W(G'^*)$ be the number of choices that satisfy at least one of (a) – (d). It follows that the number of valid backward switchings that can be applied to G'^* is $(X_\tau - i)d(S_\tau) - W(G^*)$. Thus, we bound W from above. We do so by bounding $Z(G'^*)$ from above, where $Z(G'^*)$ is the number of choices for $\{u, v, x, y\}$ such that

(a′) $x \in \overline{S}^+$ or $y \in \overline{S}^+$,

(b′) $x, y \notin \overline{S}^+$ and $uv \in E(G'^*)$,

(c′) $x, y \notin \overline{S}^+$ and $xy \in E(G'^*)$,

(d′) $x, y \notin \overline{S}^+$ and the vertices $\{u, v, x, y\}$ are not distinct,

(e′) $y \in V_\tau \cap S$.

It follows that $W \leq Z$ for all $G'^* \in A_i$. We now show that

$$Z \leq (X_\tau - i)\left(d(\overline{S}^+) + i + 3\delta^2 M\right) + d(S_\tau)d(\overline{S}^+). \tag{7.18}$$

For case (a′), there are $X_\tau - i$ choices for edge ux such that $u \in V_\tau \cap S$ and $x \in \overline{S}_\tau$. There are at most $d(\overline{S}^+)$ choices for a vertex $y \in \overline{S}^+$ and a neighbour v. Thus, there are at most $(X_\tau - i)d(\overline{S}^+)$ choices for $\{u, v, x, y\}$ such that $u \in S \cap V_\tau$, $x \in \overline{S}_\tau$, and $y \in \overline{S}^+$. Similarly, there are $d(\overline{S}^+)$ choices for a vertex $x \in \overline{S}^+$ and neighbour u, and $d(S_\tau)$ choices for $\{v, y\}$ such that $vy \in E(G'^*)$ and $v \in S_\tau$. Thus, there are at most $d(\overline{S}^+)d(S_\tau)$ choices for $\{u, v, x, y\}$ such that $x \in \overline{S}^+$. Therefore, there are at most $(X_\tau - i)d(\overline{S}^+) + d(S_\tau)d(\overline{S}^+)$ choices for $\{u, v, x, y\}$ that satisfy (a′). For case (b′), there are $X_\tau - i$ choices for the edge ux. Then, given this choice, there are at most $d(u)$ choices for a neighbour v of u such that $v \in S_\tau$. Then there are $d(v)$ choices for a neighbour y of v. Note that by assumption both u and v have degree at most $\delta\sqrt{M}$. Thus, at most $(X_\tau - i)\delta^2 M$ choices satisfy (b′). For case (c′), since we assume that $x, y \notin \overline{S}^+$, by similar reasoning at most $(X_\tau - i)\delta^2 M$ of the choices for $\{u, v, x, y\}$ have edge xy present. For part (d′), the only possibilities for non-distinct vertices are if $x = y$ or $u = y$. There are $X_\tau - i$ choices for ux and at most $d(u) + d(x)$ choices for v, since $vy \in E(G'^*)$ and either $y = x$ or $y = u$. Thus, there are at most $2(X_\tau - i)\delta\sqrt{M}$ choices for that satisfy (d′), since again $x, y \notin \overline{S}^+$. Finally, we bound the number of choices that satisfy (e′). There are exactly i choices for an edge yv such that $y \in V_\tau \cap S$ and $v \in S_\tau$, and $X_\tau - i$ choices for the edge ux such that $u \in V_\tau \cap S$ and $x \in \overline{S}_\tau$. So it follows that

$$Z \leq (X_\tau - i)\left(d(\overline{S}^+) + i + 2\delta^2 M + 2\delta\sqrt{M}\right) + d(S_\tau)d(\overline{S}^+).$$

Recall that $\delta \to 0$ and $\delta = \omega(\log^{-1} M)$. The bound in (7.18) then follows immediately. Thus, the number of backward switchings that can be applied to $G'^* \in A_i$ is at least

$$(X_\tau - i)\left(d(S_\tau) - i - d(\overline{S}^+) - 3\delta^2 M\right) - d(S_\tau)d(\overline{S}^+). \tag{7.19}$$

We know by the definition of τ that $d(S_\tau) \geq R_S - \frac{\varepsilon}{4}M - \delta\sqrt{M} \geq \frac{3}{4}\varepsilon M(1-o(1))$. Recall that $X_\tau > \beta M$, since we assume that T has more than βM open edges. The definition of τ implies that $X_{\tau-1} \leq \beta M$. Thus, $X_\tau \leq \beta M + \delta\sqrt{M} = \beta M(1+o(1))$. Also recall that $d(\overline{S}^+) \leq \delta^2 n \leq \delta M$. Therefore, for all $i \leq X_\tau/2$, the expression given in (7.19) is at least

$$(X_\tau - i)d(S_\tau)\left(1 - \frac{3}{2\varepsilon}\beta - O\left(\frac{\delta}{\varepsilon}\right)\right).$$

Recall that $\beta = 10^{-6}\varepsilon^2$. Thus, it follows for all $i \leq X_\tau/2$ that

$$\frac{|A_i|}{|A_{i+1}|} \leq \frac{(i+1)d(\overline{S}_\tau)}{(X_\tau - i)d(S_\tau)}\left(1 + 10^{-5}\varepsilon + O\left(\frac{\delta}{\varepsilon}\right)\right).$$

Define $i_0 = \frac{d(S_\tau)}{2M}X_\tau$; note that since $d(S_\tau) < M$ it follows that $i_0 < X_\tau/2$. Now suppose that $i+1 \leq i_0$. Then

$$\begin{aligned}
\frac{|A_i|}{|A_{i+1}|} &\leq \frac{\frac{d(S_\tau)}{2M}X_\tau d(\overline{S}_\tau)}{X_\tau\left(1 - \frac{d(S_\tau)}{2M}\right)d(S_\tau)}(1 + 10^{-5}\varepsilon + O(\delta/\varepsilon)) \\
&= \frac{d(\overline{S}_\tau)}{2M - d(S_\tau)}(1 + 10^{-5}\varepsilon + O(\delta/\varepsilon)) \\
&\leq \frac{d(\overline{S}_\tau)}{M + d(\overline{S}_\tau)}(1 + 10^{-5}\varepsilon + O(\delta/\varepsilon)) \\
&\leq \frac{d(\overline{S}_\tau)}{2d(\overline{S}_\tau) + d(S_\tau)}(1 + 10^{-5}\varepsilon + O(\delta/\varepsilon)) \\
&\leq \frac{1}{2+\varepsilon}(1 + 10^{-5}\varepsilon + O(\delta/\varepsilon)) \\
&< \frac{1}{2},
\end{aligned}$$

where the last inequality follows from the fact that $\delta \to 0$ and ε is fixed. So for all $i \leq i_0 - \log M$,

$$\frac{|A_i|}{|A_{i_0}|} \leq 2^{-\log M}.$$

Since $i_0 = \Theta(M)$, it follows that $i_0 > \log M$. Thus, we obtain a bound on the size of the union of A_i for all $i \leq i_0 - \log M$:

$$\sum_{i=0}^{i_0 - \log M} |A_i| \leq |A_{i_0}| \sum_{i=0}^{i_0 - \log M} \frac{1}{2^{\log M + i}} \leq 2^{1-\log M}.$$

Note that if $X_\tau > \beta M$, then

$$i_0 - \log M = \frac{d(S_\tau)}{2M}X_\tau \geq \frac{3}{8}\varepsilon\beta M(1-o(1)) - \log M > \frac{1}{5}\varepsilon\beta M.$$

Recall that G^* is a uniformly random simple pairing. Thus, if T is a partial pairing with more than βM open edges and $\Phi(T)$ is non-empty, it follows that

$$\mathbb{P}\left(|E(V_\tau \cap S, S_\tau)| \leq \frac{1}{5}\varepsilon\beta M \,\middle|\, T_\tau = T\right) \leq 2^{1-\log M}.$$

Thus, for a uniformly random $G^* \in \Phi$, it follows that

$$\mathbb{P}\left(|E(V_\tau \cap S, S_\tau)| \leq \frac{1}{5}\varepsilon\beta M\right) \leq 2^{1-\log M} + \mathbb{P}\left(X_\tau \leq \beta M\right). \tag{7.20}$$

By Lemma 7.6.8, it follows that this is $o(1)$. Therefore, (7.20) immediately implies that the probability that $G[S]$ contains no component with more than $\varepsilon\beta M/5$ edges is $o(1)$. This completes the proof. □

Thus, in the case that $R_S \geq \varepsilon M$ and $d(S^+) = o(M)$, a.a.s. $G[S\backslash S^+]$ contains a connected component containing $\Theta(M)$ edges, and thus a.a.s. $G[S]$ contains a component with $\Theta(M)$ edges. We now shift focus to the case where $d(S^+) = \Theta(M)$.

7.6.2 The case where $d(S^+) = \Theta(M)$

In this section we prove the following lemma. This, in combination with Lemmas 7.1.4 and 7.6.9, shows that in the supercritical regime there a.a.s. exists a component in $G[S]$ with $\Theta(M)$ edges, which proves Theorem 7.1.2(a).

Lemma 7.6.10. Suppose that $(\boldsymbol{d}, S)$ satisfies the conditions of Theorem 7.1.2 and also that $d(S^+) = \Theta(M)$. Then $G[S]$ a.a.s. contains a component with $\Theta(M)$ edges.

We do not analyse this case using the exploration process. Thus, we revert our discussion and our proofs back to graphs $G \in \mathcal{G}(\boldsymbol{d})$ instead of simple pairings $G^* \in \Phi$. Lemma 7.1.4 proves Lemma 7.6.10 in the case where $M = \omega(|S|)$. Thus, we only need to consider the case where $M = \Theta(|S|)$, which immediately implies that $M = \Theta(n)$ and $|S| = \Theta(n)$. A consequence of these assumptions is that $|S\backslash S^+| = |S|(1 - o(1))$, as it implies that

$$|S^+| = O(\delta^{-1}\sqrt{M}) = o(n).$$

Consequently, there exists some constant $\xi \in (0, 1)$ such that $|S\backslash S^+| \geq 2\xi n$. This implies the assumption from Theorem 7.1.2 that $d(S\backslash S^+) = \Theta(M)$.

Unlike in the previous subsection, the proof of this result does not involve moving S^+ into $\overline{S}$. Thus, we emphasise the distinction between the sets H^+, the set of all vertices with degree greater than $\delta\sqrt{M}$ in G, and the set $\overline{S}^+$, which is the subset of H^+ contained in $\overline{S}$. Naturally, since $S^+ = \Theta(M)$, the set S^+ is non-empty.

To prove that $G[S]$ a.a.s. contains a component with $\Theta(M)$ edges in this case, we use a more involved switching argument described by Hasheminezhad and McKay [65], with the relevant lemma given in Lemma 2.3.1. The idea is that we want to define a switching that takes an edge $uv \in E(G)$, where $u \in S^+$ and $v \in \overline{S}\backslash\overline{S}^+$, and replaces it with an edge between S^+ and $S\backslash S^+$. The naive way to do this is to choose some ab where $a \in S\backslash S^+$ and swap out the pair of edges $\{uv, ab\}$ for $\{ua, vb\}$. The problem comes from trying to obtain an upper bound on the number of choices for $\{u, v, a, b\}$ such that the edge ua is present. Since $u \in S^+$, it might potentially have very high degree (potentially even degree $n - 1$), and as such there might be few good choices for the vertex a such that $ua \notin E(G)$.

However, in some sense this case isn't really "bad". If some vertex $u \in S^+$ has many neighbours in $S\backslash S^+$, then each neighbour corresponds to another edge in $G[S]$ — specifically, another edge

incident to S^+. On the other hand, if some vertex $u \in S^+$ has few neighbours in $S \backslash S^+$, then there are many good choices for the edge ab to perform the switching. So if it is the case that there are very few valid switching choices, then there must be many edges between S^+ and $S \backslash S^+$. To account for this in the switching argument, we apply the framework outlined by Hasheminezhad and McKay [65] and the bound in Lemma 2.3.1.

Lemma 7.6.11. Suppose that $d(S^+) = \Theta(M)$ and the conditions of Theorem 7.1.2 are satisfied. Suppose that $M = \Theta(n)$ and $|S \backslash S^+| \geq 2\xi n$ for some fixed $\xi > 0$. Then a.a.s. there are $\Theta(M)$ edges between vertices in S^+ and vertices in S.

Proof. First we define the switching we use. For some graph G, let uv be an edge in G where $u \in S^+$ and $v \in \overline{S} \backslash \overline{S}^+$. Let (x, y) be an ordered pair of adjacent vertices such that $x \in S \backslash S^+$ and $y \notin H^+$. The switching then deletes the two edges xy and uv and replaces them with ux and vy, creating a new graph G'. This switching is considered valid if $G' \in \mathcal{G}(\boldsymbol{d})$, which occurs if and only if $ux, vy \notin E(G)$ and the vertices $\{u, v, x, y\}$ are all distinct. A diagram of this switching is given in Figure 7.10. Since we assume that $M = \Theta(n)$, we define $p, q = \Theta(1)$ such that $d(S^+) = pn$ and

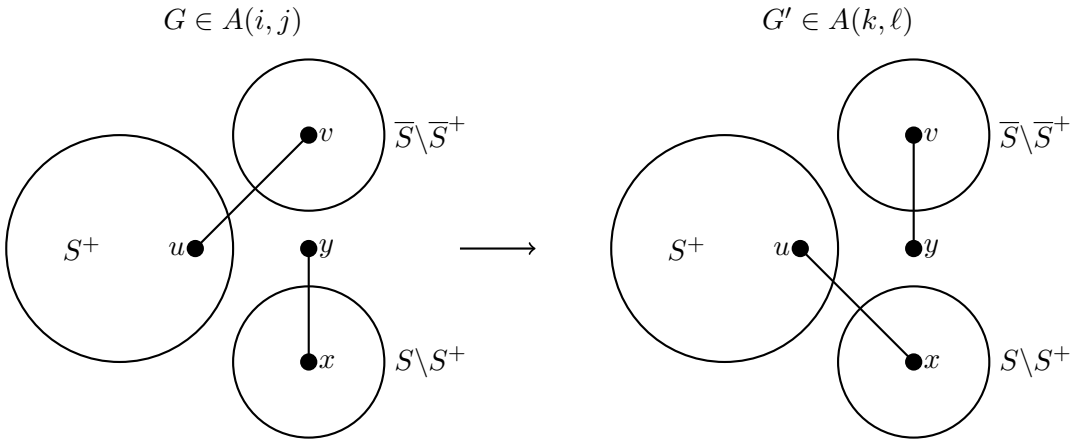

Figure 7.10: A diagram of the switching used in the proof of Lemma 7.6.11, which sends $G \in A(i, j)$ to $G' \in A(k, \ell)$. The definition of the switching implies that $k = i - 1$ and either $\ell = j$ or $\ell = j - 1$. The position of vertex y in the diagram is to denote that $y \in S_\tau$ and $y \in \overline{S}_\tau$ are both allowed.

$d(\overline{S}) = qn$ for notational convenience.

Let $N_G(S^+)$ denote the set of vertices in G with a neighbour in S^+. We now define a structure graph $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ for this switching. Define

$$A(i,j) = \left\{ G \in \mathcal{G}(\boldsymbol{d}) \ : \ |E(S^+, \overline{S} \backslash \overline{S}^+)| = i \text{ and } |S \backslash \left\{ S^+ \cup N_G(S^+) \right\}| = j \right\}.$$

Let

$$\mathfrak{V} := \{0, \dots, pn\} \times \{0, , \dots, |S \backslash S^+|\}$$

and

$$\mathfrak{E} := \{((i,j),(k,\ell)) \mid \exists G \in A(i,j) \text{ and } G' \in A(k,\ell) \text{ and a switching mapping } G \text{ to } G'\}.$$

The set $\{A(i,j)\}_{(i,j)\in\mathfrak{V}}$ forms a partition of $\mathcal{G}(\boldsymbol{d})$. Define the sets $\mathfrak{X}$ and $\mathfrak{Y}$ as follows:

$$\mathfrak{Y} := \left\{(i,j) \in \mathfrak{V} \mid i \geq \frac{pq}{\frac{1}{2}\xi^2 + q} n \text{ and } j \geq \xi^2 n\right\}$$

and

$$\mathfrak{X} := \left\{(i,j) \in \mathfrak{V} \mid i \leq \frac{pq}{\frac{2}{3}\xi^2 + q} n \text{ or } j \leq \frac{3}{4}\xi^2 n\right\}.$$

For now, we defer the proof of the following claim.

Claim 7.6.12. $\mathfrak{X}$ contains all the sinks of $\mathfrak{G}$, as well as all $A \in \mathfrak{V}$ such that $\alpha(A \to B) \geq 1$ for some $B \in \mathfrak{V}$.

Thus, $\mathfrak{X}$ satisfies the requirements of Lemma 2.3.1. The fact that $\xi > 0$ implies that $\frac{pq}{\frac{1}{2}\xi^2+q} < p - \rho$ for some constant $\rho > 0$. Thus, every graph contained in some $A(i,j)$ where $(i,j) \notin \mathfrak{Y}$ has at least ρn edges between S^+ and $S \backslash S^+$. The assumptions of this lemma imply that $\rho n = \Theta(M)$.

Note that each switching operation removes exactly one edge from $E(S^+, \overline{S} \backslash \overline{S}^+)$, and adds at most one vertex to $N_G(S^+)$ (the vertex x, since the presence of edge vy in G' does not change whether $y \in N_G(S^+)$). Thus, for all $G \in A(i,j)$ and $G' \in A(k,\ell)$ such that there exists a switching mapping G to G', it must be true that $i - k = 1$ and $j - \ell \in \{0,1\}$. As a consequence of this, not only are $\mathfrak{X}$ and $\mathfrak{Y}$ disjoint, but all paths in $\mathfrak{G}$ between $A \in \mathfrak{Y}$ and $B \in \mathfrak{X}$ must have length greater than cn, where

$$c = \frac{pq\xi}{6(1+p)\left(1 + \frac{1}{3}\xi^2 + q^2 + \frac{7}{6}q\xi^2\right)},$$

noting that

$$\frac{pq}{\frac{1}{2}\xi^2 + q} - \frac{pq}{\frac{2}{3}\xi^2 + q} = pq\left(\frac{\frac{1}{6}\xi^2}{\frac{1}{3}\xi^4 + q^2 + \frac{7}{6}\xi^2 q}\right) > c.$$

and $c < \frac{1}{4}\xi^2$ for all $p, q > 0$.

Since $c = \Theta(1)$, this means that every path in $\mathfrak{G}$ that starts in $\mathfrak{Y}$ and ends in $\mathfrak{X}$ is of length $\Theta(n)$. Another consequence of the relationships between $\{i,j,k,\ell\}$ is that there are no edges $(A,B) \in \mathfrak{E}$ such that $A \notin \mathfrak{Y}$ and $B \in \mathfrak{Y}$. Thus, the set $\mathcal{Q}_{\mathfrak{Y},3}(\mathfrak{Y},\mathfrak{Y})$ as defined in Lemma 2.3.1 contains only single-edge paths (Y_1, Y_2) where $Y_1, Y_2 \in \mathfrak{Y}$.

Recall from Section 2.3 the definition of $a(v)$ and $b(v)$ for $v \in \mathfrak{V}$, that is, on average each graph in $A(v)$ can be subject to at least $a(v)$ switchings and be created by at most $b(v)$ switchings. Next we prove the following result about $a((i,j))$ and $b((i,j))$ for all $(i,j) \in \mathfrak{V}$.

Claim 7.6.13. Every $G \in A(i+1,j)$ can be subject to at least $(i+1)(j-3\delta M)$ switchings. Every $G \in A(i,j)$ can be created by at most $(d(S^+) - i)d(\overline{S})$ switchings.

Thus, for this switching operation, for all $(A,B) \in \mathfrak{E}$ we have that

$$\alpha(A \to B) := \frac{(d(S^+) - i)d(\overline{S})}{(i+1)(j-3\delta M)}. \tag{7.21}$$

Proof of Claim 7.6.13. Let $G \in A(i+1,j)$. Then there are exactly $i+1$ choices for the edge uv such that $u \in S^+$ and $v \in \overline{S}\backslash\overline{S}^+$. Given uv, there are at least j vertices in $S\backslash S^+$ that are not adjacent to u and have at least one neighbour in $[n]\backslash S^+$. Thus, there are at least j choices for an ordered pair of adjacent vertices (x,y) such that $x \in S\backslash S^+$ and $ux \notin G$. One of these choices for $\{u,v,x,y\}$ does not correspond to a valid switching only if

(a) $y \in H^+$,

(b) $ux \in E(G)$ or $vy \in E(G)$,

(c) $\{u,v,x,y\}$ are not distinct.

Let $W(G)$ be the number of choices for $\{u,v,x,y\}$ that satisfy one of (a) – (c). Due to the way $\{u,v,x,y\}$ are chosen, these conditions are equivalent to the following:

(a′) $y \in \overline{S}^+$,

(b′) $y \notin \overline{S}^+$ and $vy \in G$,

(c′) $y \notin \overline{S}^+$ and $v = y$.

Thus, we bound $W(G)$ from above for all graphs $G \in A(i+1,j)$ by bounding the number of choices for $\{u,v,x,y\}$ that satisfy one of (a′) – (c′). For case (a′), there are $i+1$ choices for the edge uv. There are $d(\overline{S}^+)$ choices for an ordered pair of adjacent vertices (x,y) such that $y \in \overline{S}^+$. Thus, the number of choices that satisfy (a′) is at most $(i+1)d(\overline{S}^+)$. For case (b′), note that $v \in \overline{S}\backslash\overline{S}^+$ and $y \notin H^+$, and thus both vertices have maximum degree (in G) $\delta\sqrt{M}$. Again there are $i+1$ choices for the edge uv. Given v, there are $d(v)$ choices for a neighbour y of v, and then $d(y)$ choices for a neighbour x of y. Thus, the number of choices for $\{u,v,x,y\}$ that satisfy (b′) is at most $(i+1)\delta^2 M$. By similar reasoning, there are at most $(i+1)\delta\sqrt{M}$ choices for $\{u,v,x,y\}$ such that $v = y$. Thus, it follows that

$$W \leq (i+1)\left(d(\overline{S}^+) + \delta^2 M + \delta\sqrt{M}\right).$$

Since $d(\overline{S}^+) \leq \delta^2 n \leq \delta M$ and $\delta = o(1)$, it follows that $W \leq (i+1)3\delta M$. Therefore, the number of forward switchings that can be applied to some $G \in A(i+1,j)$ is at least $(i+1)(j-3\delta M)$.

The number of switchings into any element of $A(i,j)$ is at most the number of choices for an edge ux where $u \in S^+$ and $v \in S\backslash S^+$, as well as an ordered pair of adjacent vertices (v,y) where $v \in \overline{S}\backslash\overline{S}^+$. The number of choices for vertex u and neighbour x is at most $d(S^+) - i$. The number of choices for (v,y) is at most $d(\overline{S})$. Thus, the number of switchings into each $G' \in A(i,j)$ is at most $(d(S^+) - i)d(\overline{S})$. □

With these ingredients, we now apply Lemma 2.3.1. Recalling the definition of $\alpha_{\mathfrak{Y},\mathfrak{Z}}(UV)$ from Equation (2.3), this lemma says that

$$\frac{N(\mathfrak{Y})}{N(\mathfrak{X})} \leq \frac{\alpha_{\mathfrak{Y},\mathfrak{Z}}(\mathfrak{Y}\mathfrak{Z})}{1 - \alpha_{\mathfrak{Y},\mathfrak{Z}}(\mathfrak{Y}\mathfrak{Y})}.$$

Since $\mathfrak{Z} \subset \mathfrak{X}$, it follows that $\alpha_{\mathfrak{Y},\mathfrak{Z}}(\mathfrak{Y}\mathfrak{Z}) \leq \alpha_{\mathfrak{Y},\mathfrak{Z}}(\mathfrak{Y}\mathfrak{X})$. For all $u \notin \mathfrak{X}$ and v such that $uv \in \mathfrak{E}$, we show that $\alpha(uv) < 9/10$. Note that for all $((i,j),B) \in \mathfrak{E}$, the function $\alpha((i,j) \to B)$ is decreasing

in both i and j, and thus it is sufficient to check this on the boundary of $\mathfrak{X}$. That is, define $(i_0, j_0) = \left(\frac{pq}{\frac{2}{3}\xi^2+q}n, \frac{3}{4}\xi^2 n\right)$. Then

$$\begin{aligned}\alpha((i+1,j)\to B) \le \alpha\left((i_0,j_0)\to(i_0,j_0)\right) &= \frac{pqn^2 - \frac{pq}{\frac{2}{3}\xi^2+q}n\cdot qn}{\frac{pq}{\frac{2}{3}\xi^2+q}n\cdot\frac{3}{4}\xi^2 n}(1+o(1))\\ &= \frac{\left(1-\frac{q}{\frac{2}{3}\xi^2+q}\right)}{\left(\frac{\frac{3}{4}\xi^2}{\frac{2}{3}\xi^2+q}\right)}(1+o(1))\\ &= \frac{\frac{2}{3}\xi^2+q-q}{\frac{3}{4}\xi^2}(1+o(1))\\ &= \frac{2/3}{3/4}(1+o(1))\\ &< \frac{9}{10}.\end{aligned}$$

Now we combine this with the bounds on the lengths of paths in $\mathcal{Q}_{\mathfrak{Y},3}(\mathfrak{Y},\mathfrak{X})$ and $\mathcal{Q}_{\mathfrak{Y},3}(\mathfrak{Y},\mathfrak{Y})$ given earlier. Altogether this gives that

$$\max_{Q\in\mathcal{Q}_{\mathfrak{Y},3}(\mathfrak{Y},\mathfrak{X})}\alpha(Q) \le \left(\frac{9}{10}\right)^{cn} \quad \text{and} \quad \max_{Q\in\mathcal{Q}_{\mathfrak{Y},3}(\mathfrak{Y},\mathfrak{Y})}\alpha(Q) \le \frac{9}{10}.$$

Therefore, Lemma 2.3.1 implies that

$$\frac{N(\mathfrak{Y})}{N(\mathfrak{X})} \le 10\left(\frac{9}{10}\right)^{cn} = o(1).$$

Thus, a.a.s. G either has at most $\frac{pq}{\frac{1}{2}\xi^2+q}n$ edges between S^+ and $\overline{S}\backslash\overline{S}^+$ or has no more than $\xi^2 n$ vertices in $S\backslash S^+$ with no neighbours in S^+. To complete the application of Lemma 2.3.1, we now prove Claim 7.6.12.

Proof of Claim 7.6.12. Consider some pair $(i+1,j)$ such that $\alpha((i+1,j)\to B)\ge 1$. Then it follows from Equation (7.21) that

$$(i+1)(j-3\delta M) \le (d(S^+)-i)d(\overline{S}),$$

which rearranges to give that

$$i \le \frac{d(S^+)d(\overline{S}) - j + 3\delta M}{d(\overline{S}) + j - 3\delta M}.$$

If $j \le \frac{3}{4}\xi^2 n$, then $(i+1,j)\in\mathfrak{X}$ by definition. Otherwise, suppose that $j \ge \frac{3}{4}\xi^2 n$ and $\alpha((i+1,j)\to$

$B) \geq 1$. Then since $M = \Theta(n)$ it follows that

$$\begin{aligned} i \leq \frac{d(S^+)d(\overline{S}) - j + 3\delta M}{d(\overline{S}) + j - 3\delta M} &= \frac{d(S^+)d(\overline{S})}{d(\overline{S}) + j}(1 + o(1)) \\ &\leq \frac{pq}{q + \frac{3}{4}\xi^2} n(1 + o(1)) \\ &< \frac{pq}{q + \frac{2}{3}\xi^2} n. \end{aligned}$$

Thus, it follows that $(i+1, j) \in \mathfrak{X}$. Therefore, $\mathfrak{X}$ contains all $u \in \mathfrak{V}$ for which $\alpha(uv) \geq 1$ for some $uv \in \mathfrak{E}$. Furthermore, for every graph G, at least one switching can be applied to G unless either $E(S^+, \overline{S} \backslash \overline{S}^+) = \emptyset$ or, for every valid choice of uv, there are no valid choices for (x, y). From the lower bound on $a((i, j))$ found above, this can only occur if either $i = 0$ or $j \leq 3\delta M$. For all of these cases, it follows that $(i, j) \in \mathfrak{X}$ by definition, since $\delta = o(1)$ and $M = \Theta(n)$. Thus, $\mathfrak{X}$ contains all sinks of $\mathfrak{G}$. □

If G has at most $\frac{pq}{\frac{1}{2}\xi^2 + q} n$ edges between S^+ and $\overline{S} \backslash \overline{S}^+$, then since $d(S^+) = pn$ and $d(\overline{S}^+) \leq \delta^2 n$ there must be $\Theta(n)$ edges incident to S^+ that are either incident to $S \backslash S^+$ or have both endpoints in S^+. On the other hand, if G at most $\xi^2 n$ vertices in $S \backslash S^+$ with no neighbours in S^+, then at least $(\xi - \xi^2)n$ vertices in $S \backslash S^+$ have at least one neighbour in S^+ (recall that $\xi < 1$). Thus, for all $G \in A(i, j)$ where $(i, j) \notin \mathfrak{Y}$, there are at $\Theta(n)$ edges between S^+ and S. Since $M = \Theta(n)$, this completes the proof. □

Recall that $|S^+| \leq \delta^{-1}\sqrt{M}$, since every vertex in S^+ has degree at least $\delta\sqrt{M}$ in G. The previous lemma then implies that a.a.s. the total degree of S^+ in $G[S]$ is $\Theta(n)$, and thus $\Theta(M)$ by assumption. The following lemma shows that if a uniformly random graph with a fixed degree sequence has a small set of vertices with high average degree, then this graph a.a.s. has a component with $\Theta(M)$ edges.

Lemma 7.6.14. Suppose $G \sim \mathcal{G}(\boldsymbol{d})$, where $\boldsymbol{d}$ is an n-element degree sequence with total degree $M := M(\boldsymbol{d})$. Suppose G contains a set of vertices X such that $|X| \leq \delta n$ for some $\delta \to 0$, and $d(X) \geq \alpha M$ for some constant $\alpha > 0$. Then $R(\boldsymbol{d}) \geq \frac{1}{4}\alpha M$.

Proof. Without loss of generality, we suppose that $\boldsymbol{d}$ is ordered in non-decreasing order. Let $X = \{x, x+1, \ldots, n\}$ be a set of vertices with total degree αM, for some $\alpha > 0$. Since $x(x-2)$ has a minimum value of -1, we know that

$$\sum_{i=1}^{x-1} d(i)(d(i) - 2) > -M.$$

Suppose that X contains a vertex of degree at least $\frac{1}{4}\alpha M$. Then immediately $R(\boldsymbol{d}) \geq \frac{1}{4}\alpha M$, and thus the lemma is true in this case. Now suppose that X has maximum degree less than $\frac{1}{4}\alpha M$. Let (X_1, X_2) be a partition of X such that $X_1 = \{x, x+1, \ldots, x_1\}$, $X_2 = \{x_1 + 1, \ldots, n\}$ and X_1 has total degree at least $\frac{1}{4}\alpha M$. Since X has maximum degree less than $\frac{1}{4}\alpha M$, we can choose X_1 such that X_2 has total degree at least $\frac{1}{2}\alpha M$. Since $|X_1| \leq |X|$, it follows that $d(X_1)/|X_1| \geq$

$\frac{1}{4}\alpha\delta^{-1} \to \infty$. Then

$$\sum_{i=1}^{x_1} d(i)(d(i)-2) > -M + \sum_{i=x}^{x_1} d(i)(d(i)-2) \geq -M + |X_1|\frac{d(X_1)}{|X_1|}\left(\frac{d(X_1)}{|X_1|} - 2\right) > 0.$$

Thus, $j(\boldsymbol{d}) \leq x_1$, and therefore $R(\boldsymbol{d}) \geq \frac{1}{2}\alpha M$. This completes the proof. □

We now combine this result with Claim 7.2.3 and Lemma 7.6.11 to show that $G[S]$ a.a.s. contains a component with $\Theta(M)$ edges. This completes the proof of Lemma 7.6.10.

Proof of Lemma 7.6.10. Lemma 7.6.11 implies that there exists some constant $c > 0$ such that a.a.s. the set of vertices S^+ has total degree at least cM in $G[S]$. We also know that $|S^+| \leq \delta^{-1}\sqrt{M}$ by definition. Thus, if $\boldsymbol{t}$ is the degree sequence of the induced graph $G[S]$ (ordered in non-decreasing order and excluding isolated vertices), a.a.s. $\boldsymbol{t}$ contains a set of elements X (specifically, the elements corresponding to the degrees of the vertices in S^+) such that $|X| \leq \delta^{-1}M$ and $t(X) \geq cM$ for some constant $c > 0$. Conditional on the induced graph having such a degree sequence, Lemma 7.6.14 implies that $R(\boldsymbol{t}) \geq \frac{1}{4}cM(\boldsymbol{t})$. Thus, there exists a constant $\varepsilon := \frac{1}{4}c > 0$ such that a.a.s. $R(\boldsymbol{t}) \geq \varepsilon M$. Thus, Claim 7.2.3 implies that a.a.s. $G[S]$ has a component with $\Theta(M)$ edges. This completes the proof. □

7.7 From size to order, with stricter conditions

In this section we prove Theorem 7.1.2(b). Note that we are no longer assuming that $M = \Theta(|S|)$ or even $M = \Theta(n)$. Unlike when looking for giant components in $\mathcal{G}(\boldsymbol{d})$ [83], the a.a.s. existence of a component with $\Theta(M)$ edges in $G[S]$ does not imply the a.a.s. existence of a component with $\Theta(|S|)$ vertices. In fact, it can even be possible that $G[S]$ a.a.s. has a component with $\Theta(M)$ edges, but almost all vertices in $G[S]$ have degree 0 always. Some examples of such unruly pairs $(\boldsymbol{d}, S)$ are given in Section 7.8.

Thus, we unfortunately cannot always infer that $G[S]$ a.a.s. has a linear-order component even if it a.a.s. has a linear-size component. However, the counterexamples that we give in Section 7.8 all involve vertices of linear degree in $\overline{S}$. We prove that if $d(\overline{S}^+) = o(n)$, then the desired size-to-order results follow.

Recall from Remark 7.3.1 the various properties of δ. Specifically here, we assume that $d(\overline{S}^+) \leq \delta^2 n$ for some $\delta \to 0$. The proof of Theorem 7.1.2 is given in several stages. First we show that if $d(\overline{S}^+) = o(n)$, then a.a.s. $\overline{S}^+$ does not isolate many vertices in S, where a vertex $v \in S$ is "isolated" by a set U if the neighbourhood of v, $N(v)$, is not contained entirely in U. Then we leverage this to show that a.a.s. a positive fraction of the vertices in $S \backslash S^+$ are not isolated by $\overline{S}$ under the assumptions of Theorem 7.1.2.

Recall that our assumption that $d(S\backslash S^+) = \Theta(M)$ implies that $|S^+| = O(|S\backslash S^+|)$, and thus $|S\backslash S^+| = \Theta(|S|)$. This means that if a positive fraction of the vertices in $S\backslash S^+$ are a.a.s. not isolated by $\overline{S}$, then $G[S]$ a.a.s. contains $\Theta(|S|)$ non-isolated vertices. With this in mind, we do not prove any results about the number of non-isolated vertices in S^+, even though intuitively one could expect that almost none of them are isolated by $\overline{S}$ (and thus almost none of them are isolated vertices in $G[S]$). Altogether this combines to give a proof of Theorem 7.1.2(b).

Remark 7.7.1. We use the following notation throughout this section:

(1) $|S| = n \cdot g(n)$. By definition, $g(n) = O(1)$ and $g(n) = \Omega(\delta^{-1}\sqrt{M}/n)$.

(2) $|S\backslash S^+| = n \cdot h(n)$. By definition, $h(n) = O(1)$ and $h(n) = \Omega(\delta^{-1}\sqrt{M}/n)$.

(3) $d(S\backslash S^+) = aM$, where $a = \Theta(1)$.

(4) $d(\overline{S}\backslash\overline{S}^+) = bM$, where $b = \Theta(1)$.

Lemma 7.7.2. Suppose that $d(\overline{S}^+) \leq \delta^2 n$ for some $\delta \to 0$ and $|S\backslash S^+| = nh(n)$ for some $h(n)$ such that $h(n) = O(1)$ and $h(n) = \omega(n^{-1})$. Then a.a.s. the graph G has at most $\sqrt{\delta}nh(n)$ vertices in $S\backslash S^+$ with all neighbours in $\overline{S}^+$.

If $h(n) = \omega(\delta^{3/2})$, then this result is trivially true: at most $d(\overline{S}^+)$ vertices in G are isolated by $\overline{S}^+$, and $\delta^2 n \leq \sqrt{\delta}n$ for sufficiently large n. Thus, the result follows trivially in this case. However, in the cases where $h(n)$ is small (for example, if $|S\backslash S^+| = O(\delta^2 n)$), then this result is necessary to show that $\overline{S}^+$ does isolate almost all of S.

Proof of Lemma 7.7.2. This proof is a straightforward application of Lemma 2.3.1 and the associated switching framework. Let A_i be the set of $G \in \mathcal{G}(\boldsymbol{d})$ such that $S\backslash S^+$ contains exactly i vertices with all their neighbours in $\overline{S}^+$. That is, A_i contains the graphs G such that exactly i vertices in $S\backslash S^+$ are isolated by $\overline{S}^+$. We define a switching as follows. Suppose $G \in A_{i+1}$. Choose some vertex $v \in S\backslash S^+$ with all its neighbours in $\overline{S}^+$. Choose one of its neighbours u. Choose an ordered pair of vertices (x, y) such that $xy \in E(G)$, $y \notin \overline{S}^+$, edges ux and vy are not present, and all four vertices $\{u, v, x, y\}$ are distinct. Then delete the edges uv and xy and add in edges ux and vy to create a new graph $G' \in \mathcal{G}(\boldsymbol{d})$.

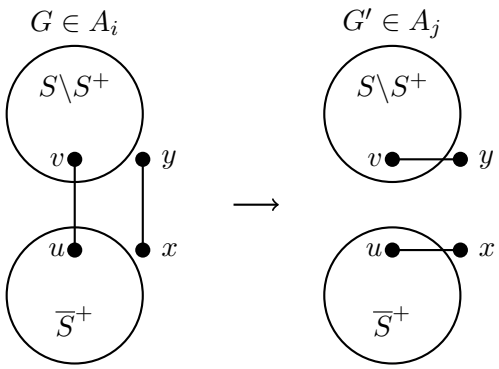

Figure 7.11: A diagram of the switching used in the proof of Lemma 7.7.2, which sends $G \in A_i$ to $G' \in A_j$. The definition of the switching implies that either $j = i - 1$ or $j = i$. The vertex v only has neighbours in $\overline{S}^+$ in G, and $y \notin \overline{S}^+$.

Now we define the structure graph corresponding to this switching. The collection of sets $\{A_i\}_{i=0}^{nh(n)}$ forms a partition of $\mathcal{G}(\boldsymbol{d})$. Then define the (directed) structure graph $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ by

$$\mathfrak{V} := \{0, \ldots, nh(n)\}$$

and

$$\mathfrak{E} := \{(i, j) \mid \exists\, G \in A_i \text{ and } G' \in A_j \text{ and a switching mapping } G \text{ to } G'\}.$$

The sets $\mathfrak{X}$ and $\mathfrak{Y}$ are defined as

$$\mathfrak{Y} := \left\{ i \in \mathfrak{V} \mid i \geq \sqrt{\delta} nh(n) \right\},$$

and

$$\mathfrak{X} := \left\{ i \in \mathfrak{V} \mid i \leq \frac{1}{2}\sqrt{\delta} nh(n) \right\}.$$

We defer the proof of the following claim.

Claim 7.7.3. $\mathfrak{X}$ contains all the sinks of $\mathfrak{G}$, as well as all $A \in \mathfrak{V}$ such that $\alpha(A \to B) \geq 1$ for some $B \in \mathfrak{V}$.

First, we prove the following estimates on $a(i+1)$ and $b(i)$.

Claim 7.7.4. Every $G \in A_{i+1}$ can be subject to at least $(i+1)(n-\delta n)$ switchings. Every $G \in A_i$ can be created by at most $d(\overline{S}^+)(|S \backslash S^+| - i)$ switchings.

Proof of Claim 7.7.4. First we bound from below the number of switchings that can be applied to each $G \in A_{i+1}$. For each $G \in A_{i+1}$, there are $i+1$ choices for v. For each choice of v, there is at least one choice for u, since v has at least one neighbour in G and all neighbours of v are in $\overline{S}^+$. Given $\{u, v\}$, there are $n-2$ choices for a vertex $x \in V(G)$ distinct from u and v. A choice for $\{u, v, x\}$ corresponds to a valid switching if and only if there exists a neighbour y of x such that

(a) $ux, vy \notin E(G)$,

(b) $y \notin \overline{S}^+ \cup \{v\}$,

(c) $\{u, v, x, y\}$ are distinct.

Since G has minimum degree at least 1, for each x there is at least 1 choice for a neighbour y of x. Given u and v, let $W(G, u, v)$ be the number of choices for x such that there does not exist a neighbour y of x such that $\{u, v, x, y\}$ satisfies (a) – (c). As we have done previously, we bound $W(G, u, v)$ from above for all graphs $G \in A_{i+1}$. Given u and v, let $Z(G, u, v)$ be the set of choices for x such that, for every choice of a neighbour y of x, the set $\{u, v, x, y\}$ does not satisfy at least one of the following conditions:

(a′) $ux \notin E(G)$,

(b′) $y \notin \overline{S}^+$,

(c′) $x \notin \overline{S}^+$.

All choices for $\{u, v, x, y\}$ that satisfy (a′) – (c′) also satisfy (a) – (c). To see this, note that if some choice $\{u, v, x, y\}$ satisfies (b′), then we know that $y \notin \overline{S}^+$. This implies that $vy \notin E(G)$, since $y \notin \overline{S}^+$ and v is isolated by $\overline{S}^+$. Condition (a′) implies that $ux \notin E(G)$, and thus the choice of $\{u, v, x, y\}$ satisfies (a). Now suppose the set $\{u, v, x, y\}$ satisfies (c′) too. Then neither x nor y are isolated by $\overline{S}^+$, nor are they in $\overline{S}^+$, and thus it follows that $\{u, v, x, y\}$ are all distinct. This implies that (c) holds, and also that (b) holds, since $y \notin \overline{S}^+$ and $y \neq v$. Therefore, this set $\{u, v, x, y\}$ also satisfies (a) – (c). However, it is possible that $x \in \overline{S}^+$ and $\{u, v, x, y\}$ may still be

a valid choice for a switching. Therefore, it follows that $W(G,u,v) \leq Z(G,u,v)$ for all graphs G. We bound $Z(G,u,v)$ from above for all graphs $G \in A_{i+1}$ and each choice for (u,v).

Since $d(\overline{S}^+) \leq \delta^2 n$ by assumption, there are at most $\delta^2 n$ choices for x such that $ux \in E(G)$. Thus, at most $\delta^2 n$ choices for x do not not satisfy (a′). For case (b′), there are at most $\delta^2 n$ choices for x such that x is isolated by $\overline{S}^+$, and thus at most $\delta^2 n$ choices for x such that there is no neighbour $y \notin \overline{S}^+$. Similarly, $d(\overline{S}^+) \leq \delta^2 n$ implies that $|\overline{S}^+| \leq \delta^2 n$, and thus the number of choices for a vertex x such that $\{u,v,x\}$ does not satisfy (c′) is at most $\delta^2 n$. Thus, there are $n - 2 - O(\delta^2 n)$ choices for a vertex x such that $ux \notin E(G)$ and x is neither isolated by, nor an element of, $\overline{S}^+$. For each of these $n - O(\delta^2 n)$ choices for x, there exists at least one choice for $y \notin \overline{S}^+$ such that $xy \in E(G)$, since x has at least one neighbour in G. This set $\{u,v,x,y\}$ satisfies (a′) – (c′). This implies that $Z(G,u,v) = O(\delta^2 n)$, and thus $W(G,u,v) = O(\delta^2 n)$, for all $G \in A_{i+1}$ and all pairs (u,v). Therefore, for every $G \in A_{i+1}$ there are at least $(i+1)(n - O(\delta^2 n))$ switchings that can be applied to G. Since $\delta \to 0$, for sufficiently large n this is at least $(i+1)(n-\delta n)$.

Now we bound from above the number of switchings that create some $G \in A_i$. There are $d(\overline{S}^+)$ choices for an ordered pair of vertices (u,x) such that $ux \in E(G)$ and $u \in \overline{S}^+$. There are at most $|S \backslash S^+| - i$ choices for a vertex v with exactly one neighbour not in $\overline{S}^+$, and given such a choice of v there is a unique choice for vertex y such that $vy \in E(G)$ and $y \notin \overline{S}^+$. Thus, the number of switchings that create some $G \in A_i$ is at most $d(\overline{S}^+)(|S \backslash S^+| - i)$. □

Thus, we define $a(i+1) := (i+1)(n-\delta n)$ and $b(i) := d(\overline{S}^+)(|S \backslash S^+| - i)$. For all $(i,j) \in \mathfrak{E}$, it follows that $i - j \in \{0,1\}$. If x has exactly one neighbour outside the set $\overline{S}^+$ (that is, vertex y), then $i - j = 0$; otherwise it must be that $i - j = 1$. It also straightforward to see that $b(i+1) \leq b(i)$. Therefore, we define

$$\alpha(i+1 \to B) := \frac{d(\overline{S}^+)(|S \backslash S^+| - i)}{(i+1)(n-\delta n)},$$

which satisfies the requirements of α from Lemma 2.3.1 and the associated framework. With this in mind it is straightforward to prove Claim 7.7.3.

Proof of Claim 7.7.3. Consider $\alpha(i \to B)$ for $i \notin \mathfrak{X}$, that is, $i > \frac{1}{2}\sqrt{\delta} n h(n)$. Then, for all B such that $(i,B) \in \mathfrak{E}$,

$$\alpha(i \to B) \leq \frac{\delta^2 n^2 h(n)}{\frac{1}{2}\sqrt{\delta} n^2 h(n)} \leq \delta.$$

Thus, for all $(i,j) \in \mathfrak{E}$, if $i \notin \mathfrak{X}$ then $\alpha((i,j)) = o(1)$. Furthermore, the lower bound on $a(i)$ implies that i is a sink in $\mathfrak{G}$ if and only if $i = 0$. Thus, Claim 7.7.3 holds. □

Since $\mathfrak{Y}$ and $\mathfrak{X}$ are distinct, the sets $\mathfrak{X}$ and $\mathfrak{Y}$ satisfy the requirements of Lemma 2.3.1.

Recall the definition of $\mathfrak{Q}_{\mathfrak{Y},\mathfrak{Z}}(A,B)$, the set of all non-trivial paths in $\mathfrak{G}$ from set A to set B, where $A, B \subset \mathfrak{V}$, that have no internal vertices in $\mathfrak{Y} \cup \mathfrak{Z}$. As mentioned earlier, all edges in $\mathfrak{E}$ are of the form $(i+1,i)$ or (i,i). Thus, all paths between $\mathfrak{X}$ and $\mathfrak{Y}$ in $\mathfrak{G}$ must be of length at least $\frac{1}{2}\sqrt{\delta} n h(n)$. Furthermore, this implies that the set $\mathfrak{Q}_{\mathfrak{Y},\mathfrak{Z}}(\mathfrak{Y},\mathfrak{Y})$ exclusively contains single-edge paths (Y_1,Y_2) where $Y_1, Y_2 \in \mathfrak{Y}$. Without loss of generality, we can choose δ such that

$\sqrt{\delta}nh(n) \to \infty$. Thus, Lemma 2.3.1 implies that

$$\frac{N(\mathfrak{Y})}{N(\mathfrak{X})} \leq \frac{\delta^{\sqrt{\delta}nh(n)}}{1-\delta} = o(1).$$

Therefore, the probability that a graph $G \in \mathcal{G}(\boldsymbol{d})$ has more than $\sqrt{\delta}h(n)n$ vertices in $S\backslash S^+$ with all neighbours in $\overline{S}^+$ is $o(1)$. This completes the proof. □

For the following lemma and corollary, let $V_0(G)$ be the set of vertices in $S\backslash S^+$ with all neighbours in $\overline{S}$ and at least one neighbour in $\overline{S}\backslash\overline{S}^+$, and let $V_1(G)$ be the set of vertices in $S\backslash S^+$ with exactly 1 neighbour in S and the rest in $\overline{S}$. The following lemma shows that a.a.s. the number of vertices in $S\backslash S^+$ isolated by $\overline{S}$ that are not isolated by $\overline{S}^+$ is at most $(1-\varepsilon)|S\backslash S^+|$ for some $\varepsilon > 0$. Notably, this lemma does not require that $d(\overline{S}^+) = o(n)$, only that it is $o(M)$.

Lemma 7.7.5. Suppose that $|S\backslash S^+| = nh(n)$ for some $h(n)$ such that $h(n) = O(1)$ and $h(n) = \omega(n^{-1})$. Suppose that $d(\overline{S}^+) \leq \delta M$, $d(S\backslash S^+) = aM$, and $d(\overline{S}) = bM$ for some $a, b = \Theta(1)$. Then a.a.s. $|S\backslash(S^+ \cup V_0(G))| = \Theta(|S\backslash S^+|)$.

Proof. Let $\mathcal{A}_k$ be the set of $G \in \mathcal{G}(\boldsymbol{d})$ such that $E(G[S])$ contains k edges uv where either $u \notin S^+$ or $v \notin S^+$ — that is, the set of $G \in \mathcal{G}(\boldsymbol{d})$ such that $G[S]$ contains k edges with at least one endpoint in S^+. These sets form a partition of $\mathcal{G}(\boldsymbol{d})$. We define a switching that maps $G \in \mathcal{A}_k$ to $G' \in \mathcal{A}_k$ (that is, the switching keeps k fixed) to apply Lemma 2.3.1.

The switching is defined as follows: for some graph $G \in \mathcal{A}_k$, choose an edge uv where $v \in V_0(G)$ and $u \in \overline{S}\backslash\overline{S}^+$. Then choose an ordered pair of adjacent vertices (x, y) such that $x \in S\backslash S^+$ and $y \in S$. Then the switching deletes edges uv and xy and replaces them with ux and vy, creating a new graph G'. Such a choice for $\{u, v, x, y\}$ corresponds to a valid switching if and only if $G' \in \mathcal{A}_k$, which occurs if and only if

(a) $vy \notin E(G)$,

(b) $ux \notin E(G)$,

(c) $\{u, v, x, y\}$ are all distinct.

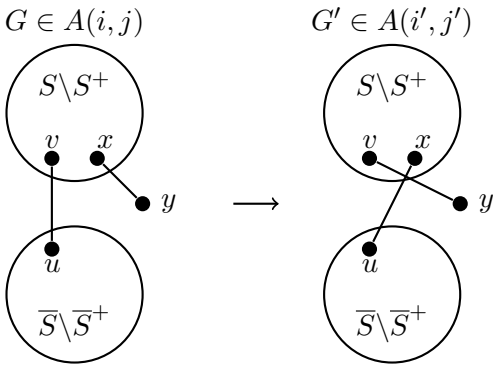

Figure 7.12: A diagram of the switching used in the proof of Lemma 7.7.5, which sends $G \in A(i,j)$ to $G' \in A(i',j')$. The definition of the switching implies that $(i',j') \in \{(i,j),(i-1,j+1),(i-1,j+2)\}$. The vertex v only has neighbours in $\overline{S}$, and at least one neighbour $u \in \overline{S}\backslash\overline{S}^+$.

The condition that $x \in S \backslash S^+$ ensures that the switching does not change k, however, $y \in S^+$ is permissible. We define a structure graph $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ based on this switching. Define

$$A(i,j) := \{G \in \mathcal{A}_k \mid \forall G \in A(i,j),\ |V_0(G)| = i \text{ and } |V_1(G)| = j\}.$$

Then $\{A(i,j)\}_{(i,j)\in[nh(n)]^2}$ forms a partition of $\mathcal{A}_k$. Thus, we define the sets $\mathfrak{V}$ and $\mathfrak{E}$ by

$$\mathfrak{V} := \{0, \ldots, nh(n)\}^2$$

and

$$\mathfrak{E} := \left\{\left\{(i,j),\ (i',j')\right\} \mid \exists\ G \in A(i,j) \text{ and } G' \in A(i',j') \text{ and a switching mapping } G \text{ to } G'\right\}.$$

This switching keeps k fixed, so all outputs of a valid switching are in some set $A(i,j)$ for $(i,j) \in \mathfrak{V}$. This switching maps $G \in A(i,j)$ to some G' in one of $A(i,j)$, $A(i-1,j+1)$, or $A(i-1,j+2)$, depending on the degree of x in $G[S]$. If x has degree 1 in $G[S]$, then $G' \in A(i,j)$. If x has degree 2 in $G[S]$, then $G' \in A(i-1,j+2)$. If x has degree at least 3 in $G[S]$, then $G' \in A(i-1,j+1)$. We prove the following lower and upper bounds on $a((i,j))$ and $b((i,j))$ respectively.

Claim 7.7.6. Every $G \in A(i+1,j)$ can be subject to at least $(i+1)(k-\delta^2 M)$ switchings. Every $G \in A(i,j)$ can be created by at most $d(\overline{S})\,(|S\backslash S^+| - i)$ switchings.

Proof of Claim 7.7.6. First we determine $b((i,j))$. Suppose $G \in A(i,j)$. There are $d(\overline{S}\backslash\overline{S}^+)$ choices for an edge ux such that $u \in \overline{S}\backslash\overline{S}^+$ and $ux \in E(G)$. By definition of $A(i,j)$ there are exactly j choices for vertex $v \in S\backslash S^+$ such that v has degree 1 in $G[S]$; for each v there is a unique choice for vertex y such that $y \in S$ and $vy \in E(G)$. Thus, the number of switchings that map to a given $G \in A(i,j)$ is at most $d(\overline{S}\backslash\overline{S}^+)j$, which is at most $d(\overline{S})\,(|S\backslash S^+| - i)$. Therefore, we set $b((i,j)) := d(\overline{S})\,(|S\backslash S^+| - i)$.

Now we bound from below the number of switchings that can be applied to some $G \in A(i+1,j)$, which determines $a((i+1,j))$. There are $i+1$ choices for the vertex v by definition of $A(i+1,j)$. For each $v \in V_0(G)$, there is at least one choice for u such that $uv \in E(G)$ and $u \in \overline{S}\backslash\overline{S}^+$. By definition of $\mathcal{A}_k$, there are at least k choices for an ordered pair of adjacent vertices (x,y) such that $x \in S\backslash S^+$ and $y \in S$. One of these choices for $\{u,v,x,y\}$ is invalid only if at least one of (a) – (c) is not satisfied.

Let W be the number of choices for $\{u,v,x,y\}$ such that at least one of (a) – (c) is not satisfied. Then the number of switchings that can be applied to a graph $G \in A(i+1,j)$ is at least $(i+1)k - W$. The fact that $v \in V_0(G)$ means that v has no neighbours in S, and thus the number of choices for $\{u,v,x,y\}$ that do not satisfy (a) is 0. Now we consider case (b). To bound the number of choices for vertices where edge ux is present, note that for each choice of uv there are at most $d(u)-1$ choices for x. Given x, there are at most $d(x)-1$ choices for y. Since $u \in \overline{S}\backslash\overline{S}^+$ and $x \in S\backslash S^+$, it follows that for each choice of $\{u,v\}$ there are at most $\delta^2 M$ choices for (x,y) that do not satisfy (b). Finally, we consider case (c). The vertex v is distinct from $\{u,x,y\}$ as $v \in V_0(G)$. The vertex u is distinct from $\{x,y\}$ since $u \in \overline{S}$. Finally, $x \neq y$ because xy is an edge in G. Thus, there are 0 choices for $\{u,v,x,y\}$ that do not satisfy (c). Thus, there are at least $a((i+1,j)) := (i+1)(k-\delta^2 M)$ switchings that can be applied to every $G \in A(i+1,j)$. □

Therefore, we set

$$\alpha((i+1,j)\to B) := \frac{d(\overline{S})\left(|S\backslash S^+|-i\right)}{(i+1)(k-\delta^2 M)}.$$

Lemma 7.2.2 implies that a.a.s. $k \geq 10^{-6}a^2M$. Thus, our analysis focuses on this case. For notational convenience, define $c := 10^{-6}a^2$. Then define the sets $\mathfrak{X}$ and $\mathfrak{Y}$ as follows:

$$\mathfrak{Y} := \left\{(i,j)\in\mathfrak{V} \mid i \geq \frac{b}{\frac{1}{2}c+b}h(n)n \text{ and } j \leq \frac{c}{200}n\right\}$$

and

$$\mathfrak{X} := \left\{(i,j)\in\mathfrak{V} \mid i \leq \frac{b}{\frac{3}{4}c+b}h(n)n \text{ or } j \geq \frac{c}{100}n\right\}.$$

These sets are disjoint. We prove the following claim about $\mathfrak{X}$ and $\mathfrak{Y}$.

Claim 7.7.7. Suppose $k \geq cM$. $\mathfrak{X}$ contains every sink of $\mathfrak{G}$, as well as the tail of every directed edge $e \in \mathfrak{E}$ for which $\alpha(e) \geq 1$.

Proof of Claim 7.7.7. If a vertex (i,j) is a sink, then there are no switchings from any graph in $A(i,j)$ to some other set $A(i',j')$. This means that, for each $G \in A(i,j)$, either there are no vertices $v \in V_0(G)$ (which means $i=0$), or for each choice of uv where $v \in V_0(G)$ there are no valid choices for a pair of adjacent vertices (x,y) such that $x \notin V_1(G)$ and $ux \notin E(G)$. By Claim 7.7.6 this implies that $|V_1(G)| \geq k-\delta^2 M$. For all $k \geq cM$, either of these situations imply that $(i,j)\in\mathfrak{X}$. Furthermore, for every $\alpha((i+1,j)\to B)$ such that $(i+1,j)\notin\mathfrak{X}$ and all $k \geq cM$,

$$\frac{d(\overline{S})\left(|S\backslash S^+|-i\right)}{(i+1)(k-o(M))} < \frac{bM\left(nh(n)-\frac{b}{\frac{3}{4}c+b}nh(n)\right)}{\frac{b}{\frac{3}{4}c+b}nh(n)cM}(1+o(1)) < \frac{3}{4}(1+o(1)) < \frac{4}{5}.$$

□

Thus, for all $\mathcal{A}_k$ such that $k \geq cM$, $\mathfrak{X}$ satisfies all the necessary conditions. Now we consider the sets of paths $\mathfrak{Q}_{\mathfrak{Y},\mathfrak{Z}}(\mathfrak{Y},\mathfrak{X})$ and $\mathfrak{Q}_{\mathfrak{Y},\mathfrak{Z}}(\mathfrak{Y},\mathfrak{Y})$. All paths between some $(i,j)\in\mathfrak{Y}$ and $(i',j')\in\mathfrak{X}$ must be of length $\Omega(nh(n))$. Also note that for all $((i,j),(i',j'))\in\mathfrak{E}$, we know that $i'\leq i$ and $j'\geq j$. This implies that there are no edges $(A,B)\in\mathfrak{E}$ such that $A\notin\mathfrak{Y}$ and $B\in\mathfrak{Y}$. This means that $\mathfrak{Q}_{\mathfrak{Y},\mathfrak{Z}}(\mathfrak{Y},\mathfrak{Y})$ only contains paths of length at most 1. Therefore, Lemma 2.3.1 implies that

$$\frac{N(\mathfrak{Y})}{N(\mathfrak{X})} \leq \frac{\max_{Q\in\mathfrak{Q}_{\mathfrak{Y}\mathfrak{X}}}\alpha(Q)}{1-\max_{Q\in\mathfrak{Q}_{\mathfrak{Y}\mathfrak{Y}}}\alpha(Q)} \leq 5\left(\frac{4}{5}\right)^{\Omega(nh(n))} = 5\left(\frac{4}{5}\right)^{\omega(1)}, \tag{7.22}$$

which is naturally $o(1)$. Define $K = \min\left\{1-\frac{b}{\frac{1}{2}c+b}, \frac{c}{200}\right\}$. Thus, it follows that the probability

that $|S\backslash(S^+ \cup V_0(G))| \geq K|S\backslash S^+|$ is

$$\begin{aligned}\mathbb{P}\left(|S\backslash(S^+ \cup V_0(G))| \geq K|S\backslash S^+|\right) &= \sum_{k\geq 0} \mathbb{P}\left(G \notin \mathfrak{Y} | \mathcal{A}_k\right) \mathbb{P}\left(\mathcal{A}_k\right) \\ &= \sum_{k<cM} \mathbb{P}\left(G \notin \mathfrak{Y} | \mathcal{A}_k\right) \mathbb{P}\left(\mathcal{A}_k\right) + \sum_{k\geq cM} \mathbb{P}\left(G \notin \mathfrak{Y} | \mathcal{A}_k\right) \mathbb{P}\left(\mathcal{A}_k\right) \\ &\geq \sum_{k\geq cM} (1-o(1))\mathbb{P}\left(\mathcal{A}_k\right).\end{aligned}$$

The inequality follows from Equation (7.22). Lemma 7.2.2 implies that

$$\sum_{k\geq cM} \mathbb{P}\left(\mathcal{A}_k\right) = 1 - o(1),$$

and thus $\mathbb{P}\left(|S\backslash(S^+ \cup V_0(G))| = K|S\backslash S^+|\right) = 1 - o(1)$ for some $K = \Theta(1)$. This completes the proof. □

This lemma suggests that problem of finding a linear-order component in $G[S]$ lies in the behaviour of vertices in $\overline{S}^+$. If the number of vertices in S isolated by $\overline{S}^+$ is a.a.s. not all or almost all of $S\backslash S^+$, then the number of isolated vertices in $G[S]$ is also a.a.s. sufficiently small. Intuitively, it is expected that S^+ contains very few isolated vertices. Thus, the important step to determining a threshold for the a.a.s. existence of a vertex-giant component in $G[S]$ appears to boil down to studying how vertices in $\overline{S}^+$ isolate vertices in $S\backslash S^+$.

Corollary 7.7.8. Suppose that $d(S\backslash S^+) = aM$ for some $a = \Theta(1)$, $d(\overline{S}^+) \leq \delta^2 n$, $|S| = ng(n)$ and $|S\backslash S^+| = nh(n)$ for $g(n)$ and $h(n)$ defined earlier. Then a.a.s. the graph $G[S]$ contains $\Theta(|S|)$ non-isolated vertices.

Proof. Lemma 7.7.2 states that G a.a.s. has at most $\sqrt{\delta}nh(n)$ vertices in $S\backslash S^+$ with all neighbours in $\overline{S}^+$. Lemma 7.7.5 states that a.a.s. $|S\backslash(S^+ \cup V_0(G))| \geq K|S\backslash S^+| = Knh(n)$. Therefore, a.a.s. the set $S\backslash S^+$ contains at least $\frac{1}{2}Knh(n)$ vertices that are not isolated in $G[S]$. Since $|S\backslash S^+| = \Theta(|S|)$ by the assumptions of Theorem 7.1.2, this completes the proof. □

This means that in the case where $M = \omega(|S|)$ and $d(\overline{S}^+) \leq \delta^2 n$, we get an easy proof of the a.a.s. existence of a vertex-giant component in $G[S]$. This proves Theorem 7.1.2(b) in the case where $M = \omega(|S|)$. The proof is a straightforward application of Theorem 2.2.15 and the above corollary, in a similar fashion to Claim 7.2.3 and Lemma 7.1.4. Thus, we omit some details.

Corollary 7.7.9. Suppose that $M = \omega(|S|)$ and the conditions of Theorem 7.1.2(b) apply. Then a.a.s. the graph $G[S]$ contains a component with $\Theta(|S|)$ vertices.

Proof sketch. The proof is almost identical to the proof of Lemma 7.1.4, but applying Theorem 2.2.15 instead of Theorem 7.2.1. Let $\boldsymbol{t} = \boldsymbol{t}(G)$ be the degree sequence of $G[S]$, ordered in non-decreasing order with isolated vertices excluded. Then Proposition 2.2.16 implies that there exists an $\varepsilon > 0$ such that a.a.s. $R(\boldsymbol{t}) \geq \varepsilon \widetilde{M}(\boldsymbol{t})$, that is, a.a.s. satisfies the supercritical criterion of Theorem 2.2.15. This means that a.a.s. the graph $G[S]$ has a component containing at least $\gamma n(\boldsymbol{t})$ vertices, where $\gamma > 0$ is a constant and $n(\boldsymbol{t})$ is the number of elements in the sequence $\boldsymbol{t}$ (equivalently, the number of non-isolated vertices in $G[S]$). By Corollary 7.7.8, under the assumptions of

this lemma the graph $G[S]$ a.a.s. has $\Theta(|S|)$ non-isolated vertices. Thus, the largest component of $G[S]$ a.a.s. contains $\Theta(|S|)$ vertices. □

The remaining case is when $M = \Theta(|S|)$, which is only possible if $M = \Theta(n)$ and $|S| = \Theta(n)$. Under these conditions, Theorem 7.1.2(a) and Corollary 7.7.8 together imply that the graph $G[S]$ a.a.s. contains $\Theta(n)$ non-isolated vertices, as well as a component containing $\Theta(M)$ edges. We give a basic argument based on Theorem 2.2.15 that, under these conditions, a.a.s. the graph $G[S]$ contains a component with a linear number of vertices. This will complete the proof of Theorem 7.1.2(b). Let $H(G)$ be the (multi)graph that results from contracting all degree 2 vertices in G. Joos et al. [83] prove the following result in their paper.

Theorem 7.7.10. ([83], Theorem 5) Let $\boldsymbol{d}$, $j_{\boldsymbol{d}}$, $R(\boldsymbol{d})$, and $\widetilde{M}(\boldsymbol{d})$ be as defined in Theorem 2.2.15. For every $\rho > 0$, there exists a $\gamma > 0$ such that for every well-behaved feasible degree sequence $\boldsymbol{d}$, the probability that $\mathcal{G}(\boldsymbol{d})$ has no component of order at least γn and $H(G(\boldsymbol{d}))$ has a component of size at least $\rho\widetilde{M}(\boldsymbol{d})$ is $o(1)$.

Lemma 7.7.11. Suppose that the conditions of Theorem 7.1.2(b) apply, and suppose that $M = \Theta(n)$ and $|S| = \Theta(n)$. Then a.a.s. the graph $G[S]$ contains a component with $\Theta(M)$ edges and $\Theta(n)$ vertices.

Proof. For each $G \in \mathcal{G}(\boldsymbol{d})$, define $\boldsymbol{t} := \boldsymbol{t}(G)$ to be the degree sequence of $G[S]$, ordered in non-decreasing order with isolated vertices excluded. Let $n(\boldsymbol{t})$ be the number of vertices in $G[S]$ with degree at least 1. For some sequence $\boldsymbol{k}$, let $A(\boldsymbol{k})$ be the set of $G \in \mathcal{G}(\boldsymbol{d})$ such that $\boldsymbol{t}(G) = \boldsymbol{k}$. Recall that Proposition 3.1.1 states that, conditional on the event that $G \in A(\boldsymbol{k})$, $G[S]$ is isomorphic to a uniformly random graph with degree sequence $\boldsymbol{k}$ (with the correct number of isolated vertices added back in). Recall the definition of $\widetilde{M}(\boldsymbol{k}) = M(\boldsymbol{k}) - 2n_2(\boldsymbol{k})$ from Theorem 2.2.15, where $n_2(\boldsymbol{k})$ is the number of elements in the sequence $\boldsymbol{k}$ equal to 2. Also recall that a sequence $\boldsymbol{k}$ is well-behaved if $\widetilde{M}(\boldsymbol{k}) = \omega(1)$. Partition $\mathcal{G}(\boldsymbol{d})$ into two parts:

$$Z := \left\{ G \in \mathcal{G}(\boldsymbol{d}) \ : \ \widetilde{M}(\boldsymbol{t}(G)) < \sqrt{M} \right\},$$
$$\mathcal{G}(\boldsymbol{d}) \backslash Z := \left\{ G \in \mathcal{G}(\boldsymbol{d}) \ : \ \widetilde{M}(\boldsymbol{t}(G)) \geq \sqrt{M} \right\}.$$

Suppose there exists some $G \in Z$ such that $G[S]$ has a component with at least cM edges for some constant $c > 0$. Since $G \in Z$, this means that $\widetilde{M}(\boldsymbol{t}) < \sqrt{M}$. Thus, a component containing cM edges must also contain at least $\frac{1}{2}(2cM - \widetilde{M}(\boldsymbol{t}))$ vertices of degree 2, and therefore must have $\Theta(n)$ vertices overall. Naturally, a component with $o(M)$ edges must also have $o(n)$ vertices, since $M = \Theta(n)$ by the assumptions of the lemma. Theorem 7.1.2(a) implies that we can choose $\rho > 0$ such that $G[S]$ a.a.s. has a component with at least ρM edges. Define a partition (Z_1, Z_2) of Z, where Z_1 contains the graphs $G \in Z$ such that $G[S]$ has a component with at least ρM edges for the above choice of ρ, and $Z_2 = Z \backslash Z_1$. Thus, $\mathbb{P}(Z_2) = o(1)$.

Let $L_1(G)$ be the number of edges in the largest component of G. The above choice of ρ also implies that either $\mathbb{P}(L_1(G[S]) \geq \rho M | \mathcal{G}(\boldsymbol{d}) \backslash Z) = 1 - o(1)$ or $\mathbb{P}(\mathcal{G}(\boldsymbol{d}) \backslash Z) = o(1)$. For each $\boldsymbol{k}$ such that $\widetilde{M}(\boldsymbol{k}) \geq \sqrt{M}$, define $A(\boldsymbol{k}) \subset \mathcal{G}(\boldsymbol{d})$ to be the set of graphs G such that $\boldsymbol{t}(G) = \boldsymbol{k}$. The sets $A(\boldsymbol{k})$ form a partition of $\mathcal{G}(\boldsymbol{d}) \backslash Z$. The definition of Z implies that every such $\boldsymbol{k}$ is well-behaved,

since $\sqrt{M} = \omega(1)$. Partition each $A(\boldsymbol{k})$ into three parts:

$$X(\boldsymbol{k}) := \left\{ G \in A(\boldsymbol{k}) : L_1(H(G[S])) \geq \frac{1}{2}\rho\widetilde{M}(\boldsymbol{k}) \right\},$$
$$Y_1(\boldsymbol{k}) := \left\{ G \in A(\boldsymbol{k}) : L_1(G[S]) \geq \rho M,\ L_1(H(G[S])) < \frac{1}{2}\rho\widetilde{M}(\boldsymbol{k}) \right\},$$
$$Y_2(\boldsymbol{k}) := \left\{ G \in A(\boldsymbol{k}) : L_1(G[S]) < \rho M,\ L_1(H(G[S])) < \frac{1}{2}\rho\widetilde{M}(\boldsymbol{k}) \right\}.$$

We know that a.a.s. $G[S]$ has a component with at least ρM edges. Thus, $\sum_{\boldsymbol{k}} \mathbb{P}(Y_2(\boldsymbol{k})) = o(1)$, where the sum is over all sequences $\boldsymbol{k}$ such that $\boldsymbol{t}(G) = \boldsymbol{k}$ for some $G \in \mathcal{G}(\boldsymbol{d})\backslash Z$. Now consider some $G \in Y_1(\boldsymbol{k})$ for an arbitrary $\boldsymbol{k}$. The extra edges in the component in $G[S]$ must have come from subdividing edges of the corresponding component in $H(G[S])$. Each subdivision increases the number of edges in the component by exactly 1, as well as increasing the number of vertices by exactly 1. Therefore, by a similar reasoning to the case where $G \in Z$, for all $G \in Y_1$ we know that all components in $G[S]$ with ρM edges must also contain at least $\frac{1}{2}\rho n$ vertices. This implies that for all graphs $G \in Z_1 \cup \bigcup_{\boldsymbol{k}} Y_1(\boldsymbol{k})$, there exists some constant $\alpha > 0$ such that $G[S]$ has a component with at least $\alpha|S|$ vertices. Finally, Theorem 7.7.10 implies that there exists some constant $\alpha' > 0$ such that for all choices of $\boldsymbol{t}$,

$$\mathbb{P}(G[S] \text{ has a component with at least } \alpha' n(\boldsymbol{t}) \text{ vertices} \,|\, X(\boldsymbol{t})) = 1 - o(1).$$

Thus, by the law of total probability, a.a.s. the graph $G[S]$ contains a giant component on a positive fraction of the non-isolated vertices. Now recall that $d(\overline{S}^+) \leq \delta^2 n$. Thus, Corollary 7.7.8 applies, which implies that $G[S]$ a.a.s. has $\Theta(|S|)$ non-isolated vertices, which is equivalent to the statement that a.a.s. $n(\boldsymbol{t}) = \Theta(|S|)$. Therefore, we can choose $\alpha, \alpha' > 0$ such that a.a.s. $G[S]$ contains a component with at least $\alpha|S|$ vertices. Since $|S| = \Theta(n)$, this component has $\Theta(n)$ vertices. This component must contain at least $\Theta(n)$ edges too, as it is connected. Since $M = \Theta(n)$, this completes the proof. □

This completes the proof of Theorem 7.1.2(b), except for the following caveat. The proof of Theorem 2.2.15 shows that in the case where a degree sequence $\boldsymbol{d}$ satisfies the supercritical criterion and $\widetilde{M} \leq n \log\log n$, a uniformly random graph $G(\boldsymbol{d})$ a.a.s. contains a component with both $\Theta(\widetilde{M})$ edges and $\Theta(n)$ vertices too. This naturally carries over when we apply their result. However, the proof does not actually show that the linear-size and linear-order components are a.a.s. the same component when $\widetilde{M} \geq n \log\log n$. Of course, this is straightforward to show, since intuitively the probability that $\mathcal{G}(\boldsymbol{d})$ contains a component with $\Theta(\widetilde{M})$ edges and $o(n)$ vertices as well as a different component with $\Theta(n)$ vertices is minuscule. This is the content of the next lemma. It is straightforward but included for completeness.

Lemma 7.7.12. Let $\boldsymbol{d}$ be a graphical sequence satisfying the supercritical case of Theorem 2.2.15 such that $M = \omega(n)$. Then if $G \sim \mathcal{G}(\boldsymbol{d})$ contains a component with $\Theta(M)$ edges and a component with $\Theta(n)$ vertices, a.a.s. they are the same component.

Proof. Note that if $M = \omega(n)$, then $\widetilde{M} = M(1 - o(1))$. Choose constants $\gamma, \alpha > 0$ such that Theorems 2.2.15 and 7.2.1 imply that $G(\boldsymbol{d})$ a.a.s. has a component with at least γM edges and a component with at least αn vertices. Let A be the subset of $\mathcal{G}(\boldsymbol{d})$ such that, for all $G \in A$,

the graph G contains a component with at least γM edges and αn vertices. Let B be the set of $G \in \mathcal{G}(\boldsymbol{d})$ such that G contains a component K_1 with at least γM edges but less than εn vertices for a sufficiently small constant $\varepsilon < \alpha$, as well as a different component K_2 with at least αn vertices. We define a switching between graphs in A and graphs in B. Suppose $G \in B$. This switching takes an ordered pair of vertices (a, b) in the component K_1 such that $ab \in E(G)$ and ab is not a bridge, as well as a pair of adjacent vertices (c, d) in $V(K_2)$. The switching then deletes the edges ab and cd and replaces them with ac and bd.

Lemma 11 of Joos et al. [83] states that for all $G \in B$ there are at most $8n^2$ switchings that map some $G' \in A$ to G. Now we bound from below the number of valid switchings that can be applied to some $G \in B$. There are at least $2\gamma M$ choices for an ordered pair of adjacent vertices (a, b) such that $a, b \in V(K_1)$, since K_1 contains at least γM edges. We call this choice valid if ab is not a bridge. For all $\varepsilon > 0$, there are at most εn bridges in K_1, since it contains less than εn vertices by the definition of B. Since $M = \omega(n)$, this implies that there are at least γM valid choices for (a, b). Furthermore, there are at least αn choices for a pair of adjacent vertices (c, d) in the component K_2, since component K_2 contains at least αn vertices, and thus at least as many edges. For all valid choices of (a, b), every choice for (c, d) corresponds to a valid switching from B to A. Thus,

$$\frac{|B|}{|A| + |B|} = \frac{\Theta(n^2)}{\Theta(nM)}.$$

Since $M = \omega(n)$, this implies that $|B| = o(|A|)$.

Let C be the set of $G \in \mathcal{G}(\boldsymbol{d})$ such that G contains either no component with at least αn vertices or no component with at least γM edges. Then the sets (A, B, C) form a partition of $\mathcal{G}(\boldsymbol{d})$. Our choices of γ and α imply that $\mathbb{P}(C) = o(1)$. Thus,

$$\mathbb{P}(A) = \frac{|A|}{|A| + |B| + |C|} = \frac{1}{1 + o(1)} = 1 - o(1).$$

Therefore, the components in G containing $\Theta(n)$ vertices and $\Theta(M)$ edges respectively are a.a.s. the same component. □

Remark 7.7.13. This proof can be easily adapted to show that, under these same conditions, the giant component in $G(\boldsymbol{d})$ a.a.s. has order $n - o(n)$. The proof is very similar to that given above, which is itself very similar to the proof of Lemma 10 in [83]. Since $M = \omega(n)$, for all $G \in B$ there are $\Theta(M)$ valid choices for (a, b). Recall that $M = nf(n)$ for some $f(n) = \omega(1)$. If $|V(K_2)| \geq n/\sqrt{f(n)}$, then it still follows that $|B| = \omega(|A|)$, and the rest of the proof follows similarly.

Thus, if $d(\overline{S}^+) = o(n)$, $d(S \backslash S^+) = \Theta(M)$, and $R_S(\boldsymbol{d}) = \Theta(M)$, then $G[S]$ a.a.s. contains a component with a linear number of edges and vertices. This completes the proof of Theorem 7.1.2(b).

7.8 Sequences and subsets that inform our restrictions

If $(\boldsymbol{d}, S)$ is a pair such that $d(\overline{S}^+) = \Omega(n)$ (in particular, if we are also assuming that $d(\overline{S}^+) = o(M)$, then this forces $M = \omega(n)$) then the situation is more complicated. In this case, it is

entirely possible that $G[S]$ consists mostly of isolated vertices. Lemma 7.7.5 shows that if $\overline{S}^+$ does not a.a.s. isolate almost all vertices in $S \backslash S^+$, then a.a.s. the number of vertices in $S \backslash S^+$ not isolated by $\overline{S}$ is at least $\varepsilon |S \backslash S^+|$ for some $\varepsilon > 0$. So the culprit is vertices in $\overline{S}^+$. We conjecture that the culprit is specifically vertices of degree $\Theta(n)$ in $\overline{S}^+$. If this is the case, then one could possibly prove the following strengthening of Theorem 7.1.2(b).

Conjecture 7.8.1. Suppose $(\boldsymbol{d}, S)$ satisfies the conditions of Theorem 7.1.2 with $d(\overline{S}^+) = o(M)$ and $\overline{S}^+$ has maximum degree δn for some $\delta \to 0$. Then a.a.s. $G[S]$ contains a component with $\Theta(|S|)$ vertices.

In this section we given some examples of $(\boldsymbol{d}, S)$ pairs that help inform our intuition and our restrictions on $\boldsymbol{d}$ and S in Theorem 7.1.2. Notably, these examples all have a common thread: in each case, $\overline{S}$ contains vertices of degree $n-1$. With this in mind, we also put forward a strengthening of the above conjecture.

Conjecture 7.8.2. Suppose $(\boldsymbol{d}, S)$ satisfies the conditions of Theorem 7.1.2 and $\overline{S}^+$ has maximum degree cn for some constant $c < 1$. Then a.a.s. $G[S]$ contains a component with $\Theta(|S|)$ vertices.

However, the restrictions that we impose on $\overline{S}^+$ cannot be totally removed. Here we give some examples of pairs $(\boldsymbol{d}, S)$ which suggest that some conditions are needed. The first example is a pair $(\boldsymbol{d}, S)$ which all of the conditions of Theorem 7.1.2(a) except the condition that $d(S \backslash S^+) = \Theta(M)$. In this case, $G[S]$ never contains a component with a linear number of edges or vertices.

Example 7.8.3. Suppose S contains $n^{1/3}$ vertices of degree $n^{3/4}$ and $\frac{1}{2}n - n^{1/3}$ vertices of degree $n^{1/50}$. Then suppose $\overline{S}$ contains $n^{1/50}$ vertices of degree $n-1$ and $\frac{1}{2}n - n^{1/50}$ vertices of degree $n^{1/12}$ (enough to make the total number of vertices exactly n). We know that $M = \Theta(n^{13/12}) = \omega(n)$, and $d(\overline{S}^+) \sim n^{51/50} = o(M)$, but we also know that $d(S \backslash S^+) = \Theta(n^{51/50}) = o(M)$. Both $d(S)$ and $d(\overline{S})$ are $\Theta(M)$, as is $d(S^+)$. Since $M = \omega(n)$, one might think that, in a similar vein to Lemma 7.1.4, the graph $G[S]$ should a.a.s. contain a component with $\Theta(M)$ edges. But in every graph $G \in \mathcal{G}(\boldsymbol{d})$, each vertex in S with degree $n^{1/50}$ is adjacent to every vertex $\overline{S}^+$. Thus, $G[S \backslash S^+]$ is always the empty graph, and there are also no edges between S^+ and $S \backslash S^+$. So the only non-singular components in $G[S]$ contain exclusively vertices in S^+. Since $|S^+| = n^{1/3}$, there are at most $n^{2/3}$ edges between all of these vertices. Thus, every component in $G[S]$ always has $O(n^{2/3})$ edges.

In the next example, we give a pair $(\boldsymbol{d}, S)$ for which there is no linear order component, despite the fact that $d(S \backslash S^+) = \Theta(M)$. This satisfies all of the conditions of Theorem 7.1.2(a) (and thus a.a.s. $G[S]$ have a component with $\Theta(M)$ edges) but the induced subgraph $G[S]$ never has a component with $\Theta(|S|)$ vertices.

Example 7.8.4. Suppose S contains $n^{3/4}$ vertices of degree $n^{1/2}$ and $\frac{1}{2}n - n^{3/4}$ vertices of degree $n^{1/20}$. Then suppose $\overline{S}$ contains $n^{1/20}$ vertices of degree $n-1$ and $\sim \frac{1}{2}n$ vertices of order $n^{1/4}$ (enough to make the total number of vertices exactly n). We know that $M = \frac{3}{2}n^{5/4}(1+o(1))$ and $d(\overline{S}^+) = n^{21/20} = o(M)$, and it is straightforward to show that $R_S = \Theta(M)$. Also, since $\sqrt{M} \geq n^{5/8}$, we can choose δ large enough such that $\delta\sqrt{M} = \omega(n^{1/2})$ and thus $S^+ = \emptyset$. This implies that $d(S \backslash S^+) \sim n^{5/4}$. Therefore, Theorem 7.1.2 implies that $G[S]$ a.a.s. contains a component of size $\Theta(M)$. However, all the degree $n^{1/20}$ vertices in S are always adjacent to every vertex with degree

$n-1$ in $\overline{S}$, so the induced graph $G[S]$ has at most $n^{3/4}(1+o(1))$ non-isolated vertices. Thus, $G[S]$ a.a.s. contains a component with $\Theta(M)$ edges, but it cannot possibly contain a component with $\Theta(|S|)$ vertices.

Chapter 8

Final remarks

8.1 Comparison of the two results for giant components

In this thesis we give two different thresholds for the existence of a giant component in $G[S]$. The threshold given in Theorems 7.1.1 and 7.1.2 is a stronger characterisation, as it works for a wider range of degree sequences than Theorem 4.4.2. This should not be surprising, as the exploration method is a more direct approach to studying this problem. Theorem 4.4.2 follows from an application of a very general result about the degree sequence of $G[S]$, which has the advantage of also giving results about other graph properties. On the other hand, Theorems 7.1.1 and 7.1.2 are proved using methods much more specific to analysing giant components, and thus give better results in that case but do not generalise to proving results about other properties of $G[S]$. In particular, allowing vertices of high degree in $\overline{S}$ created some situations where $G[S]$ a.a.s. did not contain a vertex-giant component, even if it contained a component with cM edges for some $c > 0$. This is not captured in Theorem 4.4.2, as the assumptions of the reduction method (specifically the assumption on the maximum degree of $\boldsymbol{d}$) imply that $G[S]$ a.a.s. contains $\Theta(|S|)$ non-isolated vertices.

Consider an n-element graphical sequence $\boldsymbol{d}$ and set $S \subset [n]$ such that $d(S),\ d(\overline{S}) = \Theta(M)$ and $\Delta(\boldsymbol{d}) \leq \sqrt{M}/\log^7 M$. That is, Theorem 4.4.2 applies, and one of Theorems 7.1.1 and 7.1.2 applies. Naturally, these characterisations must agree: if $(\boldsymbol{d}, S)$ satisfies the conditions of Theorem 7.1.1, then Theorem 4.4.2 should also imply that $G[S]$ a.a.s. does not contain a giant component, and vice versa. However, these conditions are extremely different, and quite hard to compare directly. For the more basic cases, it is easy to show that they agree. For example, consider the case that $M = \omega(|S|)$. Recall the definition of $\boldsymbol{d}_I$ from Definition 3.1.2, and recall that $\boldsymbol{d}'_I$ is defined to be the sequence $\boldsymbol{d}_I$ ordered in non-decreasing order. First note that if $M = \omega(|S|)$, then Theorem 7.1.2 applies and states that $G[S]$ a.a.s. contains a giant component (both a linear number of edges and vertices). On the other hand, Lemma 3.3.1 states that a.a.s. $M(\boldsymbol{d}_S) \sim M(\boldsymbol{d}_I) \sim \gamma^2 M$ (where $\gamma = d(S)/M$) and thus a.a.s. $G[S]$ has total degree $\omega(|S|)$. Therefore, Theorem 4.4.2 also implies that $G[S]$ a.a.s. contains a giant component in this case. In fact, if $M \geq 4|S|/\gamma^2$, then Claim 2.2.17 and Theorem 4.4.2 together imply that $G[S]$ a.a.s. contains a giant component. The case where $M = \Theta(|S|)$, the path is less clear, but this would be a natural direction for future research.

One noteworthy observation about both thresholds is that, in either case, it does not particularly matter which vertices of G are deleted; the only thing that matters is the total degree of the deleted vertices. Theorems 7.1.1 and 7.1.2 implies that this intuition extends to the case where

$d(\overline{S}^+)$ is small, but the examples given at the end of the previous section imply that this is not true if $\overline{S}$ contains vertices of extremely high degree. One obvious and natural direction for future work would be to extend the results given in Lemma 3.1.5 (and thus the following results given in Theorem 4.4.2), as well as the thresholds given in Theorems 7.1.1 and 7.1.2, to the case where $d(\overline{S}^+) \geq cM$ for some constant $c > 0$. This would most likely be quite difficult, as analysing switchings in the presence of many vertices of high degree would require some finesse.

8.2 Extensions of the reduction method

The first obvious direction to extend this work would be to combine what we know about $\boldsymbol{d}_S$ with other results about $\mathcal{G}(\boldsymbol{d})$. That way one could determine more properties of induced subgraphs of random graphs with given degree sequences similarly to what was done in Chapter 4. Along with their results about the connectivity and chromatic number of $\mathcal{G}(\boldsymbol{d})$, Gao and Ohapkin [56] provide a general framework for adapting results about the configuration model $\mathcal{C}(\boldsymbol{d})$ to the model $\mathcal{G}(\boldsymbol{d})$, for certain sequences $\boldsymbol{d}$. In combination with Lemma 3.1.5, this could provide a pipeline for proving results about $G[S]$ from known results about $\mathcal{C}(\boldsymbol{d})$. This could give rise to many more results about $G[S]$, as the configuration model has been well studied for many years and is significantly easier than the random graph model.

Another natural direction would be to relax the restriction that $d(S),\ d(\overline{S}) = \Theta(M)$. In particular, the case where $d(S) = o(M)$ would be interesting to study. One important thing to note in this generalisation is that the cutoff between what is considered a "big" vertex and what is considered a "small" vertex might need to change. In particular, the cutoff may need to increase as a function of $d(S)/M$. This means that there are more vertices in S_{small} with a wider range of possible degrees. Using the current methods, this would likely require a stricter maximum degree requirement on $\boldsymbol{d}$. A generalisation in this direction would be useful though, particularly for studying percolation: if the case where $d(S)/M = o(M)$ can be analysed, then we can study the site-percolated $\mathcal{G}(\boldsymbol{d})$ with survival probability p in the case that $p \to 0$. Percolation results in the model $\mathcal{G}(\boldsymbol{d})$ in this case would be novel and interesting to study, as the behaviour could be quite different to the case where p is constant.

As well as this, the results given in Lemma 3.1.5 could be extended to increase the maximum degree of G that is allowed, which would allow applications of these results for a wider range of random graphs. This would be quite challenging, and most likely would require a detailed and careful use of the switching method, as all of the switching errors in question were of order $\Omega(\Delta^2/M)$. If one was to go generalise this and allow for vertices with significantly higher degree than the current conditions, it is likely that entirely new methods would be needed. It is natural to expect that the distribution of the induced degree of a high-degree vertex would be quite concentrated, but proving this would be challenging as straightforward switchings tend to fail when the graph contains many high-degree vertices. Thus, extensions in this direction would most likely require novel methods that could be used to study many other problems concerning $\mathcal{G}(\boldsymbol{d})$.

8.3 Extensions of the exploration method

A natural direction to take this work would be to extend to the case where $\overline{S}^+$ has large total degree, that is $d(\overline{S}^+) \geq cn$ or $d(\overline{S}^+) \geq cM$ for some $c > 0$. There are some clear paths for small progress in this direction. For example, in the case where $M = \Theta(n)$, if $d(\overline{S}^+) \geq cM$ and $d(\overline{S}^+) \leq (1-\varepsilon)d(S^+)$ for constants $c,\ \varepsilon > 0$, then Lemma 7.6.11 should still hold, as the proof should follow with minor modifications. Similarly, if $d(\overline{S}^+) \geq cM$ and $d(\overline{S}^+) \leq (1-\varepsilon)d(S\backslash S^+)$, then a slight modification to the proof of Lemma 7.2.2 should imply that a.a.s. $G[S]$ still contains $\Theta(M)$ edges with at least one end in $S\backslash S^+$.

Beyond easy modifications like this, results in this area would most likely be quite dependent on the particular degrees of the vertices in $\overline{S}^+$, rather than just the total degree of the set. The intuitive idea here is that, for example, the neighbourhood of a degree $n-1$ vertex is simply the whole graph (except itself), but the neighbourhoods of two vertices with degree $\frac{n-1}{2}$ are likely to have significant overlap. Thus, results in this direction would be quite distinct from the results given in this thesis, which imply that the individual degrees of the deleted vertices are not particularly relevant. In this vein, it would be interesting to see whether Conjectures 7.8.1 and 7.8.2 are true or not, even just in the case when $d(\overline{S}^+) = o(M)$. Lemma 7.7.5 implies that if one can show that deleting the vertices in $\overline{S}^+$ a.a.s. does not create too many isolated vertices, then the graph $G[S]$ a.a.s. does not have too many isolated vertices. Therefore, progress towards these conjectures would require more involved switching arguments to study the neighbourhoods of vertices in $\overline{S}^+$, and neighbourhoods of high-degree vertices in general.

Applying the results in Chapter 7 to the percolation problem would be another direction for future research. The results in this chapter allow for a wider range of degrees than the results about percolation in Chapter 5, and in particular allow the presence of vertices of very high degree. However, the condition that $d(\overline{S}^+) = o(M)$ means that the result does not extend well to studying percolation on $\mathcal{G}(\boldsymbol{d})$: each high-degree vertex has a non-trivial probability of being deleted, which means that either their total degree must be small or the event that $\overline{S}^+$ has large total degree would be likely. This would still be a generalisation, as Theorem 5.3.3 does not apply if G contains even one such vertex. However, results in this direction would be much more general if combined with results in the case where $\overline{S}^+$ has large total degree.

Another natural extension of the exploration method, similar to what was mentioned earlier for the reduction method, would be to extend this case to the case where $d(S) = o(M)$. This would be quite challenging, particularly if the condition that $d(\overline{S}^+) = o(M)$ is also relaxed, as it might be possible that every vertex in S a.a.s. has many neighbours in $\overline{S}^+$. For both this extension and the case where $d(\overline{S}^+) \geq cM$, Lemma 7.5.5 would need to be refined to a much more precise estimate. In the current proof, we show that the contribution to $\mathbb{E}_{t-1}[d_S(w_t) - 2]$ from vertices in $\overline{S}^+$ throughout the process is minimal and can be effectively ignored. We also show that back edges do not eat up a significant portion of the open edges gained at each step. However, under either of these generalisations, neither of these claims are clear. As an example, suppose one were to analyse the exploration process on a random induced subgraph $G[S]$ where $(\boldsymbol{d}, S)$ are given in Example 7.0.1. In this case, the contribution of back edges from $\overline{S}^+$ is incredibly vital to the analysis.

One different direction would be to analyse the process much more precisely to get better upper and lower bounds on the sizes of the components in $G[S]$. This is analogous to an idea mentioned

by Joos et al. [83] as an extension of their result. In our proofs, we made no real effort to optimise the constants that we used, instead simply giving results about the existence and non-existence of giant components. This is also true of the result by Joos et al. [83] about giant components in $\mathcal{G}(\boldsymbol{d})$. Along with similar results about the specific size and order of giant components in $\mathcal{G}(\boldsymbol{d})$, it would be interesting to see exactly how the size and order of the giant component are affected by deleting the vertices in $\overline{S}$. In their paper, Joos et al. claim that intuitively the largest component in a random graph with degree sequence $\boldsymbol{d}$ has size approximately $R(\boldsymbol{d})$. By similar intuitive reasoning, it is possible that the size of the largest component in $G[S]$ is approximately $R_S(\boldsymbol{d})$. However, this could only be true under certain conditions, as the threshold itself does not apply for all pairs $(\boldsymbol{d}, S)$. It is likely that, as with the current threshold, there would need to be conditions on $\overline{S}^+$ in order to determine its asymptotic size.

Bibliography

[1] Achlioptas, D. and Naor, A. (2004). The two possible values of the chromatic number of a random graph. In *Proceedings of the Thirty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '04, pages 587–593, New York, NY, USA. Association for Computing Machinery.

[2] Ajtai, M., Komlos, J., and Szemerédi, E. (1982). Largest random component of a k-cube. *Combinatorica*, 2:1–7.

[3] Alon, N., Benjamini, I., and Stacey, A. (2004). Percolation on finite graphs and isoperimetric inequalities. *The Annals of Probability*, 32(3):1727 – 1745.

[4] Alon, N. and Krivelevich, M. (1997). The concentration of the chromatic number of random graphs. *Combinatorica*, 17:303–313.

[5] Appel, K. and Haken, W. (1977). Every planar map is four colorable. Part I: Discharging. *Illinois Journal of Mathematics*, 21(3):429 – 490.

[6] Appel, K., Haken, W., and Koch, J. (1977). Every planar map is four colorable. Part II: Reducibility. *Illinois Journal of Mathematics*, 21(3):491 – 567.

[7] Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

[8] Bates, S., Angeon, V., and Ainouche, A. (2014). The pentagon of vulnerability and resilience: A methodological proposal in development economics by using graph theory. *Economic Modelling*, 42:445–453.

[9] Blanchet, J. and Stauffer, A. (2013). Characterizing optimal sampling of binary contingency tables via the configuration model. *Random Structures & Algorithms*, 42(2):159–184.

[10] Bollobás, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Eur. J. Comb.*, 1:311–316.

[11] Bollobás, B. (1981a). Degree sequences of random graphs. *Discrete Mathematics*, 33(1):1–19.

[12] Bollobás, B. (1981b). Threshold functions for small subgraphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 90(2):197206.

[13] Bollobás, B. (1982a). The asymptotic number of unlabelled regular graphs. *Journal of the London Mathematical Society*, s2-26(2):201–206.

[14] Bollobás, B. (1982b). Vertices of given degree in a random graph. *Journal of Graph Theory*, 6(2):147–155.

[15] Bollobás, B. (1984). The evolution of random graphs. *Transactions of the American Mathematical Society*, 286(1):257–274.

[16] Bollobás, B. (1985). *Random graphs*. Academic Press.

[17] Bollobás, B. (1988). The chromatic number of random graphs. *Combinatorica*, pages 49–55.

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

[19] Bollobás, B., Borgs, C., Chayes, J., and Riordan, O. (2010). Percolation on dense graph sequences. *The Annals of Probability*, 38(1):150 – 183.

[20] Bollobás, B., Kohayakawa, Y., and Łuczak, T. (1992). The evolution of random subgraphs of the cube. *Random Structures & Algorithms*, 3(1):55–90.

[21] Bollobás, B. and Riordan, O. (2015). An old approach to the giant component problem. *Journal of Combinatorial Theory, Series B*, 113:236–260.

[22] Borgs, C., Chayes, J. T., van der Hofstad, R., Slade, G., and Spencer, J. (2006). Random subgraphs of finite graphs: iii. the phase transition for the n-cube. *Combinatorica*, 26(4):395–410.

[23] Brick, L., Gao, P., and Southwell, A. (2020). The threshold of symmetry in random graphs with specified degree sequences.

[24] Britton, T., Deijfen, M., and Martin-Löf, A. (2006). Generating simple random graphs with prescribed degree distribution. *Journal of Statistical Physics*, 124(6):1377–1397.

[25] Britton, T., Janson, S., and Martin-Löf, A. (2007). Graphs with specified degree distributions, simple epidemics, and local vaccination strategies. *Advances in Applied Probability*, 39(4):922–948.

[26] Broadbent, S. R. and Hammersley, J. M. (1957). Percolation processes: I. crystals and mazes. *Mathematical Proceedings of the Cambridge Philosophical Society*, 53(3):629–641.

[27] Brunk, N. E. and Twarock, R. (2021). Percolation theory reveals biophysical properties of virus-like particles. *ACS Nano*, 15(8):12988–12995.

[28] Burtin, Y. D. (1977). On connection probability of a random subgraph of an n-dimensional cube. *Probl. Peredachi Inf.*, 13(2):90–95.

[29] Chaitin, G. J. (1982). Register allocation & spilling via graph coloring. *SIGPLAN Not.*, 17(6):98–101.

[30] Chen, Z. and Ji, H. (2010). Graph-based clustering for computational linguistics: A survey. In *Proceedings of the 2010 Workshop on Graph-Based Methods for Natural Language Processing*, TextGraphs-5, pages 1–9, USA. Association for Computational Linguistics.

[31] Chodrow, P. S. (2020). Configuration models of random hypergraphs. *Journal of Complex Networks*, 8(3):1 – 26.

[32] Chung, F. R. K. and Lu, L. (2002). Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics*, 6:125–145.

[33] Clauset, A., Shalizi, C., and Newman, M. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703.

[34] Coja-Oghlan, A., Panagiotou, K., and Steger, A. (2008). On the chromatic number of random graphs. *Journal of Combinatorial Theory, Series B*, 98(5):980–993.

[35] Cooper, C., Frieze, A., and Reed, B. (2002a). Random regular graphs of non-constant degree: Connectivity and hamiltonicity. *Combinatorics, Probability and Computing*, 11(3):249–261.

[36] Cooper, C., Frieze, A., Reed, B., and Riordan, O. (2002b). Random regular graphs of non-constant degree: Independence and chromatic number. *Combinatorics, Probability and Computing*, 11(4):323–341.

[37] Deo, N. (1974). *Graph theory with applications to engineering and computer science.* Prentice-Hall series in automatic computation. Prentice-Hall, Englewood Cliffs, N.J.

[38] Dereich, S. and Mörters, P. (2013). Random networks with sublinear preferential attachment: The giant component. *The Annals of Probability*, 41(1):329 – 384.

[39] Diestel, R. (2017). *Graph Theory, Fifth Edition.* Graduate Texts in Mathematics. Springer, Berlin, Heidelberg.

[40] Erdős, P. and Gallai, T. (1960). Gráfok előírt fokú pontokkal. *Matematikai Lapok*, 11:264–274.

[41] Erdős, P. and Rényi, A. (1959). On random graphs i. *Publicationes Mathematicae Debrecen*, 6:290–297.

[42] Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. In *Publication of the Mathematical Institute of the Hungaruian Academy of Sciences*, pages 17–61.

[43] Erdős, P. and Rényi, A. (1963). Asymmetric graphs. *Acta Mathematica Academiae Scientiarum Hungaricae*, 14:295–315.

[44] Erdős, P. and Spencer, J. (1979). Evolution of the n-cube. *Computers & Mathematics with Applications*, 5(1):33–39.

[45] Federico, L. and van der Hofstad, R. (2017). Critical window for connectivity in the configuration model. *Combinatorics, Probability and Computing*, 26(5):660–680.

[46] Fountoulakis, N. (2007). Percolation on Sparse Random Graphs with Given Degree Sequence. *Internet Mathematics*, 4(4):329 – 356.

[47] Fountoulakis, N., Joos, F., and Perarnau, G. (2022). Percolation on random graphs with a fixed degree sequence. *SIAM Journal on Discrete Mathematics*, 36(1):1–46.

[48] Fountoulakis, N., Kang, R. J., and McDiarmid, C. (2010). The t-stability number of a random graph. *The Electronic Journal of Combinatorics*, 17.

[49] Frieze, A. (1988). Random regular graphs of non-constant degree. Technical report, Department of Mathematics, Carnegie Mellon University.

[50] Frieze, A. and Karoński, M. (2015). *Introduction to Random Graphs*. Cambridge University Press.

[51] Frieze, A., Krivelevich, M., and Martin, R. (2004). The emergence of a giant component in random subgraphs of pseudo-random graphs. *Random Structures & Algorithms*, 24(1):42–50.

[52] Frieze, A., Krivelevich, M., and Smyth, C. (2007). On the chromatic number of random graphs with a fixed degree sequence. *Combinatorics, Probability and Computing*, 16(5):733–746.

[53] Frieze, A. and Łuczak, T. (1992). On the independence and chromatic numbers of random regular graphs. *Journal of Combinatorial Theory, Series B*, 54(1):123–132.

[54] Gao, J., Buldyrev, S. V., Havlin, S., and Stanley, H. E. (2011). Robustness of a network of networks. *Phys. Rev. Lett.*, 107(19):195701.

[55] Gao, P., Isaev, M., and McKay, B. (2020). Sandwiching random regular graphs between binomial random graphs. In Chawla, S., editor, *Proceedings of the 2020 ACM/SIAM Symposium on Discrete Algorithms*, Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms, pages 690–701, United States of America. Society for Industrial & Applied Mathematics (SIAM).

[56] Gao, P. and Ohapkin, Y. (2020). Subgraph probability of random graphs with specified degrees and applications to chromatic number and connectivity. arXiv preprint.

[57] Gao, P., Su, Y., and Wormald, N. (2012). Induced subgraphs in sparse random graphs with given degree sequences. *European Journal of Combinatorics*, 33(6):1142–1166.

[58] Gilbert, E. N. (1959). Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144.

[59] Goerdt, A. (2001). The giant component threshold for random regular graphs with edge faults. *Theor. Comput. Sci.*, 259(1):307–321.

[60] Greenhill, C., Isaev, M., and McKay, B. D. (2021). Subgraph counts for dense random graphs with specified degrees. *Combinatorics, Probability and Computing*, 30(3):460–497.

[61] Greenhill, C. and McKay, B. D. (2009). Random dense bipartite graphs and directed graphs with specified degrees. *Random Structures & Algorithms*, 35(2):222–249.

[62] Grimmett, G. R. and McDiarmid, C. J. H. (1975). On colouring random graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 77(2):313–324.

[63] Gross, J., Yellen, J., and Zhang, P. (2013). *Handbook of Graph Theory, Second Edition*. Discrete Mathematics and Its Applications. Taylor & Francis.

[64] Hakimi, S. L. (1962). On realizability of a set of integers as degrees of the vertices of a linear graph. I. *J. Soc. Indust. Appl. Math.*, 10(3):496–506.

[65] Hasheminezhad, M. and McKay, B. D. (2010). *Combinatorial estimates by the switching method*, volume 531, pages 209–223. American Mathematical Society.

[66] Hatami, H. and Molloy, M. (2012). The scaling window for a random graph with a given degree sequence. *Random Structures & Algorithms*, 41(1):99–123.

[67] Havel, V. (1955). Poznámka o existenci konečných grafru. *Časopis pro pěstování matematiky*, 080(4):477–480.

[68] Heckel, A. (2018). The chromatic number of dense random graphs. *Random Structures & Algorithms*, 53(1):140–182.

[69] Heckel, A. (2021). Non-concentration of the chromatic number of a random graph. *Journal of the American Mathematical Society*, 34:245–260.

[70] Heckel, A. and Riordan, O. (2021). How does the chromatic number of a random graph vary? arXiv preprint.

[71] van der Hofstad, R. (2016). *Configuration Model*, volume 1 of *Cambridge Series in Statistical and Probabilistic Mathematics*, pages 216–255. Cambridge University Press.

[72] van der Hofstad, R. and Nachmias, A. (2017). Hypercube percolation. *Journal of the European Mathematical Society*, 19(3):725–814.

[73] Hulshof, T. and Nachmias, A. (2020). Slightly subcritical hypercube percolation. *Random Structures & Algorithms*, 56(2):557–593.

[74] Janson, S. (2009a). On percolation in random graphs with given vertex degrees. *Electronic Journal of Probability*, 14:86–118.

[75] Janson, S. (2009b). The probability that a random multigraph is simple. *Combinatorics, Probability and Computing*, 18(1-2):205–225.

[76] Janson, S. (2011). Probability asymptotics: notes on notation. arXiv.

[77] Janson, S. (2014). The probability that a random multigraph is simple. ii. *Journal of Applied Probability*, 51(A):123–137.

[78] Janson, S. (2020). Random graphs with given vertex degrees and switchings. *Random Structures & Algorithms*, 57(1):3–31.

[79] Janson, S. and Luczak, M. J. (2009). A new approach to the giant component problem. *Random Struct. Algorithms*, 34(2):197–216.

[80] Janson, S., Łuczak, T., and Ruciński, A. (2000). *Random graphs*. John Wiley and Sons.

[81] Jeffers, G., Dubson, M. A., and Duxbury, P. M. (1994). Islandtopercolation transition during growth of metal films. *Journal of Applied Physics*, 75(10):5016–5020.

[82] Joos, F. and Perarnau, G. (2018). Critical percolation on random regular graphs. *Proc. Amer. Math. Soc.*, pages 3321–3332.

[83] Joos, F., Perarnau, G., Rautenbach, D., and Reed, B. (2018). How to determine if a random graph with a fixed degree sequence has a giant component. *Probability Theory and Related Fields*, 170(1):263–310.

[84] Kamaldinov, D., Skorkin, A., and Zhukovskii, M. (2021). Maximum sparse induced subgraphs of the binomial random graph with given number of edges. *Discrete Mathematics*, 344(2):112205.

[85] Kang, M. and Seierstad, T. G. (2008). The critical phase for random graphs with a given degree sequence. *Combinatorics, Probability and Computing*, 17(1):67–86.

[86] Kargaltsev, S., Shabanov, D., and Shaikheeva, T. (2019). Two values of the chromatic number of a sparse random graph. *Acta Mathematica Universitatis Comenianae*, 88(3):849–854.

[87] Kelly, S. T. and Black, M. A. (2020). graphsim: An r package for simulating gene expression data from graph structures of biological pathways. *Journal of Open Source Software*, 5(51):2161.

[88] Kim, J. H., Sudakov, B., and Vu, V. (2007). Small subgraphs of random regular graphs. *Discrete Mathematics*, 307(15):1961–1967.

[89] Kim, J. H., Sudakov, B., and Vu, V. H. (2002). On the asymmetry of random regular graphs and random graphs. *Random Structures & Algorithms*, 21(3-4):216–224.

[90] Klamt, S., Haus, U.-U., and Theis, F. (2009). Hypergraphs and cellular networks. *PLOS Computational Biology*, 5(5):1–6.

[91] Knuth, D. E. (1976). Big omicron and big omega and big theta. *SIGACT News*, 8(2):18–24.

[92] Krivelevich, M., Sudakov, B., Vu, V., and Wormald, N. (2001). Random regular graphs of high degree. *Random Structures & Algorithms*, 18(4):346–363.

[93] Krivelevich, M., Sudakov, B., and Wormald, N. (2011). Regular induced subgraphs of a random graph. *Random Structures & Algorithms*, 38:235–250.

[94] Leighton, F. T. (1977). A graph coloring algorithm for large scheduling problems. *Journal of Research of the National Bureau of Standards*, 84(6):489–506.

[95] Lewis, R. M. R. (2021). *Guide to Graph Colouring*. Texts in Computer Science. Springer Cham, 2 edition.

[96] Li, D., Fu, B., Wang, Y., Lu, G., Berezin, Y., Stanley, H. E., and Havlin, S. (2015). Percolation transition in dynamical traffic network with evolving critical bottlenecks. *Proceedings of the National Academy of Sciences*, 112(3):669–672.

[97] Liebenau, A. and Wormald, N. (2019). Asymptotic enumeration of graphs by degree sequence, and the degree sequence of a random graph. arXiv preprint.

[98] Ling, R. F. (1975). An exact probability distribution on the connectivity of random graphs. *Journal of Mathematical Psychology*, 12(1):90–98.

[99] Łuczak, T. (1987). On the equivalence of two basic models of random graphs. In *Random graphs '87 : based on proceedings of the 3rd International Seminar on Random Graphs and Probabilistic Methods in Combinatorics, June 27 - July 3 1987, Poznań, Poland.*

[100] Łuczak, T. (1989). Sparse random graphs with a given degree sequence. In *Proceedings of the Symposium on Random Graphs, Poznań, Poland*, pages 165–182.

[101] Łuczak, T. (1990). Component behavior near the critical point of the random graph process. *Random Structures & Algorithms*, 1(3):287–310.

[102] Łuczak, T. (1991). A note on the sharp concentration of the chromatic number of random graphs. *Combinatorica*, 11(3):295 – 297.

[103] Maehara, H. (1987). On the number of induced subgraphs of a random graph. *Discrete mathematics*, 64(2-3):309–312.

[104] Matula, D. (1970). On the complete subgraph of a random graph. *Proc. Second Chapel Hill Conf. on Combinatorial Mathematics and its Applications*, 42:356–369.

[105] Matula, D. W. (1987). Expose-and-merge exploration and the chromatic number of a random graph. *Combinatorica*, 7(3):275–284.

[106] McDiarmid, C. (1990). On the chromatic number of random graphs. *Random Structures & Algorithms*, 1(4):435–442.

[107] McDiarmid, C., Scott, A., and Withers, P. (2021). The component structure of dense random subgraphs of the hypercube. *Random Structures & Algorithms*, 59(1):3–24.

[108] McKay, B. (2011). Subgraphs of dense random graphs with specified degrees. *Combinatorics, Probability and Computing*, 20:413–433.

[109] McKay, B. D. (1981). Subgraphs of random graphs with specified degrees. *Congressus Numerantium*, 33:213–223.

[110] McKay, B. D. and Wormald, N. C. (1984). Automorphisms of random graphs with specified vertices. *Combinatorica*, 4(4):325–338.

[111] McKay, B. D. and Wormald, N. C. (1990). Asymptotic enumeration by degree sequence of graphs of high degree. *European Journal of Combinatorics*, 11(6):565–580.

[112] McKay, B. D. and Wormald, N. C. (1997). The degree sequence of a random graph. i. the models. *Random Structures & Algorithms*, 11(2):97–117.

[113] Molloy, M. and Reed, B. (1995). A critical point for random graphs with a given degree sequence. *Random Struct. Algorithms*, 6(2-3):161–180.

[114] Nachmias, A. and Peres, Y. (2010). Critical percolation on random regular graphs. *Random Structures & Algorithms*, 36(2):111–148.

[115] Newman, M. E. (2003). Random graphs as models of networks. In Bornholdt, S. and Schuster, H., editors, *Handbook of Graphs and Networks: From the Genone to the Internet*. Wiley-VCH, 1st edition.

[116] Noga Alon, J. H. S. (2016). *The Probabilistic Method*. John Wiley and Sons, 4 edition.

[117] Palka, Z. (1984). On the number of vertices of given degree in a random graph. *Journal of Graph Theory*, 8(1):167–170.

[118] Palka, Z. (1986). Some remarks about extreme degrees in a random graph. *Mathematical Proceedings of the Cambridge Philosophical Society*, 100(1):167–174.

[119] Panagiotou, K. and Steger, A. (2009). A note on the chromatic number of a dense random graph. *Discrete Mathematics*, 309(10):3420–3423.

[120] Pittel, B. (2008). Edge percolation on a random regular graph of low degree. *The Annals of Probability*, 36(4):1359–1389.

[121] Read, R. (1981). The graph theorists who count—and what they count. In Klarner, D. A., editor, *The Mathematical Gardner*, pages 326–345. Springer US, Boston, MA.

[122] Robertson, N., Sanders, D., Seymour, P., and Thomas, R. (1997). The four-colour theorem. *Journal of Combinatorial Theory, Series B*, 70(1):2–44.

[123] Ruciński, A. (1985). Subgraphs of random graphs: A general approach. In Karoński, M. and Ruciński, A., editors, *Random Graphs '83*, volume 118 of *North-Holland Mathematics Studies*, pages 221–229. North-Holland.

[124] Ruciński, A. (1988). When are small subgraphs of a random graph normally distributed? *Probability Theory and Related Fields*, 78:1–10.

[125] Schürger, K. (1979). Limit theorems for complete subgraphs of random graphs. *Periodica Mathematica Hungarica*, 10(1):47–53.

[126] Shamir, E. and Spencer, J. (1987). Sharp concentration of the chromatic number on random graphs g n, p. *Combinatorica*, 7(1):121–129.

[127] Ta, X., Mao, G., and Anderson, B. D. O. (2009). On the giant component of wireless multihop networks in the presence of shadowing. *IEEE Transactions on Vehicular Technology*, 58(9):5152–5163.

[128] Tishby, I., Biham, O., Katzav, E., and Kühn, R. (2018). Revealing the microstructure of the giant component in random graph ensembles. *Phys. Rev. E*, 97(4-1):042318.

[129] Vecchio, F., Miraglia, F., Piludu, F., Granata, G., Romanello, R., Caulo, M., Onofrj, V., Bramanti, P., Colosimo, C., and Rossini, P. M. (2017). "small world" architecture in brain connectivity and hippocampal volume in alzheimer's disease: a study via graph theory from eeg data. *Brain imaging and behavior*, 11(2):473–485.

[130] Wormald, N. C. (1981). The asymptotic connectivity of labelled regular graphs. *Journal of Combinatorial Theory, Series B*, 31(2):156–167.

Appendix A

Common probability results

We give a list of several common tools from probability that we use throughout this thesis. These results themselves are quite basic, but we give them here to be as self-contained as possible.

A.1 The union bound

Lemma A.1. For a countable set of events $\{A_i\}_{i=1}^{\infty}$, $\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mathbb{P}\left(A_i\right)$.

The simplicity of this inequality belies its utility. In particular, if we have a set of events $\{A_i\}_{i\in I}$, and $\mathbb{P}\left(A_i\right) = o(|I|^{-1})$ for all $i \in I$, then $\mathbb{P}\left(\cup_{i\in I} A_i\right) = o(1)$. We use this idea extensively throughout this thesis. Typically, either $|I| = O(1)$, where we use the union bound to combine multiple lemmas about events that each occur with probability $1 - o(1)$, or $|I| = n$, where we are performing a union bound over all the vertices in a random graph with n vertices.

A.2 Concentration inequalities

In probability theory, concentration inequalities are a broad and indispensable class of results that provide bounds on how likely a random variable is to deviate from some value. These are invaluable tools in studying the behaviour of random graphs. Here we give a brief rundown of the common inequalities of this type that we use throughout this thesis. First we give Markov's inequality, which for nonnegative random variables gives a rough upper bound on the probability that they exceed some particular value.

Theorem A.1. (Markov's inequality) If X is a nonnegative random variable and $a > 0$, then $\mathbb{P}\left(X \geq a\right) \leq \frac{\mathbb{E}[X]}{a}$.

In many cases, "large" for us means greater than 0, and we use this to argue that $\mathbb{P}\left(X > 0\right) = \mathbb{P}\left(X \geq 1\right) = o(1)$ for certain discrete random variables X where $\mathbb{E}\left[X\right] \to 0$. This tactic is sometimes referred to as the First Moment Method.

The next inequality is Chebyshev's inequality. This is one of many inequalities used to show how likely a random variable is to deviate far from its expectation. In this case, the bound given is in terms of its variance.

Theorem A.2. (Chebyshev's inequality) Let X be a random variable and $a > 0$. Then

$$\mathbb{P}\left(|X - \mathbb{E}\left[X\right]| \geq a\right) \leq \frac{\mathrm{Var}\left(X\right)}{a^2}.$$

Next, we give an inequality that we vaguely refer to as "the second moment method", since it gives a bound in terms of the second moment of a random variable. We use this almost as a complement to Markov's inequality, to give a lower bound on the probability that a nonnegative random variable X (which in our case, often counts the number of a particular structure present in a random graph) is greater than 0.

Lemma A.3. (Second moment method) If X is a nonnegative random variable with finite variance, then

$$\mathbb{P}(X > 0) \geq \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]}.$$

Proof sketch. This inequality follows from the Cauchy-Schwarz inequality, which says for a nonnegative random variable X that

$$\mathbb{E}[X] = \mathbb{E}\left[X \mathbb{1}_{\{X>0\}}\right] \leq \mathbb{E}\left[X^2\right]^{1/2} \mathbb{P}(X > 0)^{1/2},$$

which can be rearranged to give the inequality in the lemma. □

The next bound is a variant of a classic bound known as a Chernoff bound. This is one variation of a wide variety of results about the concentration of a sum of many independent random variables. We use this inequality to study random variables that can be expressed as a sum of many independent Bernoulli trials (random variables with output 0 or 1).

Lemma A.4. (Chernoff inequality) Suppose that $S_n = X_1 + \cdots + X_n$ where $X_i \in [0, 1]$ and $\mathbb{E}[X_i] = \mu_i$ for all $i \in \{1, \ldots, n\}$ and $X_1, \ldots, X_n$ are all independent. Define $\mu = \sum_{i \leq n} \mu_i$. Then, for $t \geq 0$,

$$\mathbb{P}(S_n \geq (1+\varepsilon)\mu) \leq \exp\left(-\frac{\mu\varepsilon^2}{3}\right) \quad \text{and} \quad \mathbb{P}(S_n \geq (1-\varepsilon)\mu) \leq \exp\left(-\frac{\mu\varepsilon^2}{2}\right).$$

We also give another related bound called McDiarmid's inequality.

Theorem A.5. (McDiarmid's inequality) Let $X_1, \ldots, X_n$ be independent random variables, where X_i has range $\mathcal{X}_i$. Let $f : \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$ be a function such that if $(x_1, \ldots, x_n)$ and $(x'_1, \ldots, x'_n)$ differ only in the i^{th} coordinate, then

$$|f(x_1, \ldots, x_n) - f(x'_1, \ldots, x'_n)| \leq c_i.$$

Then for all $t > 0$,

$$\mathbb{P}(f(X_1, \ldots, X_n) - \mathbb{E}[f(X_1, \ldots, X_n)] \geq t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n c_i^2}\right).$$

The final bound we give is a classic tool from the theory of martingales. Briefly speaking, a martingale $\{X_k\}_{k \geq 0}$ with respect to a filtration $\{\mathcal{F}\}_{k \geq 0}$ is a sequence of random variables such that $\mathbb{E}[X_j | \mathcal{F}_i] = X_i$ for all $i \leq j$. The following martingale inequality, known as Azuma's inequality or the Azuma-Hoeffding inequality, is an indispensable tool in studying random processes that appear in the study of random graphs.

Theorem A.6. (Azuma's inequality) Suppose that $\{X_k\}_{k\geq 0}$ a martingale and $|X_k - X_{k-1}| \leq c_k$ almost surely. Then for all positive integers N and all $\alpha > 0$,

$$\mathbb{P}\left(|X_N - X_0| \geq \alpha\right) \leq 2\exp\left(\frac{-\alpha^2}{2\sum_{i=1}^{N} c_k^2}\right).$$

One particular reason that this is so useful is that we can often construct martingales to study the distribution of other random variables that are not martingales. One such construction is the Doob martingale, which gives rise to both vertex and edge exposure martingales. These are constructions for showing concentration of measure of arbitrary graph parameters by studying how the parameter changes if the graph is constructed one vertex (or edge) at a time. These martingales, as well as other similar constructions, are used throughout this thesis to show concentration of a wide variety of random variables.

Appendix B

Glossary

B.1 General notation

(1) $\boldsymbol{d}$: the degree sequence of the random graph $G \sim \mathcal{G}(\boldsymbol{d})$, with minimum degree at least 1.

(2) $\mathcal{G}(\boldsymbol{d})$: the set of all random graphs with degree sequence $\boldsymbol{d}$, equipped with the uniform probability measure.

(3) S: the set of vertices in G that induce the graph $G[S]$ that we study.

(4) $\overline{S}$: $[n]\backslash S$, the set of vertices that we delete from G.

(5) $M(\cdot)$: the *total degree* of a degree sequence, the sum of all its elements.

(6) $\widetilde{M}(\cdot)$: $M(\cdot) - 2n_2(\cdot)$, where $n_2(\cdot)$ is the number of elements in a sequence equal to 2.

(7) $R(\boldsymbol{d})$, $j_{\boldsymbol{d}}$: functions of $\boldsymbol{d}$ used for characterising the a.a.s. existence of giant components in $\mathcal{G}(\boldsymbol{d})$, defined in Theorem 2.2.15.

B.2 Notation for Chapters 3 to 5

(1) J: the cutoff between what we classify as "big" and "small" vertices in the reduction method, $J := \log M \log\log M$.

(2) S_{small}: the set of vertices in S with degree at most J in the graph G; equivalently, the set of elements in $\boldsymbol{d}$ with value at most J.

(3) S_{big}: the set of vertices in S with degree greater than J in the graph G; equivalently, the set of elements in $\boldsymbol{d}$ with value greater than J.

(4) ℓ: the index in the subsequence $S = \{i_1, \dots, i_s\}$ such that $S_{\text{small}} = \{i_1, \dots, i_\ell\}$.

(5) γ: the ratio of the total degree of S and the total degree of $\boldsymbol{d}$; that is, $\gamma = d(S)/M$.

(6) $\boldsymbol{d}_S$: the degree sequence of the induced graph $G[S]$. Note that this definition is unrelated to $\mathsf{d}_S(w)$ used in Chapters 6 and 7.

(7) Z_j: a binomial random variable, $Z \sim \text{Bin}\,(j, \gamma)$.

(8) Y_i: the number of vertices in S_{small} with degree i in the induced subgraph $G[S]$.

(9) $\tilde{y}_i$: a function of $(\boldsymbol{d}, S)$ defined in Equation (3.1), which corresponds to the expected number of vertices in S_{small} with induced degree i under the approximation that $\mathbb{P}(d_S(v) = i) = \mathbb{P}(Z_j = i)$ for all vertices v.

(10) y_i: the sequence $(y_i)_{i \le J}$ is defined as the sequence $(\tilde{y}_i)_{i \le J}$ after cascade rounding, as defined in Equation (3.2).

(11) $\boldsymbol{d}_I$: an "average" degree sequence for the induced graph $G[S]$ when S is fixed, defined in Definition 3.1.2.

(12) $\boldsymbol{d}_A$: an "average" degree sequence for the site-percolated $\mathcal{G}(\boldsymbol{d})$, defined in Definition 5.2.1.

(13) $\tilde{w}_i$: a function of $(\boldsymbol{d}, p)$ for constant $p \in (0,1)$, which corresponds to the expected number of vertices in S_{small} in the site-percolated $\mathcal{G}(\boldsymbol{d})$ with induced degree i under the approximation that $\mathbb{P}(d_S(v) = i) = \mathbb{P}(Z_j = i)$ for all vertices v.

(14) w_i: the sequence $(w_i)_{i \le J}$ is defined as the sequence $(\tilde{w}_i)_{i \le J}$ after cascade rounding.

(15) $\tilde{z}_k$: a function of $(\boldsymbol{d}, p)$ for constant $p \in (0,1)$, which corresponds to the expected number of vertices in S_{big} in the site-percolated $\mathcal{G}(\boldsymbol{d})$ with induced degree i under the approximation that $\mathbb{P}(d_S(v) = i) = \mathbb{P}(Z_j = i)$ for all vertices v.

(16) z_i: the sequence $(z_i)_{i \le J}$ is defined as the sequence $(\tilde{z}_i)_{i \le J}$ after cascade rounding.

(17) "Good" set S: in Chapter 5, where S is chosen randomly, we say a set S is "good" if $|S| = pn(1 \pm 3\sqrt{\log n}/\sqrt{pn})$ and $d(S) = pM(1 \pm 1/pM^{1/4})$.

B.3 Notation for Chapters 6 and 7

(1) $\mathsf{d}_S(\cdot)$: this is a function from the vertex set V to $\mathbb{N}$, which is equal to 1 for every $v \in \overline{S}$ and otherwise is simply $d(v)$ for $v \in S$. Note that this definition is unrelated to $\boldsymbol{d}_S$, the induced sequence of $G[S]$, used in Chapters 3 to 5.

(2) $\mathcal{C}(\boldsymbol{d})$: the configuration model, the set of all pairings with degree sequence $\boldsymbol{d}$, equipped with the uniform probability measure.

(3) $R_S(\boldsymbol{d}), j_S(\boldsymbol{d})$: functions of $\boldsymbol{d}$ used for characterising the a.a.s. existence of giant components in $\mathcal{C}(\boldsymbol{d})[S]$ or $\mathcal{G}(\boldsymbol{d})[S]$, defined in Equation (6.2).

(4) P: a pairing, an element of $\mathcal{C}(\boldsymbol{d})$.

(5) Φ: the set of all simple (no loops or multiple edges) pairings, equipped with the uniform probability measure.

(6) G^*: a simple pairing, an element of Φ.

(7) $\Phi(T)$: the set of all simple pairings such that there exists a sequence of choices for half edges for each $t' < t$ such that $T_{t-1} = T$.

(8) H^+: the set of elements in $\boldsymbol{d}$ (equivalently, the set of vertices in $G \sim \mathcal{G}(\boldsymbol{d})$ or $P \sim \mathcal{C}(\boldsymbol{d})$) with degree greater than $\delta\sqrt{M}$ for some $\delta \to 0$.

(9) S^+, $\overline{S}^+$: $S^+ = S \cap H^+$, $\overline{S}^+ = \overline{S} \cap H^+$.

(10) T_t: the partial pairing created by the exploration process at time t. $T_t = (V_t, E_t, \mathcal{X}_t)$, which correspond to the set of vertices, edges, and open edges (unpaired half edges that belong to vertices in $V_t \cap S$) contained in the partial pairing.

(11) X_t: $|\mathcal{X}_t|$, the number of open edges in the partial pairing at time t.

(12) X'_t: an upper bound on X_t, defined in (6.8).

(13) v_t: the parent vertex of the half edge whose pair is revealed at time t.

(14) w_t: the vertex explored at time t by the process.

(15) $d_S(\cdot)$: a function of a vertex w, defined in (6.1) - $d_S(w)$ is equal to w if $w \in S$, and 1 otherwise.

(16) $d^{(t-1)}(w)$: the *available degree* of a vertex w at time t. Defined in (7.3). For $w \in S$, this is 0 if $w \in V_{t-1}$ and $d(w)$ otherwise. For $w \notin S$, this is $d(w)$ minus the number of edges incident to w in E_{t-1}.

(17) $d'_{S,t}(\cdot)$: the number of back edges (edges between $w_t \cap S$ and $V_{t-1} \cap S$) revealed by the exploration process at time t.

(18) $\mathbb{P}_{t-1}(\cdot)$: the probability of some event, conditional on the partial pairing at time $t-1$.

(19) $\mathbb{E}_{t-1}[\cdot]$: the expectation of some event, conditional on the partial pairing at time $t-1$.

(20) f_t: defined in (7.5), the expected number of open edges gained at time t under the approximation that $\mathbb{P}_{t-1}(w_t = w) = d^{(t-1)}(w)$.

(21) A_t, B_t, G_t: random variables used for concentration arguments: $A_t = d_S(w_t) - \mathbb{E}_{t-1}[d_S(w_t)]$, $B_t = d'_{S,t}(w_t) - \mathbb{E}_{t-1}\left[d'_{S,t}(w_t)\right]$, $G_t = f_{t+1} - f_t - \mathbb{E}_{t-1}[f_{t+1} - f_t]$.