Generation and properties of random graphs and analysis of randomized algorithms

by

Pu Gao

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2009

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

We study a new method of generating random d-regular graphs by repeatedly applying an operation called pegging. The pegging algorithm, which applies the pegging operation in each step, is a method of generating large random regular graphs beginning with small ones. We prove that the limiting joint distribution of the numbers of short cycles in the resulting graph is independent Poisson. We use the coupling method to bound the total variation distance between the joint distribution of short cycle counts and its limit and thereby show that $O(\epsilon^{-1})$ is an upper bound of the ϵ-mixing time. The coupling involves two different, though quite similar, Markov chains that are not time-homogeneous. We also show that the ϵ-mixing time is not $o(\epsilon^{-1})$. This demonstrates that the upper bound is essentially tight. We study also the connectivity of random d-regular graphs generated by the pegging algorithm. We show that these graphs are asymptotically almost surely d-connected for any even constant $d \geq 4$.

The problem of orientation of random hypergraphs is motivated by the classical load balancing problem. Let $h > w > 0$ be two fixed integers. Let H be a hypergraph whose hyperedges are uniformly of size h. To *w-orient* a hyperedge, we assign exactly w of its vertices positive signs with respect to this hyperedge, and the rest negative. A (w, k)-orientation of H consists of a w-orientation of all hyperedges of H, such that each vertex receives at most k positive signs from its incident hyperedges. When k is large enough, we determine the threshold of the existence of a (w, k)-orientation of a random hypergraph. The (w, k)-orientation of hypergraphs is strongly related to a general version of the off-line load balancing problem.

The other topic we discuss is computing the probability of induced subgraphs in a random regular graph. Let $0 < s < n$ and H be a graph on s vertices. For any $S \subset [n]$ with $|S| = s$, we compute the probability that the subgraph of $\mathcal{G}_{n,d}$ induced by S is H. The result holds for any $d = o(n^{1/3})$ and is further extended to $\mathcal{G}_{n,\mathbf{d}}$, the probability space of random graphs with given degree sequence $\mathbf{d}$. This result provides a basic tool for studying properties, for instance the existence or the counts, of certain types of induced subgraphs.

Acknowledgements

The four years I have spent in Waterloo are very fruitful and memorable to me. Four years ago I was outside the door of mathematics and was learning computer science to which I had no interest. Thanks to my uncle Jason's encouragement and recommendation, I decided to come to Canada and work on mathematics. The research experience during my Master's and Ph.D. was joyful. I would like to especially thank my supervisor Nick, who patiently led me into the beautiful area of random graphs and consistently guided me all the time. Owing to him, I started my first research collaboration and my first research paper, and played my first games of pool and croquet. I would also like to thank Hania who always treated me like her own child and helped me to get accustomed to different cultures.

I owe great gratitude to my Ph.D. committee members, Luc, Penny, Jochen and Christopher. They gave me valuable suggestions of refining and improving my thesis.

During my four years' studies in C&O, I participated in regular group research discussions with the members of the group varying through the years. I would like to thank all group members, including Graeme, Carlos, Vishal, Cris and Xavier. With them I had a lot of interesting lunch discussions, series of topical lectures and many hours of delighted chats. I owe much gratitude to Graeme, who encouraged me consistently whenever I was in difficult times, especially when I was suffering by my clumsiness of English and worries of my lack of mathematical background when I just arrived in Canada.

I would also like to thank the friends that I met in Waterloo who accompanied me and brought me a lot of joy. I want to especially thank Hans for being so supportive and nice to me always. I want to also thank my family members, particularly my parents and my grandmother, who support my work all the time.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The general subject of this thesis is the properties and generation of random graphs. Random graphs (though not called random graphs at the time) became known because of a few proofs by Erdős around the 40's (for instance, the proof showing a lower bound of the Ramsey number $R(n, n)$). The approach followed in these proofs is known as the probabilistic method, which is normally used to prove the existence of graphs with certain properties. Random graphs were formally defined by Erdős and Rényi [23] in 1959 and soon afterwards, random graph theory developed as its own area of research. The term "random graphs" refers to a probability space with a (finite) set of graphs with a certain distribution. There are in general two ways to define a probability space of random graphs. The first way is to specifically define the sample space and the probability measure. The second way is to define a random process which generates a probability space of random graphs, in which the sample space and the probability measure are sometimes implicit. A probability space of random graphs is usually called a random graph model.

The earliest and also the most studied random graph models are the $\mathcal{G}_{n,p}$ model and the $\mathcal{G}_{n,m}$ model. The $\mathcal{G}_{n,p}$ model is often referred to as the binomial random graph model or the standard random graph model, which consists of all graphs on n vertices with each edge occurring independently with probability p. The $\mathcal{G}_{n,m}$ model defines a probability space of graphs with n vertices and m edges, each of which occurs with equal probability. It is convenient to use these two random graph models since their probability measures are of simple form, which make it relatively easy to compute the probabilities of events. Another commonly studied random graph model is $\mathcal{G}_{n,d}$, the probability space of d-regular graphs on n vertices with the uniform distribution. Computing the probabilities of events in $\mathcal{G}_{n,d}$ is usually not easy. In fact, the known methods for counting d-regular graphs on n vertices are already very technical, especially for large values of d. Random d-regular graphs with non-uniform distribution [62, Section 6] have been studied as well by various authors. They are generated by random processes, for which the underlying probability measures, determined by the random processes, do not have explicit form. This often makes the analysis even harder. Some of these probability spaces are believed but not yet proved to be "close to" $\mathcal{G}_{n,d}$. Research on $\mathcal{G}_{n,d}$ has also been extended to the probability spaces of random graphs with given degree sequences. There are also important random graphs motivated by real applications that have different looks from all of the previously introduced random graph models, for instance, the web graph interpreting the hyperlinks of webpages in the Internet, and the underlying topology of some of the social networks. A degree sequence $d_1, \ldots, d_n$ is said to obey a power law if there is a constant $\beta > 0$ such that for any nonnegative integer k, the number of d_i that equals k is (approximately) proportional to $k^{-\beta}$. These graphs are known to have the scale-free property [5] (the degree distribution following a power law) and they require new models to define the probability space. There have been a few models for generating such random graphs, among which the most popular one is called the preferential attachment model [58, 63], which defines a random process that generates a random graph whose degree sequence obeys a power law.

For any graph property A, instead of studying whether a given graph has the property A or not, we study the probability that a random graph has the property A. Commonly studied graph properties include connectivity, hamiltonicity, diameter and the existence or counts of certain subgraphs, for instance, trees, cycles, cliques, independent sets, etc.

The methods of studying these graph properties are very different from those in classical graph theory. Interestingly, a lot of graph properties (for instance, the properties that can be written as first order logic sentences) obey the 0-1 law in $\mathcal{G}_{n,p}$, in the sense that the probability of having property A in $\mathcal{G}_{n,p}$ converges either to 0 or 1 for most p as $n \to \infty$. We say a graph property A is *monotonically increasing* if G has property A whenever $G' \subset G$ has property A. Similarly we say a graph property A is *monotonically decreasing* if G has property A whenever $G' \supset G$ has property A. If furthermore the graph property A is monotonic, then there exists a critical value of p at which the probability of having property A has a sudden jump (from 0 to 1 or the other way around). We call this critical value of p, normally a function of n, a threshold function of property A. The thresholds of graph properties are studied in various random graph models other than $\mathcal{G}_{n,p}$.

Studying graph properties in random graph models is currently of significant interest. Given a probability space, if some graph property holds with high probability (i.e. the probability goes to 1 as the size of the random graph goes to infinity), we expect a graph picked from that probability space to have that property.

This thesis combines some recent results on three main topics: a new way of generating random regular graphs, thresholds of orientability of random hypergraphs, and computing the probabilities of induced subgraphs in random regular graphs.

In Chapter 3, we introduce a new method of generating random d-regular graphs by repeatedly applying an operation called pegging. The pegging operation is abstracted from the basic operation applied in a type of peer-to-peer network called the SWAN network. The pegging algorithm, which applies the pegging operation in each step, is a method of generating large random regular graphs beginning with small ones. We prove that for the resulting graphs, the limiting joint distribution[1] of the numbers of short cycles is independent Poisson. Properties of the short cycles are interesting because they are the only likely local structures other than trees and forests in a random regular graph with constant degree. The total variation distance between two distributions, formally defined in Section 2.1.4, is a two-variable function that measures the difference between two distributions. The ϵ-mixing time of the distribution of short cycle counts of these random regular graphs is the time at which the distribution reaches and maintains total variation distance at most ϵ from its limiting distribution. We use coupling, a proof technique introduced in Section 2.1.4, to bound the rate at which the distribution approaches its limit and thereby show that $O(\epsilon^{-1})$ is an upper bound of the ϵ-mixing time. The coupling involves two different, though quite similar, Markov chains that are not time-homogeneous. We also show that the ϵ-mixing time is not $o(\epsilon^{-1})$. This demonstrates that the upper bound is essentially tight. We study also the connectivity of random d-regular graphs generated by the pegging algorithm. We show that these graphs are asymptotically almost surely d-connected for any even constant $d \geq 4$.

In Chapter 4, we study the thresholds of the orientability of random hypergraphs, which is motivated by the classical load balancing problem. Let $h > w > 0$ be two fixed integers. Let H be a random uniform hypergraph whose hyperedges are of size h. To *w-orient* a hyperedge, we assign exactly w of its vertices positive signs with respect to this hyperedge, and the rest negative. A (w, k)-orientation of H consists of a w-orientation of all hyperedges of H, such that each vertex receives at most k positive signs from its incident hyperedges. When k is large enough, we determine the threshold of the existence of a (w, k)-orientation of a random hypergraph. The graph case, when $h = 2$ and $w = 1$, was solved recently by Cain, Sanders and Wormald and independently by Fernholz and Ramachandran, thereby settling a conjecture made by Karp and Saks. The (w, k)-orientation of hypergraphs is strongly related to a general version of the off-line load balancing problem.

In Chapter 5, we compute the probability of induced subgraphs in a random regular graph. Let $0 < s < n$ and H be a graph on s vertices. Let $[n]$ denote the set $\{1, 2, \ldots, n\}$. For any $S \subset [n]$ with $|S| = s$, we compute the probability that the subgraph of $\mathcal{G}_{n,d}$ induced by S is H. The result holds for any $d = o(n^{1/3})$ and is further extended to $\mathcal{G}_{n,\mathbf{d}}$, the probability space of random graphs with given degree sequence $\mathbf{d}$. This result provides a basic tool for studying properties, for instance the existence or the counts, of certain types of induced subgraphs.

Some background of these topics is discussed in Chapter 2. We also discuss in Chapter 2 a few methods used in the proofs in the later chapters, for instance, the coupling method,

[1]defined in Section 2.1.3

the differential equation method and the switching method.

The results on short cycle distribution, the upper and lower bounds on ϵ-mixing time have appeared in the joint paper [27, 26] with Wormald and the result on thresholds of orientability of random hypergraphs is also joint work with Wormald [28]. The result on estimation of the probability of induced subgraphs in random regular graphs is joint work with Su and Wormald.

Chapter 2

Background

In this chapter we give a brief description of problems we present in this thesis and explain the contribution of our work to the related areas. We also give an exposition of some of the methods we use in our proofs.

2.1 Generation and properties of random regular graphs

In Chapter 1, we defined the probability spaces $\mathcal{G}_{n,p}$, $\mathcal{G}_{n,m}$ and $\mathcal{G}_{n,d}$. Given $\mathbf{d} = (d_1, \ldots, d_n)$, $\mathcal{G}_{n,d}$ can be extended to $\mathcal{G}_{n,\mathbf{d}}$, the probability space of graphs on n vertices with degree sequence $\mathbf{d}$ and with the uniform distribution. For any sequence of properties A_n, we say A_n holds *asymptotically almost surely* (a.a.s.) if $\mathbf{P}(A_n) \to 1$ as $n \to \infty$. The probability spaces $\mathcal{G}_{n,d}$ and $\mathcal{G}_{n,\mathbf{d}}$ have certain advantages over $\mathcal{G}_{n,p}$ and $\mathcal{G}_{n,m}$ when applied to designing and analysing some networks. For $p = O(1/n)$ and $m = O(n)$, the degree distributions of $\mathcal{G}_{n,p}$ and $\mathcal{G}_{n,m}$ are far from regular. In particular, there are a.a.s. $\Theta(n)$ isolated vertices (vertices of degree 0). This is usually what researchers try to avoid in network design. Hence $\mathcal{G}_{n,d}$ and $\mathcal{G}_{n,\mathbf{d}}$ are more suitable models since they generate graphs with required degree sequences.

Computing probabilities of events in $\mathcal{G}_{n,d}$ is usually difficult. To implicitly define the probability measure of $\mathcal{G}_{n,d}$, it is required to compute (or estimate) the number of d-regular graphs. However, the work spread over years for different range of values of d, especially for large d. The first contribution to this problem was by Bender and Canfield [7], who obtained the asymptotic formula for the number of d-regular graphs for bounded d. Bollobás [9] then reproved the result where the constraint of bounded d was relaxed to $d < \sqrt{2 \log n}$. McKay [42] extended the formula to $d = o(n^{1/3})$ using a technique called switching and later McKay and Wormald [43] extended the formula to $d = o(\sqrt{n})$ by using an improved version of the switching operations in [42] which also makes the analysis easier. McKay and Wormald also obtained the asymptotic formula for dense regular graphs, where $d \approx cn$ for some constant c within certain range. However, the problem remains unresolved for the case that d is in the range between $\Theta(n^{1/2})$ and $cn/\log n$ for some $c > 0$.

Properties of random regular graphs in $\mathcal{G}_{n,d}$ have been studied by many authors. In areas such as Computer Science, researchers especially show interest in generating random graphs, or random regular graphs, with a given distribution. An interesting question is how to generate regular graphs from $\mathcal{G}_{n,d}$. The pairing model, explained in Section 2.1.1, can be used to generate random d-regular graphs. However, it only works efficiently for small values of d. Currently there is no known efficient algorithm to generate random d-regular graphs with uniform distribution when d is $\Omega(n^{1/3})$. There are some alternative algorithms that efficiently (linear time) generate random regular graphs whose distributions are non-uniform but are believed to be "close to" uniform. Two of the well known algorithms are the d-process and the d^*-process. The d-process starts with an empty graph on n vertices and repeatedly adds a random edge without having the degree of any vertex exceed d. The d^*-process starts with an empty graph on n vertices and in each step it chooses a random vertex v whose degree is below d and then repeatedly joins v by an edge to another vertex chosen uniformly at random (u.a.r.) from those whose degrees are less than d and who are non-adjacent to v, until the degree of v reaches d. It was shown in [55, 50, 49] that the d-process and the d^*-process a.a.s. terminate with a graph that is d-regular for bounded d.

Two probability spaces $\mathcal{G}_n$ and $\mathcal{G}'_n$ are *contiguous* if and only if for any event A_n, A_n is a.a.s. true in $\mathcal{G}_n$ if and only if A_n is a.a.s. true in $\mathcal{G}'_n$. Proving or disproving that the probability space generated by the d-process or the d^*-process is contiguous with $\mathcal{G}_{n,d}$ is still open.

The random d-regular graphs in $\mathcal{G}_{n,d}$ have nice properties. For $d \geq 3$, they are a.a.s. d-connected [62, Theorem 2.10], hamiltonian [62, Theorem 2.26] and with diameter $O(\log n)$ [62, Theorem 2.13]. Some of these almost sure properties have been proved to hold in other random regular graph models, like the random regular graphs generated by the d-process and the d^*-process. Various types of subgraphs in $\mathcal{G}_{n,d}$ have been investigated, which we discuss more in detail in Section 2.3 and Chapter 5. The distribution of the number of short cycles, a particular type of subgraphs, have been examined independently by Bollobás [9] and Wormald [59]. The short cycles caught special attention because they are the only local structures that are likely to occur in $\mathcal{G}_{n,d}$ besides trees and forests. The joint distribution of short cycle counts is asymptotically independent Poisson in various random regular graphs models. The joint distribution of short cycle counts can be used as part of a proof of contiguity between two random regular graph models in some cases (for instance, it was used in [35] to prove that the uniform model of random d-regular graphs is contiguous to the model of random d-regular graphs obtained by adding a uniformly random perfect matching to a uniformly random $(d-1)$-regular graph conditional on that the resulting graph is simple). The results on the distribution of short cycle counts supports the belief of the contiguity between the random regular graph models (the uniform model, the d-process, the d^*-process) mentioned above. The difficulty of proving the contiguity is discussed in Chapter 6.

2.1.1 The pairing model

The pairing model (also called the configuration model) was first introduced by Bollobás [9] to enumerate d-regular graphs on n vertices when dn is an even number. In the pairing model, let the n vertices be represented as n buckets each containing d points. Take a random partition of all points into $dn/2$ pairs. Such a partition is called a *pairing*. Then contract each bucket and represent it as a vertex whereas each pair corresponds to an edge. It can be easily shown that the resulting graph, if it is simple, is a d-regular graph chosen u.a.r. from all d-regular graphs on n vertices. Let $\mathcal{M}(n,d)$ denote the set of pairings generated by the pairing model. It follows easily that $|\mathcal{M}(n,d)| = (nd-1)(nd-3)\cdots 1$. Each simple d-regular graph corresponds to the same number, $d!^n$, of pairings in $\mathcal{M}(n,d)$. Hence enumerating d-regular graphs is equivalent to computing the probability that a pairing in $\mathcal{M}(n,d)$ corresponds to a simple graph. Let $\mathbf{P}(\text{simple})$ denote this probability. Then Bender and Canfield [7] showed that

$$\mathbf{P}(\text{simple}) \sim \exp\left(\frac{1-d^2}{4}\right), \tag{2.1.1}$$

with bounded $d \geq 1$. The constraint of bounded d was later relaxed by various authors [9, 42, 43].

The advantage of using the pairing model is that computing probabilities of events in $\mathcal{M}(n,d)$ is easy. In general, let A denote a graph property of d-regular graphs (any graph property can be considered to be a set of graphs), and let $A' \subset \mathcal{M}(n,d)$ denote the set of pairings, each of which corresponds to a graph in A. Then we can compute $\mathbf{P}_{\mathcal{M}(n,d)}(A')$, which denotes the probability of A' in $\mathcal{M}(n,d)$. Knowing the asymptotic formula of $\mathbf{P}(\text{simple})$, we can now estimate the asymptotic value of $\mathbf{P}_{\mathcal{G}_{n,d}}(A)$, since

$$\mathbf{P}_{\mathcal{G}_{n,d}}(A) = \mathbf{P}_{\mathcal{M}(n,d)}(A')/\mathbf{P}(\text{simple}).$$

Instead of computing $\mathbf{P}_{\mathcal{M}(n,d)}(A')$, sometimes it is easier to compute the probability of $A'' \subset \mathcal{M}(n,d)$, where any pairing in A'' corresponds to a multigraph with property A. For example, consider A to be the property that a graph (multigraph) is Hamiltonian. Computing $\mathbf{P}_{\mathcal{M}(n,d)}(A'')$ is easier than $\mathbf{P}_{\mathcal{M}(n,d)}(A')$ because we do not have to restrict pairings in A'' to those corresponding to simple graphs. Since $A' \subset A''$,

$$\mathbf{P}_{\mathcal{G}_{n,d}}(A) \leq \mathbf{P}_{\mathcal{M}(n,d)}(A'')/\mathbf{P}(\text{simple}).$$

For any fixed $d > 0$, $\mathbf{P}(\text{simple}) = \Theta(1)$ by (2.1.1). Thus, $\mathbf{P}_{\mathcal{M}(n,d)}(A'') = o(1)$ implies $\mathbf{P}_{\mathcal{G}_{n,d}}(A) = o(1)$. It follows that to show any graph property that holds a.a.s. in $\mathcal{G}_{n,d}$ it is enough to show that the corresponding property holds a.a.s. in $\mathcal{M}(n,d)$.

2.1.2 SWAN and the pegging algorithm

Random regular graphs have recently arisen in a peer-to-peer ad-hoc network, called the SWAN (Small-World Wide Area Network), introduced by Bourassa and Holt [14]. The SWAN consists of a group of peers that connect to each other. It uses a random regular graph as its underlying topology. In the random regular graph, each vertex represents a peer and each edge represents a point-to-point connection between peers. Currently the SWAN is implemented on a random 4-regular graphs although the basic operations are defined for any even $d \geq 4$. In the SWAN, peers arrive and leave randomly. When a peer joins the network, SWAN randomly chooses $d/2$ disjoint connections (edges) and interposes the new peer (vertex) on each of them. This operation that models a peer joining the network is called "clothespinning", whereas the reverse operation of "clothespinning" models a peer leaving the network. Occasionally some adjustment is required to repair the network if these operations cannot cope, for instance, if the network is too small (with less than $d + 1$ peers), or if a peer breaks down without being able to notify his neighbours first, or a peer departs from the network with all his neighbours mutually linked, etc. The result is a random graph whose distribution is not fully understood. Bourassa and Holt found experimentally that it has good connectivity and diameter properties. More recently, Cooper, Dyer and Greenhill [21] defined a Markov chain on d-regular graphs with randomised size to model (a simplified version of) the SWAN network. Each move of the Markov chain is a clothespinning operation or the reverse and no other operations used in [14] are considered. They showed that, conditional on the graph having certain size, the stationary distribution is uniform, and they bounded the mixing time of the chain. However, the Markov chain they defined is restricted such that the sizes of the graphs are bounded in probability.

In our work, we define an algorithm which generates random d-regular graphs for constant d. The algorithm simply repeats clothespinning (which we call *pegging*) operations, without performing the reverse. The definition of the algorithm will be formally defined in Chapter 3. We will focus mainly on even d, in which case a pegging operation can be visualised as binding the middles of $d/2$ nonadjacent edges together using a new vertex. Thus the size of the graph increases linearly with the number of operations. This gives an extreme version of the SWAN network, in which no peer ever leaves the network. Since the analysis of [21] does not apply if the network undergoes net long-term growth, by studying this extreme case we hope to gain knowledge of which properties of the random SWAN network are not sensitive to long-term growth.

2.1.3 Method of moments

Let ℓ be a fixed positive integer and let $X_{n,1},\ldots,X_{n,\ell}$ be ℓ discrete random variables. We say that the joint distribution of $X_{n,1},\ldots,X_{n,\ell}$ asymptotically has the probability distribution function $\sigma : \mathbb{Z}^\ell \to \mathbb{R}$, if for any $(x_1,\ldots,x_\ell) \in \mathbb{Z}^\ell$,

$$\mathbf{P}(X_{n,1} = x_1,\ldots,X_{n,\ell} = x_\ell) \to \sigma(x_1,\ldots,x_\ell), \quad \text{as } n \to \infty.$$

For any integers X and $i \geq 0$, let $[X]_i$ denote the i-th falling factorial of X, i.e. $[X]_i = X(X-1)\cdots(X-i+1)$. By convention, $[X]_0 = 1$ and $[X]_i = 0$ if $i > X$.

In general, the method of moments in probability theory is a method of proving that a sequence of random variables $(X_i)_{i\geq 1}$ converge in distribution to a random variable X. It can be described as follows. If the expectation of every moment of X exists and the distribution of X is completely determined by its moments, then we have that if for any $k \geq 1$, $\mathbf{E}(X_i^k) \to \mathbf{E}(X^k)$, then X_i converges to X in distribution. For a complete statement of the result and its history, readers can refer to [48, Section 4.1, Proposition 6].

The following theorem, a simpler form of which is stated in [3, Theorem 8.3.1], is a special case of the method of moments in general and is often used to prove that a sequence of vectors of random variables converges in distribution to the independent Poisson.

Theorem 2.1.1 *Let l be a fixed positive integer. Given l nonnegative constants $\mu_1,\ldots,\mu_l$, let $X_{n,1},\ldots,X_{n,l}$ be l nonnegative random variables such that for any $j_1 \geq 0,\ldots,j_l \geq 0$,*

$$\lim_{n\to\infty} \mathbf{E}\left(\prod_{i=1}^{l} [X_{n,i}]_{j_i}\right) = \prod_{i=1}^{l} \mu_i^{j_i},$$

then $X_{n,1},\ldots,X_{n,l}$ are asymptotically independent Poisson random variables with means $\mu_1,\ldots,\mu_l$.

2.1.4 Markov chain and Coupling

A random process $(X_i)_{i\geq 0}$ is called a discrete time *Markov chain* if the following Markov property is satisfied. For any $t \geq 0$,

$$\mathbf{P}(X_{t+1} = x_{t+1} \mid X_t = x_t,\ldots,X_0 = x_0) = \mathbf{P}(X_{t+1} = x_{t+1} \mid X_t = x_t).$$

The set of possible values of X_i for any $i \in \mathbb{N}$ is a countable set, which is called the *state space* of the Markov chain.

A Markov chain is called *irreducible* if for any states i and j, there exists $n > 0$, such that

$$\mathbf{P}(X_n = j \mid X_0 = i) > 0.$$

For any state i, the period of i is defined to be $\gcd\{n : \mathbf{P}(X_n = i \mid X_0 = i) > 0\}$. A Markov chain is said to be aperiodic if the period of each state in the chain is 1.

A Markov chain is called *ergodic* if it is irreducible and aperiodic, or equivalently, a Markov chain is ergodic if for any states i and j, there exists $N > 0$ such that for all $n > N$,

$$\mathbf{P}(X_n = j \mid X_0 = i) > 0.$$

A distribution π on the sate space of a Markov chain is called a *stationary distribution* if for any $t \geq 0$, if the distribution of the Markov chain at step t is π, then its distribution at step $t+1$ is also π.

A well known property of an ergodic finite state Markov chain is that the distribution of X_t converges to a unique stationary distribution regardless of the initial distribution of X_0. This result can be found in any text book of probability theory, for instance, [19, Chapter 10].

Let σ and π be probability distributions on the same countable state space $\mathcal{S}$. The *total variation distance* between σ and π is defined as

$$d_{TV}(\sigma, \pi) = \sup_{A \subset \mathcal{S}} \{\sigma(A) - \pi(A)\}. \tag{2.1.2}$$

Equivalently,

$$d_{TV}(\sigma, \pi) = \frac{1}{2} \sum_{x \in \mathcal{S}} |\sigma(x) - \pi(x)|.$$

As we discussed before, an ergodic finite state Markov chain converges to a unique stationary distribution. Let σ_t denote the distribution of the Markov chain at step t and let π denote the stationary distribution. For a given small constant $\epsilon > 0$ (1/4 is often chosen as the value of ϵ in engineering or other applied areas), the *mixing time* of the Markov chain is defined to be the minimum t such that the total variation distance between σ_t and π remains below ϵ in all following steps. To denote the dependence on ϵ, the mixing time t is a function of ϵ. We call this t the *ϵ-mixing time*, though this is not a standard definition.

The coupling method can be used to bound the mixing time and ϵ-mixing time of a Markov chain. In general, a *coupling* of two random variables X_1 and X_2 with the same state space (but not necessarily defined on the same probability space) is a construction of X_1 and X_2 simultaneously on the same probability space. With only a slight abuse of notation, we write this as a random pair (X_1, X_2), where the marginal distribution of X_i as the i-th coordinate is the same as the distribution of the original variable X_i ($i = 1$ and 2). The coupling method as a proof technique was first introduced by Doeblin in the 1930s. It becomes a popular technique in proving the mixing times of ergodic Markov chains by various authors [1, 2, 37]. Lindvall [40, pp. 9–12] gave an elaborate definition of coupling that is equivalent for our purposes, and gave a corresponding general coupling lemma which we may state as follows.

Lemma 2.1.2 (Coupling Lemma) *Let (X_1, X_2) be a coupling and let σ_i denote the distribution of X_i. Then*

$$d_{TV}(\sigma_1, \sigma_2) \leq \mathbf{P}(X_1 \neq X_2).$$

Proof Since $\mathcal{S}$ is countable, there are probability functions p_1 and p_2 for distributions σ_1 and σ_2. Define

$$A = \{s \in \mathcal{S} : p_1(s) > p_2(x)\}.$$

Then clearly $A \subset \mathcal{S}$ maximizes the right hand size of (2.1.2) and hence the supremum can be achieved. Then

$$\begin{aligned} d_{TV}(\sigma_1, \sigma_2) &= \mathbf{P}(X_1 \in A) - \mathbf{P}(X_2 \in A) \\ &\leq \mathbf{P}(X_1 \in A) - \mathbf{P}(X_1 = X_2 \wedge X_1 \in A) \\ &= \mathbf{P}(X_1 \in A) + \mathbf{P}(X_1 \neq X_2 \vee X_1 \notin A) - 1 \\ &\leq \mathbf{P}(X_1 \in A) + \mathbf{P}(X_1 \neq X_2) + \mathbf{P}(X_1 \notin A) - 1 \\ &= \mathbf{P}(X_1 \neq X_2). \quad \blacksquare \end{aligned}$$

If $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are two random processes in the same state space, a random process $\big((X_t, Y_t)\big)_{t\geq 0}$ is a *coupling of the two processes* if (X_t, Y_t) is a coupling of X_t and Y_t for all $t \geq 0$.

The coupling method, as it is usually applied to bound the mixing time of a Markov chain, can be described briefly as follows. Assume $(X_t)_{t\geq 0}$ is a Markov chain starting with state $X_0 = x_0$. Construct another Markov chain $(X'_t)_{t\geq 0}$ with the same transition probabilities as $(X_t)_{t\geq 0}$ for all $t \geq 0$, but starting with X'_0 as a random variable with the stationary distribution of the Markov chain. Then X'_t has the same distribution for all $t \geq 0$. Construct a coupling of the two Markov chains $(X_t, X'_t)_{t\geq 0}$. Then by the coupling lemma, the total variation distance between the distribution of X_t and X'_t is bounded by $\mathbf{P}(X_t = X'_t)$, whilst the distribution of X'_t remains the stationary distribution. Since the transition probabilities are the same in $(X_t)_{t\geq 0}$ and $(X'_t)_{t\geq 0}$, we can construct the coupling in such a way that as long as $X_t = X'_t$ for some $t \geq 0$, X_t and X'_t take the same transition for all the following steps. That is to say, $\mathbf{P}(X_{t+1} = X'_{t+1}) \geq \mathbf{P}(X_t = X'_t)$ for all $t \geq 0$. It then follows that for any $\epsilon > 0$, if $d_{TV}(X_\tau, X'_\tau) < \epsilon$ for some $\tau > 0$, then $d_{TV}(X_t, X'_t) < \epsilon$ for any $t \geq \tau$. Hence the ϵ-mixing time is bounded by the minimum t such that $\mathbf{P}(X_t \neq X'_t) < \epsilon$ and this minimum t can usually be estimated by analysing the process $(X_t, X'_t)_{t\geq 0}$.

The coupling method can be applied to more general settings. For instance, it can be used to bound the total variation distance of two random processes that converge to the same limiting distribution. Since the transition probabilities in these two random processes may be different, we cannot construct a coupling that forces X_t and X'_t to take the same transition even if $X_t = X'_t$. This can make the analysis more complicated. In our work of getting an upper bound of the ϵ-mixing time of the joint distribution of the short cycle counts in the random regular graphs generated by the pegging algorithm, we applied the coupling method to bound the ϵ-mixing time of two slightly different Markov chains. Because part of the proof is included in my Master's research paper [25], it is not presented in this thesis. We only state the main result of the upper bound since we need it to prove some other results in the thesis. Interested readers can refer to [27].

2.1.5 Small subgraph conditioning method

The small subgraph conditioning method can be used to prove some almost sure graph property in a random regular graph, or the contiguity between two random regular graph probability spaces. We assume readers are familiar with the first and second method methods. Let $\mathcal{G}$ be a probability space of random d-regular graphs and let $Y = Y(n)$ be a non-negative random variable in $\mathcal{G}$. Assume $\mathbf{E}Y(n) \to \infty$ as $n \to \infty$. If we wish to show that a.a.s. $Y > 0$ (by $Y > 0$ a.a.s. we mean there exists $\delta > 0$ such that $\mathbf{P}(Y(n) > \delta) \to 1$ as $n \to \infty$), the first try would be to compute the second moment of Y. It comes immediately from the second moment method that a.a.s. $Y > 0$ if the second moment of Y is asymptotically the square of $\mathbf{E}Y$. However, the second moment method fails if the second moment of Y is at least $c(\mathbf{E}Y)^2$ for some $c > 1$. The small subgraph conditioning method can be used to show that $Y > 0$ a.a.s. in some situation in which the "large" deviation of Y is caused by the appearance of short cycles and by conditioning on the numbers of cycles of any finite size, we obtain "small" variance. The method can be described as follows. Let $\mathcal{G}^{(Y)}$ be the probability space obtained from $\mathcal{G}$ by letting $\mathbf{P}_{\mathcal{G}^{(Y)}}(G) = Y(G)\mathbf{P}_{\mathcal{G}}(G)/\mathbf{E}Y$ for every $G \in \mathcal{G}$, where $\mathbf{E}Y$ is the expectation of Y in $\mathcal{G}$. Recall that two probability spaces $\mathcal{G}_1$ and $\mathcal{G}_2$ are contiguous, denoted by $\mathcal{G}_1 \approx \mathcal{G}_2$, if and only if for any event that is a.a.s. true in one probability space, it is also a.a.s. true in the other. If we can show that $Y/\mathbf{E}Y$ converges in distribution to some random variable W as $n \to \infty$, and $W > 0$ a.s., it follows that $\mathcal{G}$ is contiguous with $\mathcal{G}^{(Y)}$ which immediately implies $Y > 0$ a.a.s.. (see Janson, Łuczak and Ruciński [36, P. 266] for a discussion of this). The following theorem gives the conditions under which $Y/\mathbf{E}Y$ converges in distribution to some random variable W.

Theorem 2.1.3 ([35]) *Let $\lambda_i > 0$ and $\delta_i \geq -1$, $i \in \mathbb{Z}_+$ be real numbers and $X_i = X_i(n)$, $Y = Y(n)$ be random variables on the same probability space $\mathcal{G} = \mathcal{G}(n)$, where n is an implicit parameter in these random variables and probability spaces under consideration. Suppose X_i are non-negative integer valued and Y is non-negative and $\mathbf{E}Y > 0$ for sufficiently large n. Suppose furthermore that*

(a) for every $k \geq 1$, X_i, $1 \leq i \leq k$ are asymptotically independent Poisson random variables with $\mathbf{E}X_i \to \lambda_i$;

(b) for all non-negative integers $j_1 \geq 0, \ldots, j_k \geq 0$,

$$\frac{\mathbf{E}(Y[X_1]_{j_1} \cdots [X_k]_{j_k})}{\mathbf{E}Y} \to \prod_{i=1}^{k} (\lambda_i(1+\delta_i))^{j_i};$$

(c) $\sum_{i=1}^{\infty} \lambda_i \delta_i^2 < \infty$;

(d)

$$\frac{\mathbf{E}Y(n)^2}{(\mathbf{E}Y(n))^2} \leq \exp\left(\sum_{i=1}^{\infty} \lambda_i \delta_i^2\right) + o(1), \quad \text{as } n \to \infty.$$

Then

$$\frac{Y(n)}{\mathbf{E}Y(n)} \to W = \prod_{i=1}^{\infty} (1+\delta_i)^{Z_i} e^{-\lambda_i \delta_i} \quad \text{in distribution, as } n \to \infty,$$

where the variables Z_i are independent Poisson variables with means λ_i for $i \in \mathbb{Z}_+$. Moreover, this convergence and the convergence of the X_i to the Z_i as in condition (a) all hold jointly.

The following theorem shows that to prove $Y > 0$ a.a.s. or $\mathcal{G} \approx \mathcal{G}^{(Y)}$, it is enough to check that certain conditions hold.

Theorem 2.1.4 ([35]) *If all the conditions of Theorem 2.1.3 are satisfied, then*

$$\mathbf{P}(Y(n) > 0) = \exp\left(-\sum_{i:\delta_i=-1} \lambda_i\right) + o(1),$$

and, provided $\sum_{i:\delta_i=-1} \lambda_i < \infty$, $\overline{\mathcal{G}}^{(Y)} \approx \overline{\mathcal{G}}$, where $\overline{\mathcal{G}}$ denotes the probability space obtained from $\mathcal{G}$ by conditioning on the event $\bigwedge_{i:\delta_i=1}(X_i = 0)$ ($X_i = 0$ for all i such that $\delta_i = -1$).

These two theorems also explain why the studies of the short cycle distributions in various random regular graph models are important. Assume $\mathcal{G}$ is a probability space of random d-regular graphs and Y is a random variable defined in $\mathcal{G}$. If we can show that the numbers of short cycles in $\mathcal{G}$ are asymptotically independent Poisson variables with means λ_i, $i \in \mathbb{Z}_+$, these random variables immediately give X_i used in Theorem 2.1.3 (a). If furthermore, we can show that the numbers of short cycles in $\mathcal{G}^{(Y)}$ are also asymptotically independent Poisson variables with means λ'_i, $i \in \mathbb{Z}_+$, then condition (b) of the theorem is satisfied by taking $\delta_i = \lambda'_i/\lambda_i - 1$. Having λ_i and δ_i, it is straightforward to check whether condition (c) is satisfied. However, it is sometimes difficult to verify condition (d), which requires the computation of the second moment of Y. We will discuss more about the usage of the small subgraph conditioning method in Chapter 6. Interested readers can refer to [51] by Robinson and Wormald, where the small subgraph conditioning method was used to show that $\mathcal{G}_{n,3} \approx \mathcal{G}_{n,3}^{(H_n)}$, where H_n is a random variable in $\mathcal{G}_{n,3}$, denoting the number of Hamilton cycles.

2.2 Orientation of hypergraphs and the load balancing problem

An *h-hypergraph* is a hypergraph whose hyperedges are of size uniformly h. Let $h > w$ be two given positive integers. A hyperedge is said to be *w-oriented* if exactly w distinct vertices in it are marked with positive signs with respect to the hyperedge. The *indegree* of a vertex is the number of positive signs it receives. Let k be a positive integer. A (w, k)-orientation of an h-hypergraph is a w-orientation of all hyperedges such that each vertex has indegree at most k. If such a (w, k)-orientation exists, we say the hypergraph is (w, k)-orientable. Of course, being able to determine the (w, k)-orientability of an h-hypergraph H for all k solves the optimisation problem of minimising the maximum indegree of H. If a graph ($h = 2$) is $(1, k)$-oriented, we may orient each edge towards its vertex of positive sign, and we say the graph is *k-oriented*.

2.2.1 Application to load balancing

The hypergraph orientation problem is motivated by the classical load balancing problem, which has appeared in various guises in computer networking. The aim is to spread the work among a group of computers, hard drives, CPUs, or other resources. The on-line version of the load balancing problem can be considered as jobs coming sequentially and being assigned to a group of machines. To save time and storage space, for each job, the load balancer decides which machine it goes to without knowing the complete information of the current load of all machines. The goal is to minimize the heaviest load, Ł, of the machines. Mitzenmacher, Richa and Sitaraman [45] have surveyed the history, applications and techniques related to this area. A simple method that is widely used in load balancing is known as the two-choice or multi-choice paradigm. It can be explained by the following problem studied by Azar, Broder, Karlin and Upfal [4]. Consider assigning n sequentially coming jobs to n machines. For each job that arrives, the multi-choice algorithm randomly chooses h machines, and then assigns the job to the one with the lightest load. Surprisingly, Ł decreases significantly by changing the value of h from 1 to 2. They have shown that when $h = 1$, a.a.s. Ł $\sim \log n/ \log\log n$, and when $h \geq 2$, a.a.s. Ł $\sim \log\log n/ \log h$. There are also linear time algorithms that achieve Ł $= O(m/n)$ [22, 38, 44] or Ł $= m/n+O(\log\log n)$ [4, 8] for m jobs and n machines with $h \geq 2$.

Another application of the load balancing mentioned by Cain, Sanders and Wormald [15] is the disk scheduling problem, in which w out of h pieces of data are required to reconstruct a logical data block. Each piece of data is initially stored in a randomly chosen disk. The purpose of such design is to guarantee that the logical data blocks can be successfully reconstructed even if there is some damage of one or a few disks. It also helps to balance the load. Whenever a request of logical data block arrives, the scheduler lets it retrieve data from the w least busy disks among the h disks that store the required information. Interested readers can refer to [56] for further references.

In the off-line version, jobs or requests are not processed sequentially. The balancer (scheduler) processes all jobs (requests) received in a time period. Take the disk scheduling problem as an example, let m denote the number of requests to be scheduled and n the number of disks. Each request j retrieves data from w out of h disks that are randomly chosen from the n disks. The goal is to minimize the heaviest load Ł. When $h = 1$, the performance is the same as the on-line version because there is no choice. However, the off-line version performs better when $h \geq 2$. It was shown in the case of $h = 2$, that if $m < cn$ requests are to be scheduled for some constant $c > 0$, then a.a.s. the maximum load is c' whose value depends on c. An optimal allocation can be achieved in polynomial time ($O(m^2)$) by solving a maximum flow problem. As explained in [15], it is desirable to find fast algorithms (linear time algorithms for instance) that achieve maximum load close to the optimal. There are linear time algorithms that achieve maximum load $O(m/n)$ [22, 38, 44].

2.2.2 Orientability thresholds

Problems of orientability of graphs have been studied by many authors. Results on orientability of graphs with prescribed in-degree out-degree sequences or with given in-degree and out-degree bounds have been surveyed in [57, Section 61.1]. Some of these proofs use the flow technique. Problems of orientability of graphs with other restrictions such as edge connectivity are also surveyed in [57, Chapter 61].

As explained by Cain et al. [15], a graph model represents the off-line load balancing problem for $h = 2$. Again, we take the disk scheduling problem as an example. The disks are vertices and a request that can retrieve data from u or v is represented as an undirected edge $\{u, v\}$. Recall that a graph can be k-oriented if there exists an orientation of all edges such that the maximum in-degree is at most k. Note that the graph can be k-oriented if and only if there is a scheduling such that each disk receives at most k requests. In [15], and simultaneously by Fernholz and Ramachandran [24], the sharp threshold $m(n)$ was found for k-orientability in $G_{n,m}$, i.e. the function such that a graph in $G_{n,m}$ is a.a.s. k-orientable when $m < m(n) - \epsilon n$, and a.a.s. not when $m > m(n) + \epsilon n$. This is the first proof of a conjecture of Karp and Saks [38], that this threshold coincides with the threshold at which the $(k+1)$-core has mean degree at most $2k$. Note that the k-core of a graph is the subgraph with maximum number of edges and with minimum vertex degree at least k. Hakimi's theorem [32, Theorem 4] implies that a graph can be k-oriented if and only if there does not exist a subgraph with average degree more than $2k$. Hence the k-orientability thresholds immediately provide the thresholds of appearance of subgraphs with average degree at least $2k$. Both proofs in [15] and [24] analyse a linear time algorithm that finds a k-orientation a.s.s. when the mean degree of the $(k+1)$-core is slightly less than $2k$. The proof in [24] was significantly simpler, which was made possible chiefly because they analysed a different algorithm which uses the trick of "splitting vertices" to postpone decisions.

The above graph model can be generalized to the hypergraph model that represents the multi-choice allocation problem for $h \geq 2$. Let the vertices represent disks and let each hyperedge have size h and represent the set of disks from which a request would retrieve the data it needs. The hypergraph can be (w, k)-oriented if and only if there is a scheduling such that each request is scheduled to retrieve data from w disks and no disk receives more than k requests.

Let $\mathcal{G}_{n,m,h}$ denote the probability space of random h-hypergraphs on n vertices and m hyperedges with the uniform distribution. In the thesis, we determine the orientability thresholds of $\mathcal{G}_{n,m,h}$. We also generalise Hakimi's theorem [32, Theorem 4] to hypergraphs and therefore we determine the threshold of appearance of certain type of subgraphs with certain density. Note that a trivial upper bound for the k-orientability thresholds is $m = kn/w$ by counting the positive signs that are to be assigned to all vertices.

In all previous work of determining the orientability thresholds of random graphs, a central role is played by the k-core of random graphs. The k-orientability of random graphs has the same threshold as the $(k+1)$-core reaching a certain density. In common with these approaches, we first find what we call the $(w, k+1)$-core of a hypergraph, which is a generalisation of the k-core. We use the differential equation method[1] to analyse the size and density of the $(w, k+1)$-core in $\mathcal{G}_{n,m,h}$. Then we show that the (w, k)-orientability

[1]The details of the differential equation method are given in Section 2.2.3.

of $\mathcal{G}_{n,m,h}$ has the same threshold as the $(w, k+1)$-core reaching a certain type of density. However, our approach of determining the threshold is quite different from that in previous work. The technique in [24] does not seem to apply at all in the hypergraph case, and the algorithm used by [15] is already very complicated when $h = 2$ and $w = 1$. Unlike [15, 24], the proof is not based on analysing an algorithm which finds an orientation, but instead uses the natural representation of the problem in terms of flows and uses the max-flow min-cut theorem. We prove the orientability threshold of random hypergraphs analogous to the one conjectured by Karp and Saks [38] for the random graph case, for all sufficiently large k. For its special case $h = 1$ and $w = 1$, i.e. the graph case, our method provides a simpler proof (for sufficiently large k) than the proofs of [15] and [24].

Determining thresholds of orientability of random hypergraphs is useful since we can predict the maximum load, knowing the density of the hypergraph. Although the algorithm in [24] does not seem to generalise to the hypergraph case, the one in [15] has many possible generalisations and we presume that an asymptotically optimal algorithm for the hypergraph orientation problem can be obtained in this way. However, this seems formidable to analyse in the general setting. We hope that our method in the proof can potentially lead to some fast algorithm that a.a.s. finds an optimal orientation.

2.2.3 The differential equation method

Let $\mathcal{P} = (p_t)_{t\geq 0}$ denote a discrete time random process with p_t an element in a probability space Ω_t. There is a parameter n implicit in those probability spaces. For instance, consider some algorithm that is run on a random graph $G \in \mathcal{G}_{n,p}$. The algorithm generates a random process $\mathcal{P}$ with n as a hidden parameter in each step. Let $H_t = (p_0, \dots, p_t)$, which is the history of the process up to step t. Let $a > 0$ be a fixed integer and for every $1 \leq l \leq a$, let $Y_t^{(l)} : H_t \to \mathbb{R}$ be a random variable defined on the histories of the process. The differential equation method (DE method) is used to show the concentration of variables $Y_t^{(l)}$ over the process if there are only small changes of each random variable in each step and the expected change of each random variable in each step satisfies certain conditions.

Given $a > 0$ as a fixed integer, a multivariable function $f : \mathbb{R}^a \to \mathbb{R}$ is said to satisfy some Lipschitz condition if there exists an absolute constant $C > 0$ such that for any $(x_1, \dots, x_a) \in \mathbb{R}^a$ and $(y_1, \dots, y_a) \in \mathbb{R}^a$,

$$|f(x_1, \dots, x_a) - f(y_1, \dots, y_a)| \leq C \sum_{i=1}^{a} |x_i - y_i|.$$

Theorem 2.2.1 (Wormald [60]) *Let $\mathcal{P}$ be a random process and H_t the history of $\mathcal{P}$ up to step t. Let a be fixed. For $1 \leq l \leq a$, let $Y_t^{(l)}$ be a random variable depending on H_t and*

$f_l : \mathbb{R}^{(a+1)} \to \mathbb{R}$ be a multivariable function such that for every l and for some constant C, $Y_t^{(l)} \le Cn$ always for every t. Suppose for some $m = m(n)$,

(i) there exists a constant $C' > 0$ such that for all $0 \le t \le m$ and every H_t, $|Y_{t+1} - Y_t| < C'$;

(ii) for all $0 \le t \le m$ and all $1 \le l \le a$ and every H_t,

$$\mathbf{E}(Y_{t+1}^{(l)} - Y_t^{(l)} \mid H_t) = f_l(t/n, Y_t^{(1)}/n, \ldots, Y_t^{(a)}/n) + o(1);$$

(iii) for every $1 \le l \le a$, the function f_l is continuous with a Lipschitz condition on a bounded connected open set D where D contains the intersection of $(t, z^{(1)}, \ldots, z^{(a)} : t \ge 0)$ with some neighbourhood of $(0, z^{(1)}, \ldots, z^{(a)} : \mathbf{P}(Y_0^{(l)} = z^{(l)}n, 1 \le l \le a) \ne 0$ for some $n)$.

Then

(a) For any $(0, \hat{z}^{(1)}, \ldots, \hat{z}^{(a)}) \in D$, the differential equation system

$$\frac{d\ z_l}{d\ s} = f_l(s, z_1, \ldots, z_a), \quad l = 1, \ldots, a$$

has a unique solution in D for $z_l : I\!R \to I\!R$ with the initial conditions

$$z_l(0) = \hat{z}^{(l)}, \quad l = 1, \ldots, a,$$

where the solution is extended arbitrarily close to the boundary of D.

(b) A.a.s.

$$Y_t^{(l)} = n z_l(t/n) + o(n)$$

uniformly for all $0 \le t \le m$

Numerous examples of using the DE method to analyse random processes can be found in [61]. In Chapter 4, we will apply Theorem 2.2.1 to analyse a randomized algorithm that generates a random process, which starts from a probability space of random hypergraphs and terminates with outputting the $(w, k+1)$-core of the input.

2.3 Probabilities of (induced) subgraphs in random regular graphs

Properties of subgraphs and induced subgraphs in random graph models have been investigated by various authors. Ruciński [52, 54] studied the distribution of the count of small subgraphs in $\mathcal{G}_{n,p}$ and the condition under which the distribution converges to the normal distribution. He also studied properties of induced subgraphs in [53].

The techniques for proving results in the standard random graph model $\mathcal{G}_{n,p}$ do not apply in the random regular graph model $\mathcal{G}_{n,d}$. For any graph G on vertex set $[n]$ and $S \subset [n]$, let G_S denote the subgraph of G induced by the vertex set S. Let H be a graph on the vertex set S. For a random graph G, we study the probability that G_S equals H, denoted by $\mathbf{P}(G_S = H)$, for any $S \subset [n]$ and any H on the vertex set S.

Even though computing $\mathbf{P}(G_S \supseteq H)$ or $\mathbf{P}(G_S = H)$ in $\mathcal{G}_{n,p}$ is trivial, computing these probabilities in $\mathcal{G}_{n,d}$ is not easy, especially when the degree d goes to infinity as n goes to infinity. McKay [41] estimated lower and upper bounds of $\mathbf{P}(G_S \supseteq H)$ in $\mathcal{G}_{n,d}$ when the degree sequence of H and d satisfy certain conditions. These bounds are useful in estimating the asymptotic value of $\mathbf{P}(G_S \supseteq H)$ when d is not too large or H is small. However, the result and the technique in the proof do not apply to the induced subgraph case. Gao and Wormald [29] proved that the distribution of the number of small subgraphs with certain restrictions converges to the normal distribution in $\mathcal{G}_{n,d}$. However, no such results on induced subgraphs are known. On the other hand, for very dense regular graphs, Krivelevich, Sudakov and Wormald [39] computed $\mathbf{P}(G_S = H)$ in $\mathcal{G}_{n,d}$ when n is odd, $d = (n-1)/2$ and $|V(H)| = o(\sqrt{n})$.

An asymptotic formula of the probability that $G_S = H$ or $G_S \supseteq H$ in a random bipartite graph with a specified degree sequence has been derived by Bender [6] when the maximum degree is bounded. The result was extended further by Bollobás and McKay [11] and by McKay [42] when the maximum degree goes to infinity slowly as n goes to infinity. Greenhill and Mckay [31] recently derived the asymptotic formula for the case when the random bipartite graph is sufficiently dense and H is sparse enough.

Let $\mathbf{d} = (d_1, \ldots, d_n)$ be a vector of nonnegative integers. Recall that $\mathcal{G}_{n,\mathbf{d}}$ denotes the class of graphs with degree sequence $\mathbf{d}$ and the uniform distribution, which generalises $\mathcal{G}_{n,d}$.

In Chapter 5, we compute the probability that $G_S = H$ in $\mathcal{G}_{n,\mathbf{d}}$ when $d_{\max} = o(M^{1/4})$, where $d_{\max}$ is the maximum degree and M is the degree sum. The power of this result is that there is no restriction on the size or density of H. Computing this probability is useful as a basic tool of studying the properties of induced subgraphs. For instance, given H, we can use this result to compute the expected number of copies of induced subgraphs in $\mathcal{G}_{n,\mathbf{d}}$ that are isomorphic to H. In Section 5.2, as a direct application of our main result, we compute the probability that a given set of vertices in $\mathcal{G}_{n,d}$ is an independent set.

2.3.1 Graph enumeration and the switching method

Enumerating combinatorial objects is a broad subject in general, which has its own research interest and has meanwhile been used in many other areas. We limit our discussion only to enumerating graphs with given degree sequences in this thesis. We have introduced in Section 2.1 the result on enumerating d-regular graphs, using the pairing model as illustrated in Section 2.1.1. The essential idea is to compute $\mathbf{P}$(simple), the probability that a pairing in $\mathcal{M}(n, d)$ corresponds to a simple graph. The estimation of this probability given by Bollobás using the pairing model holds only for $d < \sqrt{2 \log n}$. McKay extended this result to $d = o(n^{1/3})$ using the switching method. In general, a switching operation applied to a graph G refers to replacing a (finite) set of edges in G by another set of edges not in G such that the degree sequence of G does not change after the operation is applied. A switching operation applied to a pairing refers to choosing a (finite) set of pairs in the pairing and rematching the pairs in some certain way.

A pair $\{u, v\}$ in a pairing is called a *loop* if u and v are contained in the same vertex. Two pairs $\{u_1, v_1\}, \{u_2, v_2\}$ are called a *double pair* if u_1, u_2 are in one vertex and v_1, v_2 are in another vertex. Similarly we can define a triple pair (or a multipair with any given multiplicity m).

Figure 2.1 and 2.2 illustrate the type of switching operations used by McKay [42] applied to a pairing. In both figures, the switching operation switches two pairs, turning P_1 to P_2. Figure 2.1 shows the switching operation that removes a loop from P_1 and Figure 2.2 shows the switching operation that removes a double pair.

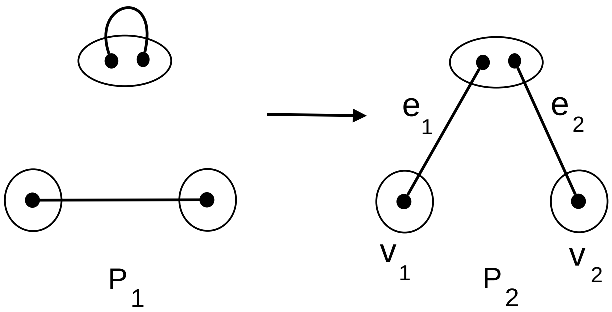

Figure 2.1: *A switching operation for removing a loop*

Next we discuss the usage of switching operations for computing $\mathbf{P}$(simple). First we can show that a random pairing in $\mathcal{M}(n, d)$ for $d = o(n^{1/3})$ a.a.s. contains no double loops or triple pairs, using the first moment method. Then it is enough to consider pairings in $\mathcal{M}(n, d)$ containing only loops and double pairs other than simple pairs. Let $\mathcal{C}_{l,d}$ denote the set of d-regular graphs containing l loops and d double pairs and no double loops or

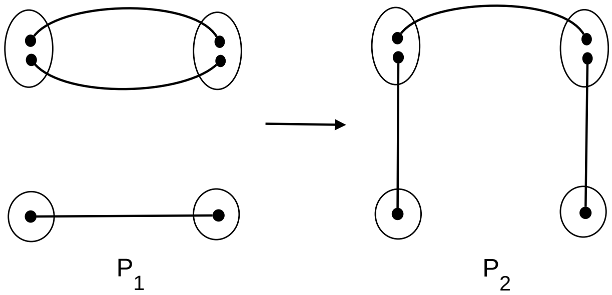

Figure 2.2: *A switching operation for removing a double pair*

triple pairs. The switching method can be summarised as follows. For a given pairing $P_1 \in \mathcal{C}_{l,d}$, we estimate N, the number of ways to apply the switching operation illustrated in Figure 2.1 such that a loop is removed after the operations are applied without any other loops or double pairs being created or destroyed simultaneously. We also estimate N', the number of inverse switching operations such that a loop is created after the inverse operations are applied to $P_2 \in \mathcal{C}_{l-1,d}$, without any other loops or double pairs being created or destroyed. Then both $|\mathcal{C}_{l-1,d}|\mathbf{E}N$ and $|\mathcal{C}_{l,d}|\mathbf{E}N'$ count the number of pairs (P_1, P_2), where $P_1 \in \mathcal{C}_{l,d}$ and $P_2 \in \mathcal{C}_{l-1,d}$ such that P_2 is obtained by applying a switching operation to P_1 (such a pair (P_1, P_2) is called a *closely related* pair in McKay's paper [42]). Then $|\mathcal{C}_{l,d}|/|\mathcal{C}_{l-1,d}| = \mathbf{E}N'/\mathbf{E}N$. Similarly we can estimate $|\mathcal{C}_{l,d}|/|\mathcal{C}_{l,d-1}|$. With the estimates of such ratios, we can estimate $\mathbf{P}(\text{simple})$ by using

$$\frac{1}{\mathbf{P}(\text{simple})} = (1 + o(1)) \sum_{l=0}^{\infty} \sum_{d=0}^{\infty} \frac{|\mathcal{C}_{l,d}|}{|\mathcal{C}_{0,0}|}.$$

This is because $\sum_{d=0}^{\infty} |\mathcal{C}_{l,d}|$ counts all pairings (except for pairings with double loops or triple pairs, which are negligible as discussed before) whilst $|\mathcal{C}_{0,0}|$ counts the simple graphs. The error term $o(1)$ in the above formula accounts for neglecting pairings containing double loops or triple pairs. If the above sum has exponentially small tail, which often happens, it is enough to sum over the main part and bound the error caused by cutting the tail.

The difficulty in analysing the switching operations in [42] is that N and N' depend on some local structures. For instance, if we count the inverse switchings in Figure 2.1 that can be applied to a given pairing $\mathcal{P}$, which do not create other loops or double pairs, it is required to choose two adjacent pairs e_1 and e_2 in P_2 such that there is no pair between v_1 and v_2, since otherwise after the switching operation is applied, a double pair will be created between v_1 and v_2 other than the created loop. Therefore the number of such 2-paths depends on the number of triangles in P_1.

McKay and Wormald [43] improved the switching operations in [42] for which the number of ways to perform a switching operation does not depend on such local structures.

Using these improved switching operations makes it easier to get more precise estimation for larger values of d and allows the constraint on d to be relaxed to $d = o(\sqrt{n})$. For d in this range, a random pairing in $\mathcal{M}(n, d)$ may contain triple pairs, too. The switching operations used by McKay and Wormald are illustrated in Figure 2.3, removing/creating loops, double pairs and triple pairs respectively.

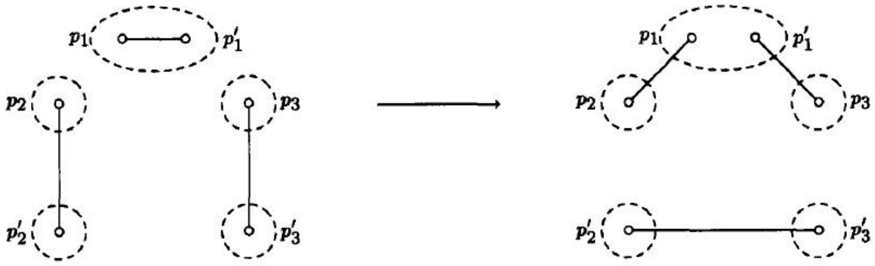

Figure 2.3: *A switching operation for removing a loop (figure from [43])*

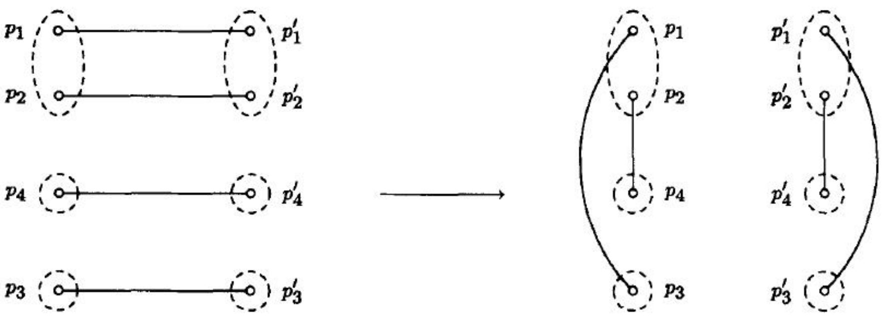

Figure 2.4: *A switching operation for removing a double pair (figure from [43])*

As we will see later in Chapter 5, computing probabilities of induced subgraphs relates to enumerating a certain type of graphs, called the B-graphs, with a given degree sequence. The B-graphs with given degree sequences can be enumerated using the pairing model and we compute $\mathbf{P}$(simple) using different types of switching operations.

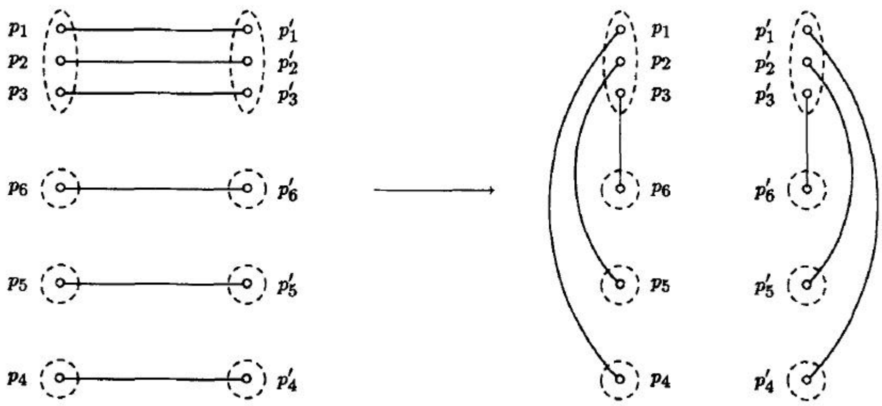

Figure 2.5: *A switching operation for removing a triple pair (figure from [43])*

Chapter 3

Random regular graphs and a peer-to-peer network

3.1 Introduction

We first introduce the *pegging algorithm*, motivated by the SWAN network, which generates random d-regular graphs for constant d. In Section 3.2 we study the joint distribution of short cycle counts in the random d-regular graph generated by pegging, for $d = 4$. As mentioned in Chapter 1, in most models of random regular graphs (for instance, the uniform model, the random d-regular graphs generated by the d-process or the d^*-process), there are unlikely to be subgraphs with more edges than vertices. The short cycles are the only small subgraphs that are likely to occur other than the trivial structures such as trees and forests.

Bollobás [9] and Wormald [59] independently showed that the joint distribution of the short cycle counts is asymptotically the independent Poisson in the uniform model. In Section 3.2, we show that the joint distribution of the short cycle counts in random regular graphs generated by the pegging algorithm also converges to the independent Poisson as $t \to \infty$ but with different means from the uniform model. We also determine the rate at which the short cycle distribution converges to its limit in Section 3.3. In Section 3.4, we study the connectivity of a random d-regular graph generated by the pegging algorithm. This is indicative of the connectivity properties of the SWAN network under long-term growth. It is well known [10, 59] that the random d-regular graphs are a.a.s. d-connected in the uniform model for any fixed constant $d \geq 3$. We prove this property holds a.a.s. in random d-regular graphs generated by the pegging algorithm for $d = 4$.

There are some preliminary results in this Chapter that have been done for my Master's research paper [25]. The expected number of k-cycles in the pegging model, for any fixed

$k \geq 3$ and even $d \geq 4$, was computed. For $k = 3$, there is an attempt of proving a Poisson limiting distribution of the number of triangles and an upper bound of the ϵ-mixing time (which is called the pseudo mixing time in [25]) using the coupling. However, the argument is not very rigorous there. For instance, Table 1 and Table 2 given in the proof of [25, Theorem 3.1] that define a coupling of two random variables (X_t, Y_t) do not work for all values of X_t and Y_t. Thus, there needs to be a separate argument bounding the probability that the value of X_t or Y_t is too large for the coupling given by Table 1 and Table 2 to be applied. The proof has been improved and made rigorous during my Ph.D. studies and the result has been extended to any fixed $k \geq 3$ and any $d \geq 3$, by coupling two random vectors. However, the main ideas (including coupling two random vectors for any fixed k) are the same as what was presented in [25]. Thus, we do not present it here in this thesis. The improved coupling argument can be found in [27] as a joint work with Wormald. In the thesis, we approach the limiting distribution using the method of moments, which is a separate argument of proving an upper bound of the ϵ-mixing time using the coupling. There are two reasons of doing this. One is that this is a much simpler proof and also a more standard way of showing the limiting distribution than using the coupling. The other is that, as just explained, the tables in the proof of [25, Theorem 3.1] do not work for large values of X_t and Y_t, and we need the higher moments of these random variables instead of their expectation to bound the probability that their values are too large. There are a few more points that need to be addressed. Some lemmas and corollaries (e.g. the lemmas and corollaries numbered from 3.2.4 to 3.2.9) are based on rather loose arguments in the proof of [25, Theorem 3.1], which bounds the probability of rare events. These lemmas are indeed interesting since the property of having small probabilities of certain types of subgraphs is shared in other random regular graph models as well. We generalise and strengthen these arguments in the thesis and put them into separate lemmas and corollaries. We also use them to prove the limiting distribution using the method of moments. We improve the error bounds in [25, Lemmas 3.2 and 3.3] from $O(n_t^{-3/4})$ to $O(n_t^{-1})$ in Lemma 3.2.3.

We first define and study the pegging algorithm applied to d-regular graphs when d is even, in particular for $d = 4$. We discuss the case of general $d \geq 3$ in Chapter 6. We define the *pegging operation* on a d-regular graph, where d is even, as follows.

Pegging Operation for Even d
Input: a d-regular graph G, where d is even.
Choose a set E_1 of $d/2$ pairwise non-adjacent edges in $E(G)$ uniformly at random.
Let $\{u_1, u_2, \ldots, u_d\}$ denote the vertices incident with edges in E_1.
$V(H) := V(G) \cup \{u\}$, where u is a new vertex.
$E(H) := (E(G) \setminus E_1) \cup \{uu_1, uu_2, uu_3, \ldots, uu_d\}$.
Output: H.

The newly introduced vertex u is called the *peg vertex*, and we say that the edges deleted are *pegged*. Figure 3.1 illustrates the pegging operation with $d = 4$.

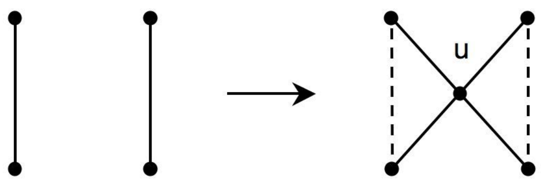

Figure 3.1: *Pegging operation when* $d = 4$

We first verify that the pegging operation is well defined. It is enough to show that for any d-regular graph, there exists a matching of size $d/2$. Given any d-regular graph G on n vertices, $n \geq d + 1$. Let $\mathcal{M}$ be a maximum matching of G. We show that $|\mathcal{M}|$ is at least $d/2$. Assume not, then $|\mathcal{M}|$ is at most $d/2 - 1$. Since $n \geq d + 1$, there exists a vertex $v \in G$ that is unmatched (a vertex is unmatched if it is not incident to any edge in $\mathcal{M}$). An *alternating path* is a path in which the edges belong alternatively to $\mathcal{M}$ and not to $\mathcal{M}$. Let P be a longest alternating path starting from the vertex v. Let u be the other end vertex of P. Since $\mathcal{M}$ is maximum, u is matched (incident to an edge in $\mathcal{M}$). Since P is a longest alternating path starting from v, all the neighbours of u are in P. Since $|\mathcal{M}| \leq d/2 - 1$, the number of vertices in P is at most $2(d/2 - 1) + 1 = d - 1$. This implies that u has at most $d - 2$ neighbours, contradicting G being d-regular. Hence $|\mathcal{M}| \geq d/2$. This shows that there always exists a set of $d/2$ non-adjacent edges in any d-regular graph, and so the pegging operation is well defined.

It is immediate that the graph outputted by the pegging operation is also d-regular. The *pegging algorithm* starts from a nonempty d-regular graph G_0 ($d \geq 4$ and even), for example, K_{d+1}, and repeatedly applies pegging operations. For $t > 0$, G_t is defined inductively to be the graph resulting when the pegging operation is applied to G_{t-1}. Clearly, G_t contains $n_t := n_0 + t$ vertices. We denote the resulting random graph process $(G_0, G_1, \ldots)$ by $\mathcal{P}(G_0)$, or $\mathcal{P}(G_0, d)$ if we wish to specify the degree d of the vertices of G_0.

The simplest and perhaps most interesting case is for $d = 4$. Here, the algorithm starts from a 4-regular graph G_0 with n_0 vertices. At each step, it randomly chooses two non-adjacent edges, deletes them, and joins a newly created vertex to the four end vertices of the deleted edges. Thus G_t contains $2n_t$ edges.

3.2 The short cycle distribution

In this section, we consider only the case $d = 4$. Let $Y_{t,k}$ denote the number of k-cycles in G_t. We show in this section that the limiting joint distribution of $Y_{t,3}, \ldots, Y_{t,k}$ for any fixed $k \geq 3$ is the independent Poisson with means defined in 3.2.1. Note that initially, the number of triangles might be as big as $2n_0$. (Note that this number is at most $2n_0$ because each vertex is contained in at most 6 triangles and there are n_0 vertices, whereas each triangle is counted three times by considering the number of vertices it contains, and so the number of triangles is at most $6n_0/3 = 2n_0$. This upper bound $2n_0$ is achieved when n_0 is a multiple of 5 and G_0 is a union of a set of K_5.) However, as we will show later, in such an extreme case, $\mathbf{E}Y_{t,3}$ will decrease quickly in the early stage of the algorithm.

For $k \geq 3$ we define

$$\mu_k = \frac{3^k - 9}{2k}. \tag{3.2.1}$$

Theorem 3.2.1 *Let G_0 and $k \geq 3$ be fixed. Then in $\mathcal{P}(G_0, 4)$,*

$$\mathbf{E}Y_{t,k} = \mu_k + O\left(n_t^{-1}\right).$$

Moreover, the joint distribution of $Y_{t,3}, \ldots, Y_{t,k}$ is asymptotically that of independent Poisson variables with means $\mu_3, \ldots, \mu_k$.

We begin with a simple technical lemma that will be used several times in this chapter. The notations $O()$ occurring in the following lemma and subsequently are defined as follows: for each occurrence of the notation $O(f)$, where f is a function of t and $G_0, \ldots, G_t$, there exists a constant C, depending only on n_0 and k, such that the absolute value of the term denoted $O(f)$ is at most $C|f|$. In particular, this is for all n in the following.

Lemma 3.2.2 *Let $c > 0$, p, a and ρ be constants with $p < c$. If $(a_n)_{n\geq 1}$ is a sequence of nonnegative real numbers with a_1 being an absolute constant (independent of n), such that*

$$a_{n+1} = \left(1 - cn^{-1} + O(n^{-2})\right) a_n + \rho n^{-p} + \gamma(n)$$

for all $n \geq 1$, then

$$a_n = \begin{cases} \big(\rho/(c-p+1)\big)n^{-p+1} + O(n^{-p}) & \text{if } \gamma(n) = O(n^{-(p+1)}), \\ \big(\rho/(c-p+1)\big)n^{-p+1} + o(n^{-p+1}) & \text{if } \gamma(n) = o(n^{-p}). \end{cases}$$

In particular $a_n = O(n^{-p+1})$ for all $n \geq 1$.

Proof When $\gamma(n) = O(n^{-(p+1)})$, we have

$$a_{n+1} = \exp\left(-\frac{c}{n} + O(n^{-2})\right) a_n + \frac{\rho}{n^p} + O(n^{-(p+1)}). \tag{3.2.2}$$

Iterating this gives

$$\begin{aligned}
a_n &= a_1 \exp\left(-\sum_{i=1}^{n-1} \frac{c}{i} + O(i^{-2})\right) + \sum_{i=1}^{n-1} \exp\left(-\sum_{j=i+1}^{n-1} \frac{c}{j} + O(j^{-2})\right)\left(\frac{\rho}{i^p} + O(i^{-(p+1)})\right) \\
&= a_1 \exp\left(-c\log n + O(1)\right) + \sum_{i=1}^{n-1} \exp\left(-c\log(n/i) + O(i^{-1})\right)\left(\frac{\rho}{i^p} + O(i^{-(p+1)})\right) \\
&= O(n^{-c}) + \sum_{i=1}^{n-1} \frac{\rho i^{c-p}}{n^c}\left(1 + O(i^{-1})\right) \\
&= \frac{\rho}{(c-p+1)} n^{-p+1} + O(n^{-p}).
\end{aligned}$$

When $\gamma(n) = o(n^{-p})$, by simply modifying the above computation we obtain

$$\begin{aligned}
a_n &= \exp\left(-\frac{c}{n} + O(n^{-2})\right) a_n + \frac{\rho}{n^p} + o(n^{-p}) \\
&= a_1 \exp\left(-\sum_{i=1}^{n-1} \frac{c}{i} + O(i^{-2})\right) + \sum_{i=1}^{n-1} \exp\left(-\sum_{j=i+1}^{n-1} \frac{c}{j} + O(j^{-2})\right)\left(\frac{\rho}{i^p} + o(i^{-p})\right) \\
&= O(n^{-c}) + \sum_{i=1}^{n-1} \frac{\rho i^{c-p}}{n^c} (1 + o(1)) \\
&= O(n^{-c}) + \left(\frac{\rho}{(c-p+1)} n^{-p+1} + O(n^{-p})\right)(1 + o(1)) \\
&= \frac{\rho}{(c-p+1)} n^{-p+1} + o(n^{-p+1}).
\end{aligned}$$

Lemma 3.2.2 follows. ∎

To show that $Y_{t,3}, Y_{t,4}, \ldots, Y_{t,k}$ are asymptotically independent Poisson random variables, it is enough, by the method of moments (introduced in Section 2.1.3), to check that their moments are asymptotic to those of independent Poisson random variables with fixed means.

Lemma 3.2.3 *For $k \geq 3$,*

$$\mathbf{E}Y_{t,k} = \mu_k + O\left(n_t^{-1}\right).$$

Proof Our analysis is based on the graph G_t produced in step t. For step $t+1$, two non-adjacent edges e_1 and e_2 are chosen in the pegging operation. There are $2n_t$ choices for e_1, and then $2n_t - 7$ choices for e_2 to be non-adjacent to e_1. So the number of ways to choose an ordered pair (e_1, e_2) is $2n_t\,(2n_t - 7)$, and hence the total number of ways to do a pegging operation in step $t+1$ is

$$N_t = \frac{2n_t\,(2n_t - 7)}{2} = n_t\,(2n_t - 7)\,. \tag{3.2.3}$$

We prove by induction on k that, for k fixed, $\mathbf{E}Y_{t,k} = \mu_k + O(n_t^{-1})$ for all $t \geq 0$. Note that in the inductive hypothesis, the notation $O()$ implicitly contains a constant that depends on k. For the base case, we consider $k = 3$.

For this and many similar calculations, to estimate the expected change in a variable counting copies of some subgraph, we consider the number of copies of the subgraph created in one step, and separately subtract the number destroyed. In particular, if a subgraph contains either of the pegged edges, it is destroyed.

We need to consider the creation of a new triangle. Given an edge e of G_t not in a triangle, a new triangle is created containing e if and only if the two pegged edges e_1 and e_2 are both adjacent to e. Of course, in view of the definition of pegging, they must be incident with different end-vertices of e. Since G_t is 4-regular, the number of ways to choose such e_1 and e_2 is precisely 9. Note also that only one edge of a given new triangle was already present in G_t. It follows that the expected number of new triangles created is at least $9\,(2n_t - 3Y_{t,3})\,/N_t$, with N_t given above. An obvious upper bound is $9 \cdot 2n_t/N_t$.

To destroy a triangle, either e_1 or e_2 must lie in the triangle, and there are of course $2n_t - 7$ choices for another edge to be pegged. So for each triangle in G_t, the probability that it is destroyed is $3\,(2n_t - 7)\,/N_t$. Thus, the expected number of existing triangles destroyed is $3Y_{t,3}\,(2n_t - 7)\,/N_t = 3Y_{t,3}/n_t$.

It follows that the expected value of $Y_{t+1,3} - Y_{t,3}$, given G_t, satisfies

$$\frac{18}{2n_t - 7} - \frac{3Y_{t,3}}{n_t}\left(1 + \frac{9}{2n_t - 7}\right) \leq \mathbf{E}\,(Y_{t+1,3} - Y_{t,3} \mid G_t) \leq \frac{18}{2n_t - 7} - \frac{3Y_{t,3}}{n_t}.$$

Thus

$$\mathbf{E}\,(Y_{t+1,3} \mid G_t) = \left(1 - \frac{3 + O(n_t^{-1})}{n_t}\right) Y_{t,3} + \frac{9}{n_t} + O\left(n_t^{-2}\right). \tag{3.2.4}$$

Taking expectation of both sides and applying the Tower Property of conditional expectations, we obtain

$$\mathbf{E}\,Y_{t+1,3} = \left(1 - \frac{3 + O(n_t^{-1})}{n_t}\right) \mathbf{E}\,Y_{t,3} + \frac{9}{n_t} + O\left(n_t^{-2}\right)$$

where the $O()$ terms are to be read as stated prior to the lemma statement, and in particular they are independent of $G_0, \ldots, G_t$. Putting $\lambda_{t,3} = \mathbf{E}(Y_{t,3} - 3)$ gives

$$\lambda_{t+1,3} = \left(1 - \frac{3}{n_t}\right) \lambda_{t,3} + O\left(\frac{1 + \lambda_{t,3}}{n_t^2}\right).$$

Applying Lemma 3.2.2, we have $\lambda_{t,3} = O(n_t^{-1})$ and hence $\mathbf{E}Y_{t,3} = 3 + O(n_t^{-1})$. This establishes the base case of the induction, i.e. for $k = 3$.

Now assume the inductive hypothesis is true of all integers smaller than k. There are two ways that one pegging operation can create a k-cycle. The first way occurs when two non-adjacent edges are pegged such that some $(k-1)$-cycle contains exactly one of them. The expected number of k-cycles created in this way is

$$\frac{(k-1)Y_{t,k-1}(2n_t - k - 3)}{N_t}.$$

The second way occurs when the two end edges of a k-path are chosen for pegging. The number of paths of length k in G_t starting from a fixed vertex v is at most $4 \cdot 3^{k-1}$, so the number of k-paths in G_t is at most $2 \cdot 3^{k-1} n_t$. This counts all walks of length k that do not immediately retrace a step, so is an over-count due to repeated vertices in the cases that the walk contains at least one cycle. There are $\sum_{i=1}^{k} Y_{t,i}$ cycles of size at most k in G_t. If we pick an edge in each of those cycles and exclude all walks containing the selected edges, we have an upper bound on the number of walks counted that are not paths. The number of selected edges is at most $\sum_{i=1}^{k} Y_{t,i}$, and each edge is contained in at most $k3^{k-1}$ walks. So G_t contains at least $2 \cdot 3^{k-1} n_t - k3^{k-1} \sum_{i=1}^{k} Y_{t,i}$ different k-paths. Thus the expected number of k-cycles created in this way, given G_t, is

$$\frac{2 \cdot 3^{k-1} n_t + O\left(\sum_{i=1}^{k} Y_{t,i}\right)}{N_t}.$$

Note that $N_t = 2n_t^2(1 + O(n_t^{-1}))$ and, by induction, $\mathbf{E}\, Y_{t,i} = O(1)$ for $i < k$. It thus follows from the two cases above that the expected number of new k-cycles created in going from G_t to G_{t+1} is

$$\frac{3^{k-1} + (k-1)\mathbf{E}Y_{t,k-1}}{n_t} + O\left(\frac{1 + \mathbf{E}Y_{t,k}}{n_t^2}\right).$$

Similar to the case of $k = 3$, given G_t, a given k-cycle is destroyed if and only if some edge contained in the k-cycle is pegged. The probability for that to occur is $k(2n_t - 7)/N_t - k(k-3)/(2N_t)$, where $k(k-3)/(2N_t)$ accounts for the over-counting in the first term when both pegged edges are in the k-cycle. Hence the expected number of k-cycles destroyed is $kY_{t,k}/n_t + O(Y_{t,k}/n_t^2)$. Combining the creation and destruction cases, we find that

$$\mathbf{E}Y_{t+1,k} - \mathbf{E}Y_{t,k} = \frac{3^{k-1} + (k-1)\mathbf{E}Y_{t,k-1} - k\mathbf{E}Y_{t,k}}{n_t} + O\left(\frac{1+\mathbf{E}Y_{t,k}}{n_t^2}\right) \tag{3.2.5}$$

By induction, $\mathbf{E}Y_{t,k-1} = \mu_{k-1} + O(n_t^{-1})$, so

$$\begin{aligned}
\mathbf{E}Y_{t+1,k} &= \left(1 - \frac{k}{n_t} + O(n_t^{-2})\right)\mathbf{E}Y_{t,k} + \frac{3^{k-1}}{n_t} + \frac{(k-1)\,\mathbf{E}Y_{t,k-1}}{n_t} + O(n_t^{-2}) \\
&= \left(1 - \frac{k}{n_t} + O(n_t^{-2})\right)\mathbf{E}Y_{t,k} + \frac{3^{k-1}}{n_t} + \frac{k-1}{n_t}\left(\frac{3^{k-1}-9}{2} + O(n_t^{-1})\right) + O(n_t^{-2}) \\
&= \left(1 - \frac{k}{n_t} + O(n_t^{-2})\right)\mathbf{E}Y_{t,k} + \frac{k\mu_k}{n_t} + O(n_t^{-2}).
\end{aligned}$$

Letting $\lambda_{t,k} = \mathbf{E}Y_{t,k} - \mu_k$ gives

$$\lambda_{t+1,k} = \left(1 - \frac{k}{n_t}\right)\lambda_{t,k} + O\left(\frac{1+\lambda_{t,k}}{n_t^2}\right),$$

and so by Lemma 3.2.2, we have $\lambda_{t,k} = O(n_t^{-1})$, and hence $\mathbf{E}Y_{t,k} = \mu_k + O(n_t^{-1})$ for any constant $k \geq 3$. Lemma 3.2.3 follows. ∎

Define $\Psi(i,r)$ to be the set of graphs with i vertices, minimum degree at least 2, and excess r, where the excess of a graph is the number of edges minus the number of vertices. Define $W_{t,i,r}$ to be the number of subgraphs of G_t in $\Psi(i,r)$. In the uniform model, $\mathbf{E}W_{t,i,r}$ is $o(1)$ when $r \geq 1$ and is bounded when $r = 0$ [62, Lemma 2.7]. The following lemma shows that this property holds also in the pegging model. For the following lemma the constants implicit in $O()$ depend on i.

Lemma 3.2.4 *For fixed $i > 0$ and $r \geq 0$,*

$$\mathbf{E}W_{t,i,r} = O(n_t^{-r}).$$

We first give a sketch of the proof. The proof is obtained by induction on r and i. Any graph in $\Psi(i,r)$ contains at least one cycle since it has minimum degree at least 2. Thus $\Psi(i,r) = \emptyset$ for $i = 1, 2$. The base case is $r = 0$ and $i = 3$. So $H \in \Psi(3,0)$ is a triangle. Hence the base case holds by Lemma 3.2.3. By checking ways of destroying subgraphs in $\Psi(i,0)$ or creating subgraphs in $\Psi(i,0)$ we can inductively prove that $\mathbf{E}W_{t,i,0} = O(1)$. For any fixed $r \geq 1$, the number of ways of destroying or creating subgraphs in $\Psi(i,r)$ depends on subgraphs in $\Psi(i', r-1)$ for some $i' \leq i$. Then by induction on both i and r, the lemma follows. This lemma will be used frequently in this chapter when we need to bound the probability of the occurrence of certain subgraphs.

Proof We proceed by induction on r and i. Any graph in $\Psi(i, r)$ contains at least one cycle since it has minimum degree at least 2. Thus $\Psi(i, r) = \emptyset$ for $i = 1, 2$. The base case is $r = 0$ and $i = 3$. So $H \in \Psi(3, 0)$ is a triangle. Hence the base case holds by Lemma 3.2.3.

Assume $W_{t,i-1,0} = O(1)$ for any $i \geq 4$. Let H be any graph in $\Psi(i, 0)$. Since the excess of H is 0, every component of H is a cycle.

We bound the expected number of subgraphs in $\Psi(i, 0)$ created when going from G_t to G_{t+1}. We omit some simple details that are virtually the same as those in the proof of Lemma 3.2.3. We also note that for any fixed i, $|\Psi(i, 0)| < \infty$, namely, there are only finitely many graphs in $\Psi(i, 0)$.

As in the proof of Lemma 3.2.3, by linearity of expectation we can deal separately with the expected numbers of subgraphs created and destroyed in a single step. Any new subgraph, which is a union of cycles, in $\Psi(i, 0)$ can be created either by pegging an edge of a short cycle with any other edge (to make a cycle with length increased by 1), or by pegging together the end edges of a short path.

Case 1: One edge in a graph H' in $\Psi(i-1, 0)$ is pegged. (Hence one cycle in H' gets longer.) Since each $H' \in \Psi(i-1, 0)$ contains $i-1$ vertices, thus $i-1$ edges, there are $O(W_{t,i-1,0})$ ways to choose an edge contained in $\Psi(i-1, 0)$ and at most $2n_t$ choices for the other edge to be pegged. The expected number of $\Psi(i, 0)$ arising this way is $O(W_{t,i-1,0}/n_t)$. By the inductive hypothesis that $\mathbf{E}W_{t,i-1,0} = O(1)$, the total expected number of graphs created in $\Psi(i, 0)$ due to this case is $O(1/n_t)$.

Case 2: A new cycle of size at most i is created by pegging two edges within distance i, which, together with a graph in $\Psi(i', 0)$ for some $i' < i$ will form a new graph in $\Psi(i, 0)$. There are $O(n_t)$ paths of length at most i. So the expected number of $\Psi(i, 0)$ created this way is at most $O(W_{t,i',0}/n_t)$. The number of choices of $i' < i$ is bounded. So summing over all possible value of i', and again by induction, the total contribution from this case is $O(1/n_t)$.

Since subgraphs are destroyed if they contain a pegged edge, the expected number of graphs in $\Psi(i, 0)$ destroyed is at least $W_{t,i,0}/n_t$.

Putting it all together, we have

$$\mathbf{E}W_{t+1,i,0} \leq \left(1 - \frac{1}{n_t}\right) \mathbf{E}W_{t,i,0} + O(n_t^{-1}).$$

and hence $\mathbf{E}W_{t,i,0} = O(1)$ by Lemma 3.2.2. By induction, we obtain $\mathbf{E}W_{t,i,0} = O(1)$ for any $i \geq 3$.

Next we fix any $r \geq 1$, and $i \geq 3$, and assume that $\mathbf{E}W_{t,j,r-1} = O(n_t^{-(r-1)})$ for any $j \geq 3$ and $\mathbf{E}W_{t,j,r} = O(n_t^{-r})$ for any $j \leq i$.

We use the same procedure to prove $\mathbf{E}W_{t,i,r} = O(n_t^{-r})$. Consider the expected number of subgraphs in $\Psi(i,r)$ created in going from G_t to G_{t+1}, treating separate cases for creation as above.

Case 1: Similar to the first case above, a subgraph in $\Psi(i,r)$ arises from a subgraph in $\Psi(i-1,r)$, so by induction, we have the total contribution as $O(\mathbf{E}W_{t,i-1,r}/n_t) = O(n_t^{-(r+1)})$.

Case 2: One subcase is that the end edges of a path of length at most i are pegged, which will convert some graph in $\Psi(i',r)$ to one in $\Psi(i,r)$, where $i' < i$. The only other case is that the edges pegged are both within distance i of some graph in $\Psi(j', r-1)$, where $j' < i$. For any fixed subgraph of G_t in $\Psi(j', r-1)$, there are only finite many choices for two such edges to be pegged. So the expected increase in this case will be a sum of a finite number of terms of the form $O(W_{t,i',r}/n_t) + O(W_{t,j',r-1}/n_t^2)$. By induction, $\mathbf{E}W_{t,i',r} = O(n_t^{-r})$ and $\mathbf{E}W_{t,j',r-1} = O(n_t^{-(r-1)})$, and again the total contribution from this case is $O(n_t^{-(r+1)})$.

Analogous to the case of $\Psi(i,0)$, the expected number of subgraphs in $\Psi(i,r)$ destroyed in a single step is at least $W_{t,i,r}/n_t$. Thus

$$\mathbf{E}W_{t+1,i,r} \leq \left(1 - \frac{1}{n_t}\right) \mathbf{E}W_{t,i,r} + O(n_t^{-(r+1)}).$$

Hence $\mathbf{E}W_{t,i,r} = O(n_t^{-r})$ by Lemma 3.2.2. ∎

In later arguments, we especially need to bound the number of subgraphs consisting of two cycles sharing at least one edge. Of course the number of such subgraphs is bounded above by $\sum_{i=1}^{2k} W_{t,i,1}$, where k is length of the longer cycle given. Define $W_{t,k}^* = \sum_{i=1}^{2k} W_{t,i,1}$. Since $\mathbf{E}W_{t,i,1} = O(n_t^{-1})$, and the summation is taken over finitely many values of i, the following comes immediately from Lemma 3.2.4.

Corollary 3.2.5 $\mathbf{E}W_{t,k}^* = O(n_t^{-1})$.

Gearing up for the proof of Theorem 3.2.1, we next give some simple lemmas bounding some rare events. Let $\mathbf{Y}_t^{(l)} := (Y_{t,3}, Y_{t,4}, \ldots, Y_{t,l})$. In the following lemmas, the choice of the norm $\|\mathbf{Y}_t^{(l)}\|$ does not change the strength of the statement, and one may for instance settle on the L^∞ norm.

Lemma 3.2.6 *Fix the graph G_t. For any fixed $k \geq 3$, the probability that more than one cycle of length at most $k+1$ in G_{t+1} contains the peg vertex is $O(\|\mathbf{Y}_t^{(2k)}\|^2/n_t^2 + W_{t,k}^*/n_t)$.*

Proof There are several cases to consider. The first case is that one edge pegged is contained in more than one cycle of length at most k, so that at least two cycles of length at most k will pass through the peg vertex. Since the subgraph consisting of two cycles of length at most k sharing common edges is involved, the probability this happens is at most $O(W^*_{t,k}/n_t)$. The second case is that one edge pegged is contained in a cycle of length at most k, and the other edge pegged is of distance at most k from the first edge. In this case, a new cycle is created using the path joining the pegged edges. There are at most $O(\|\mathbf{Y}_t^{(k)}\|)$ ways to choose the first edge, and for each such choice, there are at most $2d^k = O(1)$ ways to choose the second edge. So the probability that this case happens is $O(\|\mathbf{Y}_t^{(k)}\|/n_t^2)$. The third case is that the two edges pegged are both contained in some cycle of length at most $2k$. The probability this happens is $O(\|\mathbf{Y}_t^{(2k)}\|/n_t^2)$, since there are at most $O(\|\mathbf{Y}_t^{(2k)}\|)$ ways to choose such two edges. The fourth case is that each of the two edges pegged is contained in a cycle of length at most k. The probability for this to happen is $O(\|\mathbf{Y}_t^{(k)}\|^2/n_t^2)$. Then Lemma 3.2.6 follows. ∎

We will use Lemma 3.2.6 to show that the only significant things that can happen with respect to short cycles are (a) an edge of a short cycle is pegged and no other cycles are created or destroyed, or (b) a short cycle is created by pegging the ends of a short path and no other short cycles are created or destroyed.

Note that a cycle is destroyed only if at least one of its edges is pegged. So to create or destroy more than one k-cycle in one step, there must be at least two cycles of length at most $k+1$ containing the peg vertex after the pegging operation has been applied. Hence the following result comes immediately from Lemma 3.2.6.

Corollary 3.2.7 $\mathbf{P}(|Y_{t+1,k} - Y_{t,k}| > 1 \mid \mathbf{Y}_t^{(2k)}, W^*_{t,k}) = O\big(\|\mathbf{Y}_t^{(2k)}\|^2/n_t^2 + W^*_{t,k}/n_t\big)$.

By taking expectation of both sides of the equation in the above corollary, and by Lemma 3.2.3 and Corollary 3.2.5, we have the following corollary.

Corollary 3.2.8 $\mathbf{P}(|Y_{t+1,k} - Y_{t,k}| > 1) = O\left(1/n_t^2\right)$.

Similarly, we may bound the simultaneous creation and destruction of cycles, except for a special case. (The following bounds are sufficient for our purposes and can easily be improved by examining the cases in the proof of Lemma 3.2.6.)

Corollary 3.2.9 *For any fixed integers l_1, $l_2 \geq 3$, such that $l_1 \neq l_2 + 1$, the probability of creating a new l_1-cycle and simultaneously destroying an existing l_2-cycle in the same step is $O(\|\mathbf{Y}_t^{(k)}\|^2/n_t^2) + O(W^*_{t,k}/n_t)$, where $k = \max\{l_1, l_2\}$.*

Proof The peg vertex is contained in the l_1-cycle that is created, and also in at least one of the edges in the l_2-cycle that is destroyed. If only one edge in this l_2-cycle is pegged, then a new (l_2+1)-cycle is created which contains the peg vertex. Since $l_1 \neq l_2+1$, the peg vertex is contained in both the l_1-cycle and the (l_2+1)-cycle. By Lemma 3.2.6, this happens with probability $O(\|\mathbf{Y}_t^{(k)}\|^2/n_t^2) + O(W_{t,k}^*/n_t)$. If two edges in this l_2-cycle are pegged, then two short cycles containing the peg vertex and the rest of the edges of this l_2-cycle are created. By Lemma 3.2.6, this happens with probability $O(\|\mathbf{Y}_t^{(k)}\|^2/n_t^2) + O(W_{t,k}^*/n_t)$. Thus Corollary 3.2.9 follows. ∎

Next we estimate the moments $\mathbf{E}[Y_{t,3}]_j$ of $Y_{t,3}$, for any fixed $j \geq 0$. We set $Y_{t,2} = 0$ for any t, since the random graph generated is simple.

Lemma 3.2.10 *For any fixed nonnegative integer j,*

$$\mathbf{E}([Y_{t,3}]_j) = 3^j + O(n_t^{-1}).$$

Proof The proof is by induction on j. The base case $j = 0$ is trivial. Assume that $j \geq 1$ and $\mathbf{E}([Y_{t,3}]_{j-1}) \to 3^{j-1}$.

Instead of calculating $[Y_{t,3}]_j$ directly, we will calculate $[Y_{t,3}]_j/j!$, which is the number of j-sets of distinct i-cycles. We first consider the creation of a new j-set of triangles in moving from G_t to G_{t+1}, beginning with the j-sets that use an existing $(j-1)$-set of triangles, together with one newly created triangle.

We know that the expected number of triangles created at step t is $9/n_t + O(Y_{t,3})/n_t^2$. Each such new triangle creates a new j-set with each $(j-1)$-set of existing triangles except for those that simultaneously have one of their triangles destroyed. So the expected number of new j-sets created this way is

$$\left(\frac{9 + O(Y_{t,3}/n_t)}{n_t}\right)\frac{[Y_{t,3}]_{j-1}}{(j-1)!} + O\left(\frac{Y_{t,3}^2}{n_t^2} + \frac{W_{t,3}^*}{n_t}\right)\frac{[Y_{t,3}]_{j-1}}{(j-1)!}.$$

Here, the first term arises from the assumption that no existing triangles in the $(j-1)$-set are destroyed when the new triangle is created. The second, purely error term bounds the expected number of j-sets counted in the main term that should be discounted because, simultaneously with the new triangle being created, one of the triangles in the existing $(j-1)$-set is destroyed. The factor $O(Y_{t,3}^2/n_t^2 + W_{t,3}^*/n_t)$ comes from Corollary 3.2.9 for the probability of simultaneously creating and destroying triangles, and is multiplied by a bound on how many $(j-1)$-sets of existing triangles can contain one of the (bounded number of) triangles destroyed.

There are also j-sets that include more than one newly created triangle. It is straightforward to observe that in one step it is possible to create at most four triangles, and destroy at most six. By Corollary 3.2.7, the probability of creating more than one triangle in one step, given $Y_{t,3}$ and $Y_{t,4}$, is $O(\|\mathbf{Y}_t^{(4)}\|^2/n_t^2 + W^*_{t,3}/n_t)$. Hence, the expected number of new j-sets created this way is at most

$$O\left(\frac{\|\mathbf{Y}_t^{(4)}\|^2}{n_t^2} + \frac{W^*_{t,3}}{n_t}\right)\sum_{i=2}^{6}\frac{[Y_{t,3}]_{j-i}}{(j-i)!}.$$

Note that $W^*_{t,3}[Y_{t,3}]_{j-i}$ is bounded above by $W_{t,3(j-i)+4,1}$ (representing the structures that come from the union of a structure in $\Psi(3(j-i)+4,1)$). By Lemma 3.2.4 the expected number of such complex structures of bounded size is $O(n_t^{-1})$. Thus, using also the first part of Lemma 3.2.4,

$$\begin{aligned}
\mathbf{E}\left(\left(\frac{Y_{t,3}^2}{n_t^2} + \frac{W^*_{t,3}}{n_t}\right)[Y_{t,3}]_{j-2}\right) &= O(n_t^{-2}),\\
\mathbf{E}\left(O\left(\frac{\|\mathbf{Y}_t^{(4)}\|^2}{n_t^2} + \frac{W^*_{t,3}}{n_t}\right)\sum_{i=2}^{6}[Y_{t,3}]_{j-i}/(j-i)!\right) &= O(n_t^{-2}).
\end{aligned}$$

Now consider destroying an existing j-set. Firstly, assuming the j triangles are disjoint, then pegging any edge contained in those edges with any other non-adjacent edge will destroy the j-set. It follows that the expected number of j-sets being destroyed, given $Y_{t,3}$, is

$$\frac{3j[Y_{t,3}]_j/j!}{n_t} + \sum_{i'\le 3j}\frac{O(W_{t,i',1})}{n_t}.$$

The error term $O(W_{t,i',1}/n_t)$ accounts for the case that the j-set of triangles share some common edges. So by Lemma 3.2.4

$$\mathbf{E}\left(\sum_{i'\le 3j}\frac{O(W_{t,i',1})}{n_t}\right) = O(n_t^{-2}).$$

Thus

$$\begin{aligned}
\mathbf{E}([Y_{t+1,3}]_j/j!) - \mathbf{E}([Y_{t,3}]_j/j!) &= \left(\frac{9+O(n_t^{-1})}{n_t}\right)\mathbf{E}([Y_{t,3}]_{j-1}/(j-1)!)\\
&\quad -\frac{3j\mathbf{E}([Y_{t,3}]_j/j!) + O(n_t^{-1})}{n_t} + O(n_t^{-2}).
\end{aligned}$$

By the inductive assumption, we have $\mathbf{E}([Y_{t,3}]_i) = 3^i + O(n_t^{-1})$ for any $i \le j-1$. Then

$$\mathbf{E}([Y_{t+1,3}]_j/j!) = \left(1 - \frac{3j}{n_t}\right)\mathbf{E}([Y_{t,3}]_j/j!) + \left(\frac{9\cdot 3^{j-1}}{(j-1)!n_t}\right) + O(n_t^{-2}).$$

Applying Lemma 3.2.2 with $c = 3j$ and $p = 1$ gives $\mathbf{E}([Y_{t,3}]_j) = 3^j + O(n_t^{-1})$ as required. ∎

Proof of Theorem 3.2.1. By Theorem 2.1.1, it is enough to show that, for any fixed constant $k \geq 3$, and a given sequence of nonnegative integers $(j_3, j_4, \ldots, j_k)$,

$$\lim_{t\to\infty} \mathbf{E}([Y_{t,3}]_{j_3}[Y_{t,4}]_{j_4} \cdots [Y_{t,k}]_{j_k}) = \prod_{i=3}^{k} u_i^{j_i}.$$

We prove this by induction on the sequence of $(j_3, j_4, \ldots, j_k)$. The base case is $(j_3, 0, \ldots, 0)$, for any nonnegative integer j_3. Lemma 3.2.3 shows that $\mathbf{E}(Y_{t,j_3}) \to \mu_3^{j_3}$.

Let $\mathcal{S}(j_3, j_4, \ldots, j_k)$ denote the family of all collections (in whatever graph is under consideration) consisting of a j_3-set of distinct 3-cycles, a j_4-set of distinct 4-cycles,..., and a j_k-set of distinct k-cycles. Let $\#(t, j_3, j_4, \ldots, j_k)$ be the number of elements of $\mathcal{S}(j_3, j_4, \ldots, j_k)$ in G_t. Note that

$$\#(t, j_3, j_4, \ldots, j_k) = \prod_{i=3}^{k} \frac{[Y_{t,i}]_{j_i}}{j_i!}.$$

We will estimate

$$\Delta := \mathbf{E}\left(\prod_{i=3}^{k} \frac{[Y_{t+1,i}]_{j_i}}{j_i!}\right) - \mathbf{E}\left(\prod_{i=3}^{k} \frac{[Y_{t,i}]_{j_i}}{j_i!}\right) = \mathbf{E}\#(t+1, j_3, j_4, \ldots, j_k) - \mathbf{E}\#(t, j_3, j_4, \ldots, j_k).$$

Case 1: Analogous to the creation of new triangles considered in the proof of Lemma 3.2.10, if a new i-cycle is created by pegging together the end edges of an i-path, then a new element of $\mathcal{S}(j_3, j_4, \ldots, j_k)$ can be created from an existing element of $\mathcal{S}(j_3, \ldots, j_i{-}1, \ldots, j_k)$ together with the new i-cycle. The argument is similar to the proof of Lemma 3.2.10, and we omit the precise error terms since they have a similar nature and can be bounded in the same way. Instead we find that the contribution to Δ is

$$\sum_{i=3}^{k} \frac{3^{i-1}}{n_t} \mathbf{E}\left(\#(t, j_3, \ldots, j_i - 1, \ldots, j_k)\right) + O(n_t^{-2}).$$

Here we use the convention that $[x]_{-1} = 0$ for all x.

Case 2: If an edge of an $(i-1)$-cycle is pegged, for $i \leq k$, then a new element of $\mathcal{S}(j_3, j_4, \ldots, j_k)$ can be created in several ways. The typical way is from an element of

$\mathcal{S}(j_3,\ldots,j_{i-1}+1,j_i-1,\ldots,j_k)$, for some $4 \le i \le k$, that contains the $(i-1)$-cycle pegged. The expected number of elements of $\mathcal{S}(j_3,j_4,\ldots,j_k)$ created in this way is

$$\sum_{i=4}^{k} \frac{(i-1)(j_{i-1}+1)+O(n_t^{-1})}{n_t} \#(t,j_3,\ldots,j_{i-1}+1,j_i-1,\ldots,j_k)$$
$$+O\Big(\sum W_{t,i',0}/n_t^2 + \sum W_{t,i',1}/n_t\Big).$$

The error term $O(n^{-2})$ accounts for the approximation for the number of possible peggings as before. The event that the two edges pegged are both in short cycles is accounted for by $O\big(\sum W_{t,i',0}/n_t^2\big)$. This also accounts for the case that a new short cycle is created as in Case 1 at the same time that an edge of a short cycle is pegged. The final error term accounts for the case that cycles in an element of $\mathcal{S}(j_3,\ldots,j_{i-1}+1,j_i-1,\ldots,j_k)$ share common edges, one of which is pegged; these cases should be discounted. It also accounts for other atypical ways to produce an element of $\mathcal{S}(j_3,\ldots,j_k)$, where an edge in two or more short cycles is pegged. These sums are taken over finitely many possible i'. By Lemma 3.2.4, the expected value of the error terms is $O(n_t^{-2})$.

An existing $\mathcal{S}(j_3,j_4,\ldots,j_k)$ can be destroyed by pegging any of its edges. Arguing as in the proof of Lemma 3.2.10, the contribution to Δ from destroying these configurations is

$$-\left(\sum_{i=3}^{k} i j_i\right)\frac{1}{n_t}\mathbf{E}\left(\#(t,j_3,j_4,\ldots,j_k)\right)+O(n_t^{-2}).$$

So we get

$$\begin{aligned}\Delta &= \sum_{i=3}^{k}\frac{3^{i-1}}{n_t}\mathbf{E}\left(\#(t,j_3,\ldots,j_i-1,\ldots,j_k)\right)\\ &\quad+\sum_{i=4}^{k}\frac{(i-1)(j_{i-1}+1)}{n_t}\mathbf{E}\left(\#(t,j_3,\ldots,j_{i-1}+1,j_i-1,\ldots,j_k)\right)\\ &\quad-\left(\sum_{i=3}^{k} i j_i\right)\frac{1}{n_t}\mathbf{E}\left(\#(t,j_3,j_4,\ldots,j_k)\right)+O(n_t^{-2}).\end{aligned}$$

By induction,

$$\mathbf{E}\left(\#(t,j_3,\ldots,j_i-1,\ldots,j_k)\right)\to\frac{u_3^{j_3}}{j_3!}\cdots\frac{u_i^{j_i-1}}{(j_i-1)!}\cdots\frac{u_k^{j_k}}{j_k!}\qquad\text{for all } 3\le i\le k.$$

$$\mathbf{E}\left(\#(t,j_3,\ldots,j_{i-1}+1,j_i-1,\ldots,j_k)\right)\to\frac{u_3^{j_3}}{j_3!}\cdots\frac{u_{i-1}^{j_{i-1}+1}}{(j_{i-1}+1)!}\frac{u_i^{j_i-1}}{(j_i-1)!}\cdots\frac{u_k^{j_k}}{j_k!}$$

for all $4 \le i \le k$. So arguing as in the proof of Lemma 3.2.10, we set $\Delta = 0$ and obtain

$$\begin{aligned}
\mathbf{E}\left(\#(t, j_3, j_4, \ldots, j_k)\right) &\to \left(\frac{1}{\sum_{i=3}^k i j_i}\right)\left(\sum_{i=3}^k 3^{i-1} \frac{u_3^{j_3}}{j_3!} \cdots \frac{u_i^{j_i-1}}{(j_i-1)!} \cdots \frac{u_k^{j_k}}{j_k!}\right. \\
&\quad \left. + \sum_{i=4}^k (i-1)(j_{i-1}+1) \frac{u_3^{j_3}}{j_3!} \cdots \frac{u_{i-1}^{j_{i-1}+1}}{(j_{i-1}+1)!} \frac{u_i^{j_i-1}}{(j_i-1)!} \cdots \frac{u_k^{j_k}}{j_k!}\right) \\
&= \prod_{i=3}^k \frac{u_i^{j_i}}{j_i!} \left(\frac{1}{\sum_{i=3}^k i j_i}\right)\left(\sum_{i=3}^k 3^{i-1} \frac{j_i}{\mu_i} + \sum_{i=4}^k (i-1)(j_{i-1}+1) \frac{\mu_{i-1}}{j_{i-1}+1} \frac{j_i}{\mu_i}\right).
\end{aligned}$$

We only need to prove that

$$\sum_{i=3}^k 3^{i-1} \frac{j_i}{\mu_i} + \sum_{i=4}^k (i-1)(j_{i-1}+1) \frac{\mu_{i-1}}{j_{i-1}+1} \frac{j_i}{\mu_i} = \sum_{i=3}^k i j_i.$$

By calculating the left hand side, we get

$$\begin{aligned}
&\sum_{i=3}^k 3^{i-1} \frac{j_i}{\mu_i} + \sum_{i=4}^k (i-1)(j_{i-1}+1) \frac{\mu_{i-1}}{j_{i-1}+1} \frac{j_i}{\mu_i} \\
&\quad = \sum_{i=3}^k \frac{2 i j_i}{3^i - 9} 3^{i-1} + \sum_{i=4}^k (i-1) j_i \frac{3^{i-1}-9}{2(i-1)} \frac{2i}{3^i - 9} \\
&\quad = \sum_{i=3}^k \frac{2 i j_i}{3^i - 9} 3^{i-1} + \sum_{i=4}^k \frac{2 i j_i}{3^i - 9} \frac{3^{i-1}-9}{2} \\
&\quad = \sum_{i=3}^k i j_i.
\end{aligned}$$

So we have shown that

$$\mathbf{E}\left(\prod_{i=3}^k \frac{[Y_{t,i}]_{j_i}}{j_i!}\right) \to \prod_{i=3}^k \frac{\mu_i^{j_i}}{j_i!},$$

and hence

$$\mathbf{E}\left(\prod_{i=3}^k [Y_{t,i}]_{j_i}\right) \to \prod_{i=3}^k \mu_i^{j_i}.$$

Theorem 3.2.1 then follows. ∎

3.3 Rate of convergence

In this section, we study the rate at which the joint distribution of $Y_{t,3}, \dots, Y_{t,k}$ converges to its limit for any constant $k \geq 3$ in the process $\mathcal{P}(G_0, 4)$. The standard definition of the mixing time $\tau(\epsilon)$ of a Markov chain with state space $\mathcal{S}$ is the minimum time t, such that after t steps, the total variation distance (defined in Section 2.1.4. See (2.1.2)) between P_x^t, the distribution at time t starting from state x, and the stationary distribution π, is at most pre-specified constant ϵ, $0 < \epsilon < 1$. Formally,

$$\tau(\epsilon) = \max_{x \in \mathcal{S}} \min\{T : d_{TV}\left(P_x^t, \pi\right) \leq \epsilon \text{ for all } t \geq T\}.$$

Recall that $Y_{t,k}$ denotes the number of k-cycles in $G_t \in \mathcal{P}(G_0)$. In practice, for mixing time results one chooses a fixed value of ϵ (for instance, $\epsilon = 1/4$), and obtains results such as "the mixing time is $O(n \log n)$", where n refers to the size (the number of states) of a Markov chain. This makes sense if the mixing time is a logarithmic function of ϵ, in which case $\tau(\epsilon^k)$ can be approximately estimated by $k\tau(\epsilon)$. This case usually occurs when the Markov chain is time-homogeneous, i.e. the transition probabilities in every step are invariant. Hence, given the value of ϵ, the mixing time, determined by the probability transition matrix, is only a function of n. However, the random process $(Y_{t,k})_{t\geq 0}$ does not belong to this category. As we will see later, the transition probabilities of the process depend on the time t, which makes it unlikely to have the mixing time as a logarithmic function of ϵ. Furthermore, the size of G_t is a linearly increasing function of t, making the size of $(Y_{t,k})_{t\geq 0}$ grow in the whole process. Thus, we cannot define a mixing time as a function of the size of the process. Indeed, the random process $(Y_{t,k})_{t\geq 0}$, for some constant k, is not a Markov chain, since the distribution of $Y_{t,k}$ depends not only on $Y_{t-1,k}$, but also on the underlying graph G_{t-1}, and as a result is not independent of $Y_{t-2,k}$ given $Y_{t-1,k}$. Instead, we wish to consider the total variation distance between the random variable $Y_{t,k}$ and the limiting distribution of $Y_{t,k}$ (if it exists). Therefore, we consider ϵ-mixing time instead, which is informally defined in Section 2.1.4. We give a formal definition as follows. Let $(\sigma_t)_{t\geq 0}$ be a sequence of distributions which converge to the distribution π_k^*. The *ϵ-mixing time* of $(\sigma_t)_{t\geq 0}$ is

$$\tau_\epsilon^*\big((\sigma_t)_{t\geq 0}\big) = \min\{T \geq 0 : \; d_{TV}\left(\sigma_t, \pi_k^*\right) \leq \epsilon \text{ for all } t \geq T \;\}. \tag{3.3.1}$$

We now focus on a particular sequence of distributions. For any fixed k, let $\sigma_{t,k}$ denote the joint distribution of $Y_{t,3}, \dots, Y_{t,k}$. The next theorem from [27, Theorem 2.2] shows that the ϵ-mixing time of $(\sigma_{t,k})_{t\geq 0}$ is $O(1/\epsilon)$. We cite this theorem because it is used a few times in this chapter to prove other results.

Theorem 3.3.1 *For fixed G_0 and $k \geq 3$, the ϵ-mixing time of the sequence of short cycle joint distributions in $\mathcal{P}(G_0, 4)$ satisfies*

$$\tau_\epsilon^* \left((\sigma_{t,k})_{t \geq 0} \right) = O \left(\epsilon^{-1} \right).$$

The proof uses the coupling method introduced in Section 2.1.4. The idea of the proof is summarized as follows. For any given integer $k \geq 3$, they study the random process $(\mathbf{Y}_{t,k})_{t \geq 0} = (Y_{t,3}, \ldots, Y_{t,k})_{t \geq 0}$. Note that $(\mathbf{Y}_{t,k})_{t \geq 0}$ is not a Markov chain. They first construct a Markov chain $(\mathbf{Z}_{t,k})_{t \geq 0} = (Z_{t,3}, \ldots, Z_{t,k})_{t \geq 0}$ such that the distribution of $\mathbf{Y}_{t,k}$ is the same as $\mathbf{Z}_{t,k}$ for any $t \geq 0$. Then they construct another Markov chain $(\mathbf{X}_{t,k})_{t \geq 0} = (X_{t,3}, \ldots, X_{t,k})_{t \geq 0}$ whose transition probabilities are close to those of $(\mathbf{Z}_{t,k})_{t \geq 0}$. Let $(\mathbf{X}_{t,k})_{t \geq 0}$ start from its stationary distribution. Then they construct a coupling of the two Markov chains $(\mathbf{Z}_{t,k})_{t \geq 0}$ and $(\mathbf{X}_{t,k})_{t \geq 0}$, from which they obtain an upper bound of the ϵ-mixing time of $(\mathbf{Y}_{t,k})_{t \geq 0}$ by applying Lemma 2.1.2, which gives Theorem 3.3.1.

The main goal of this section is to show the tightness of the upper bound in Theorem 3.3.1. The following theorem shows that the upper bound of ϵ-mixing time given by Theorem 3.3.1 is eventually tight.

Theorem 3.3.2 *For fixed G_0 and $k \geq 3$, the ϵ-mixing time of the sequence of short cycle joint distributions in $\mathcal{P}(G_0, 4)$ satisfies $\tau_\epsilon^* \left((\sigma_{t,k})_{t \geq 0} \right) \neq o \left(\epsilon^{-1} \right)$.*

By considering just the events measurable in the σ-algebra generated by $Y_{t,3}$, or equivalently, by considering only the events $Y_{t,3} = j$, $j \geq 0$, and any union of these events, we see immediately that

$$d_{TV} \left(\sigma_{t,3}, \pi_3 \right) \leq d_{TV} \left(\sigma_{t,k}, \pi_k \right)$$

where π_k is the limit of $\sigma_{t,k}$. Hence, it suffices to show that the ϵ-mixing time for $\sigma_{t,3}$, which is the distribution of $Y_{t,3}$, is not $o(\epsilon^{-1})$. For convenience, in the rest of this section we use the notation Y_t to denote $Y_{t,3}$ as in the proof of Theorem 3.3.1.

Let C_4^* denote the graph consisting of a 4-cycle plus a chord (i.e. K_4 minus an edge), and let W_t denote the number of subgraphs of G_t that are isomorphic to C_4^*. Lemma 3.2.4 implies that a.a.s. $W_t = 0$. That is, a.a.s. all triangles are isolated, where an *isolated triangle* is a 3-cycle that shares no edges with any other 3-cycle. In order to prove Theorem 3.3.2, we also need more information on the distribution of the number of isolated triangles in the presence of one copy of C_4^*. In the following lemma, we show that this has the same asymptotic distribution as Y_t. This distribution is to be expected, since the creation of a copy of C_4^* will leave an asymptotically Poisson number of isolated triangles. Until the C_4^* disappears due to some pegging operation, this Poisson number of isolated triangles will undergo transitions with similar rules to Y_t and will therefore remain asymptotically

Poisson. Instead of fleshing this argument out into a proof, it seems simpler to provide a complete argument using the method of moments, although this conceals the coincidence to a greater extent.

Lemma 3.3.3 *Conditional on $W_t = 1$, the random variable $Y_t - 2$ has a limiting distribution that is Poisson with mean* 3.

Proof Let $U_{t,j}$ denote $[Y_t - 2]_j I\{W_t = 1\}$, i.e. the product of the j-th falling factorial of $Y_t - 2$ and the indicator random variable of the event that $W_t = 1$. Note that if we can show

$$\mathbf{E}(U_{t,j}) \to 3^j\, \mathbf{P}(W_t = 1), \tag{3.3.2}$$

then $\mathbf{E}([Y_t - 2]_j \mid W_t = 1) \to 3^j$. Lemma 3.3.3 then follows by the method of moments applied to the probability space obtained by conditioning on $W_t = 1$. So we only need to compute $\mathbf{P}(W_t = 1)$ and $\mathbf{E}(U_{t,j})$. We show that $\mathbf{P}(W_t = 1) = 27/(4n_t) + O(n_t^{-2})$, and show by induction on j that

$$\mathbf{E}(U_{t,j}) = \frac{27}{4n_t} 3^j + O(n_t^{-2}), \tag{3.3.3}$$

for any integer $j \geq 0$. This gives (3.3.2) as required.

Consider $\mathbf{P}(W_t = 1)$ first. Our way of estimating this quantity is by computing separately the expected numbers of copies of C_4^* that are created, or destroyed, in each step. There are two ways to create a C_4^*. One way is through the creation of a new triangle which shares an edge with an existing triangle, which we will call C. This requires two edges adjacent to different vertices of C (but not being edges of C) to be pegged. This is illustrated in Figure 3.2, where v is the peg vertex, and the two dashed edges e_1 and e_2 are pegged. Given C, if C is an isolated triangle, there are exactly 12 ways to choose such two edges. Otherwise, C is part of an existing C_4^* and the number of pegging operations using such a type of C is $O(W_t)$. Overall, the expected number of C_4^* created in this way is therefore $\big(12\hat{Y}_t + O(W_t)\big)/N_t$, where $\hat{Y}_t$ is the number of isolated triangles in G_t. The other way of creating a C_4^* from a triangle C is as illustrated in Figure 3.3, where e_1 is an edge in C, and e_2 is incident with some vertex of C, but not adjacent to e_1. Given C, there are 3 ways to choose e_1, and for each chosen e_1, there are 2 ways to choose e_2. Hence, there are 6 ways to choose the pair (e_1, e_2), and the expected number of C_4^* created in this way is $6Y_t/N_t$.

Clearly $Y_t = \hat{Y}_t + O(W_t)$. So the expected number of C_4^* created in each step is $18\hat{Y}_t/N_t + O(W_t/N_t) = 9Y_t/n_t^2 + O(n_t^{-3}) + O(W_t n_t^{-2})$.

The expected number of C_4^* destroyed in each step is easily seen to be $5W_t(2n_t - 7)/N_t = 5W_t/n_t$. Thus

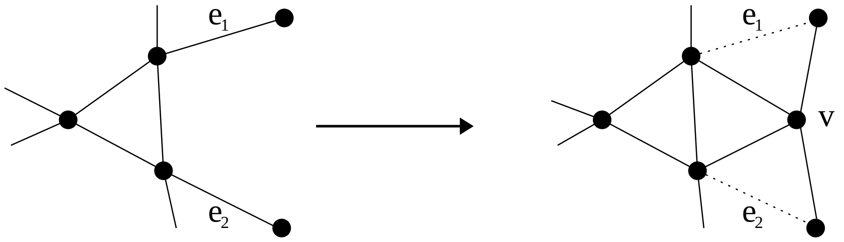

Figure 3.2: *pegging operation to create a* C_4^**, first case*

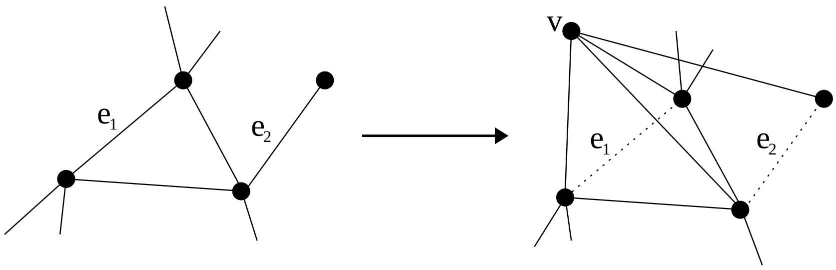

Figure 3.3: *pegging operation to create a* C_4^**, second case*

$$\mathbf{E}(W_{t+1} - W_t \mid W_t) = \frac{9Y_t}{n_t^2} - \frac{5W_t}{n_t} + O(W_t n_t^{-2} + n_t^{-3}).$$

Taking expected values and using the tower property of conditional expectation, this gives

$$\mathbf{E}W_{t+1} - \mathbf{E}W_t = \frac{9\mathbf{E}Y_t}{n_t^2} - \frac{5\mathbf{E}W_t}{n_t} + O(\mathbf{E}W_t n_t^{-2} + n_t^{-3}).$$

Since $\mathbf{E}Y_t = 3 + O(n_t^{-1})$, and $\mathbf{E}W_t = O(n_t^{-1})$, this yields

$$\mathbf{E}W_{t+1} = \left(1 - \frac{5}{n_t}\right)\mathbf{E}W_t + \frac{27}{n_t^2} + O(n_t^{-3}).$$

Applying Lemma 3.2.10 and Lemma 3.2.2 with $c = 5$, $p = 2$ and $\rho = 27$, we obtain that $\mathbf{E}W_t = 27/(4n_t) + O(n_t^{-2})$. Since $\mathbf{P}(W_t = i) \leq \mathbf{E}([W_t]_i) = O(n_t^{-i})$ by Lemma 3.2.4,

$$\mathbf{P}(W_t = 1) = 27/(4n_t) + O(n_t^{-2}). \tag{3.3.4}$$

Next we compute $\mathbf{E}(U_{t,j})$ by induction on $j \geq 0$. The base case is $j = 0$, for which we begin by noting that $\mathbf{E}(U_{t,0}) = \mathbf{P}(W_t = 1) = 27/(4n_t) + O(n_t^{-2})$ as shown above. Now assume that $j \geq 1$ and that (3.3.3) holds for all smaller values of j. Given the graph G_t, the expected change in $U_{t,j}/j!$ when t changes to $t+1$ is, as explained below,

$$\begin{aligned}\mathbf{E}\left(\frac{U_{t+1,j}}{j!}-\frac{U_{t,j}}{j!}\,\Big|\, G_t\right) &= \left(\left(\frac{9+O((1+Y_t+Y_{t,4})/n_t)}{n_t}\right)\frac{[Y_t-2]_{j-1}}{(j-1)!}\right) I\{W_t=1\}\\ &\quad+\left(\frac{9}{n_t^2}+O\left(n_t^{-3}\right)\right)\frac{(j+1)[Y_t]_{j+1}}{(j+1)!}I\{W_t=0\}\\ &\quad+f(j,G_t)\\ &\quad-\left(\frac{(3j+5)[Y_t-2]_j/j!}{n_t}+O(n_t^{-2})\right)I\{W_t=1\}, \qquad (3.3.5)\end{aligned}$$

where $f(j,G_t)$ denotes some assorted "error" terms described below. Note that, given $W_t=1$, $[U_{t,1}]_j/j!$ is simply the number of subgraphs of G_t containing precisely j isolated triangles, so we may just compute the change in the number of such subgraphs in those cases where no copies of C_4^* are created or destroyed. The first term on the right in (3.3.5) is the positive contribution when $W_t=1$ and the pegging step creates one new isolated triangle. Any set of $j-1$ isolated triangles, together with the new triangle, can potentially form a new set of j isolated triangles. A new triangle is created from pegging the two end-edges of a 3-path, the number of which in G_t is $4\cdot 3\cdot 3\cdot n_t/2+O(Y_t)=18n_t+O(Y_t)$. Dividing this by N_t gives rise to the main term. The error term $O(1+Y_t+Y_{t,4})$ accounts for choices of such edges which, when pegged, create two or more triangles (when both edges pegged are contained in a 4-cycle) or cause some existing triangle, including possibly the C_4^*, to be destroyed, or cause the new triangle or an existing one not to be isolated.

The second term on the right in (3.3.5) accounts for the contribution when $W_t=0$ due to the creation of a C_4^*, when the set of j isolated triangles are all pre-existing. We have noted above that a new C_4^* can be created only from a triangle. So, when $W_t=0$, a positive contribution to $U_{t+1,j}-U_{t,j}$ can arise from each set of $j+1$ isolated triangles, such that a new C_4^* comes from pegging near one of these triangles as in Figure 3.2 and 3.3. There are $[Y_t]_{j+1}/(j+1)!$ different $(j+1)$-sets of triangles, and for each $(j+1)$-set, there are $j+1$ ways to choose one particular triangle. There are 18 ways to peg two edges to create a C_4^* from any given triangle. This, together with $N_t=2n_t^2(1+O(n_t^{-1}))$, explains the significant part of this term and the first error term. There is also a correction required when the pegging that creates a C_4^* also "accidentally" destroys one or more of the other triangles in the $(j+1)$-set. This occurs only if the two triangles destroyed are near each other, so they create a small subgraph with more edges than vertices. This correction term is a sum of terms of the form $[Y_t]_{j'}W_{t,i',1}/n_t^2$ for a few different values of i' and j', whose expected value is $O(n_t^{-3})$.

The third term, $f(j,G_t)$, is a function that accounts for all other positive contributions, i.e. counts all other cases of newly created sets of j isolated triangles together with a copy of C_4^*. The situations included here are those in which

(a) $W_t = 1$ and $j' \geq 2$ new triangles are created, which only happens if both edges pegged are contained in a 4-cycle, contributing $O(I\{W_t = 1\}[Y_t]_{j-j'}Y_{t,4}/n_t^2)$, or

(b) $W_t = 1$, the copy of C_4^* is destroyed (leaving behind a new isolated triangle) and simultaneously another is created, contributing $O(I\{W_t = 1\}[Y_t]_{j-1}/n_t^2)$ or

(c) $W_t \geq 2$, and all but one of the copies of C_4^* are destroyed, possibly creating a number of isolated triangles and possibly destroying one. This contributes terms of the form $O(I\{W_t \geq 2\}[Y_t]_{j'}/n_t)$ for various $j' \leq j+1$, or

(d) $W_t = 0$, a C_4^* is created along with an isolated triangle, which is contained in the set of j isolated triangles. When this happens, there must be a triangle sharing a common edge with a 4-cycle, so that the triangle turns into C_4^* when two edges of the 4-cycle are pegged, whilst the other edge of the 4-cycle together with two new edges forms an isolated triangle. Figure 3.4 illustrates how this works. This case contributes $O(I\{W_t = 0\}[Y_t]_{j-1}W_{t,5,1}/n_t^2)$.

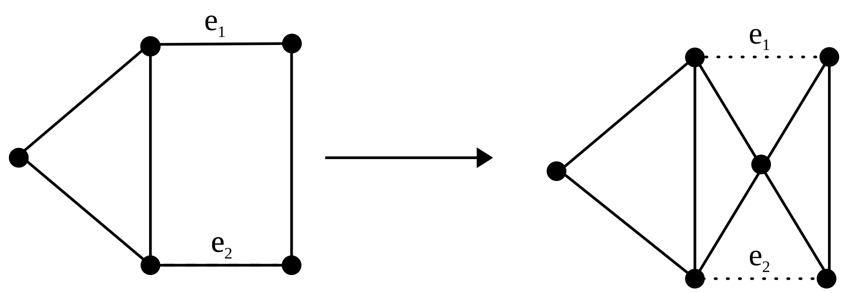

Figure 3.4: *pegging operation to create a C_4^* and a new triangle.*

We note here for later use that each of these cases involves a subgraph with excess at least 1, and at least 2 in the case (c). For instance $I\{W_t = 1\}[Y_t]_{j-j'}Y_{t,4} \leq W_t[Y_t]_{j-j'}Y_{t,4}$ counts subgraphs with $j - j'$ distinct triangles, a 4-cycle and a copy of C_4^*. Such subgraphs have at most $3(j - j') + 8$ vertices and excess at least 1. By Lemma 3.2.4, the expected number of such subgraphs is $O(n_t^{-1})$. Using this argument, we find that $\mathbf{E}(f(j, G_t)) = O(n_t^{-3})$.

The last term in (3.3.5) accounts for the negative contribution to $U_{t+1,j} - U_{t,j}$. Let F_i be the class of subgraphs consisting of i isolated triangles, for some fixed i. Then $U_{t,j}/j!$ counts the number of copies of subgraphs of G_t that are contained in F_j if $W_t = 1$, and is counted as 0 if $W_t \neq 1$. The negative contribution comes when an edge contained in some copy of a member of F_j is destroyed, or an edge contained in the C_4^* is destroyed. In the first case, each copy of an $f \in F_j$ in G_{t+1} that is destroyed contributes -1. The number of subgraphs of G_t that are in F_j is $[Y_t - 2]_j/j!$, and for each copy there are $3j$ ways to choose an edge. Hence the expected contribution of this case is $-3j[Y_t-2]_j/(j!n_t)$. In the second case, the destruction of C_4^* kills the contribution of any copy of $f \in F_j$ to $U_{t+1,j}$, since W_{t+1} becomes 0. Hence the negative contribution is $-[Y_t-2]_j/j!$, the number subgraphs in F_j. There are 5 edges in C_4^*, hence the probability that the C_4^* is destroyed is $5/n_t$. So the expected negative contribution by destroying the C_4^* is $-5[Y_t-2]_j/(j!n_t)$.

Taking expectation of both sides of (3.3.5) and using the tower property of conditional expectation, we have

$$\begin{aligned}\mathbf{E}\left(\frac{U_{t+1,j}}{j!}\right)-\mathbf{E}\left(\frac{U_{t,j}}{j!}\right) &= \frac{9}{n_t}\mathbf{E}\left(\frac{U_{t,j-1}}{(j-1)!}\right)+\frac{9(j+1)}{n_t^2}\mathbf{E}\left(\frac{[Y_t]_{j+1}I\{W_t=0\}}{(j+1)!}\right)\\ &\quad -\frac{3j+5}{n_t}\mathbf{E}\left(\frac{U_{t,j}}{j!}\right)+O(n_t^{-3}).\end{aligned}$$

Note the error term $O(n_t^{-3})$ includes $\mathbf{E}(f(j,G_t))$ (as estimated above), as well as $\mathbf{E}((1+Y_t+Y_{t,4})[Y_t-2]_{j-2}I\{W_t=1\}/(j-2)!n_t^2)$, $\mathbf{E}([Y_t]_{j+1}I\{W_t=0\}/(j!n_t^3))$ and $\mathbf{E}(I\{W_t=1\}/n_t^2)$. This bound holds because $Y_t[Y_t-2]_{j-2}I\{W_t=1\}/(j-2)!$ counts subgraphs with $j-1$ triangles and a copy of C_4^*, $Y_{t,4}[Y_t-2]_{j-2}I\{W_t=1\}/(j-2)!$ counts subgraphs with one 4-cycle, $j-1$ triangles and a copy of C_4^*, and $[Y_t]_{j+1}I\{W_t=0\}/j!$ counts subgraphs with $j+1$ triangles, and hence by Lemma 3.2.4 $\mathbf{E}((1+Y_t+Y_{t,4})[Y_t-2]_{j-2}I\{W_t=1\}/(j-2)!)=O(n_t^{-1})$, $\mathbf{E}([Y_t]_{j+1}I\{W_t=0\}/j!)=O(1)$, and $\mathbf{E}(I\{W_t=1\})=\mathbf{P}(W_t=1)=O(n_t^{-1})$.

Clearly for all fixed $j \geq 0$,

$$\mathbf{E}([Y_t]_jI\{W_t=0\})=\mathbf{E}([Y_t]_j+O([Y_t]_jI\{W_t\geq 1\}))=\mathbf{E}([Y_t]_j)+O(\mathbf{E}([Y_t]_jW_t)). \quad (3.3.6)$$

Hence by Lemma 3.2.10 we have $\mathbf{E}([Y_t]_jI\{W_t=0\})=3^j+O(n_t^{-1})$. Together with $\mathbf{E}(U_{t,j-1})=27/(4n_t)3^{j-1}+O(n_t^{-2})$ by the induction hypothesis, we obtain

$$\mathbf{E}(U_{t+1,j}/j!)=\left(1-\frac{3j+5}{n_t}\right)\mathbf{E}(U_{t,j}/j!)+\frac{9}{n_t}\cdot\frac{27}{4n_t}\cdot\frac{3^{j-1}}{(j-1)!}+\frac{9}{n_t^2}\cdot\frac{3^{j+1}}{j!}+O(n_t^{-3}).$$

By Lemma 3.2.2 we obtain (3.3.3) as required. ∎

Proof of Theorem 3.3.2: As mentioned above, it is enough to show that the ϵ-mixing time for $\sigma_{t,3}$, i.e. the distribution of Y_t, is not $o(\epsilon^{-1})$.

A random walk $(X_t)_{t\geq 0}$ was defined in the proof of Theorem 3.3.1 as follows, and was used to obtain the upper bound of the ϵ-mixing time (in the special case $k=3$) by the coupling method. In our proof, we use the same random walk $(X_t)_{t\geq 0}$ to study the lower bound of the ϵ-mixing time. However, we do not construct a coupling of $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ in our case. Instead, we let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ walk independently and we only compare the total variation distance between X_t and Y_t.

Define $\mathbf{B}_{t,3} := \{i \in \mathbb{Z}_+ : (9+3i)/n_t \le 1\}$, and the boundary of $\mathbf{B}_{t,3}$ to be $\partial\mathbf{B}_{t,3} := \{i \in \mathbf{B}_{t,3} : i+1 \notin \mathbf{B}_{t,3}\}$. The notation w.p. denotes "with probability."

For $X_t \in \mathbf{B}_{t,3} \setminus \partial\mathbf{B}_{t,3}$,

$$X_{t+1} = \begin{cases} X_t - 1 & \text{w.p. } 3X_t/n_t \\ X_t & \text{w.p. } 1 - 3X_t/n_t - 9/n_t \\ X_t + 1 & \text{w.p. } 9/n_t. \end{cases}$$

For $X_t \in \partial\mathbf{B}_{t,3}$,

$$X_{t+1} = \begin{cases} X_t - 1 & \text{w.p. } 3X_t/n_t \\ X_t & \text{w.p. } 1 - 3X_t/n_t. \end{cases}$$

For $X_t \notin \mathbf{B}_{t,3}$,

$$X_{t+1} = X_t \quad \text{w.p. } 1.$$

It was shown in the proof of Theorem 3.3.1 that that the Poisson distribution with mean 3, $\mathbf{Po}(3)$, is a stationary distribution of the Markov chain $(X_t)_{t\ge 0}$. For the completeness, we reproduce the proof as follows.

Assuming X_t has distribution $\mathbf{Po}(3)$, we have

$$\mathbf{P}(X_t = i) = e^{-3}\frac{3^i}{i!} \qquad \text{for all } i \in \mathbb{Z}_+$$

where $\mathbb{Z}_+$ denotes the set of nonnegative integers. Let $\mathbf{P}_{ij} = \mathbf{P}(X_{t+1} = j \mid X_t = i)$. For $j \in \mathbf{B}_t \setminus \partial\mathbf{B}_t$, we have

$$\begin{aligned} \mathbf{P}(X_{t+1} = j) &= \sum_{i\in\mathbb{Z}_+} \mathbf{P}(X_t = i)\mathbf{P}_{ij} \\ &= e^{-3}\frac{3^{j-1}}{(j-1)!}\frac{9}{n_t} + e^{-3}\frac{3^j}{j!}\left(1 - \frac{9}{n_t} - \frac{3j}{n_t}\right) + e^{-3}\frac{3^{j+1}}{(j+1)!}\frac{3\,(j+1)}{n_t} \\ &= e^{-3}\frac{3^j}{j!}. \end{aligned}$$

For $j \in \mathbb{Z}_+$, such that $j \in \partial\mathbf{B}_t$, we have

$$\begin{aligned} \mathbf{P}(X_{t+1} = j) &= \sum_{i\in\mathbb{Z}_+} \mathbf{P}(X_t = i)\mathbf{P}_{ij} \\ &= e^{-3}\frac{3^{j-1}}{(j-1)!}\frac{9}{n_t} + e^{-3}\frac{3^j}{j!}\left(1 - \frac{3j}{n_t}\right) \\ &= e^{-3}\frac{3^j}{j!}. \end{aligned}$$

For $j \in \mathbb{Z}_+$, such that $j \notin \mathbf{B}_t$, we have

$$\mathbf{P}(X_{t+1} = j) \quad = \quad \sum_{i \in \mathbb{Z}_+} \mathbf{P}(X_t = i)\mathbf{P}_{ij} = e^{-3}\frac{3^j}{j!}.$$

Thus $\mathbf{Po}(3)$ is invariant, and so by definition it is a stationary distribution.

Let the random walk X_t starts with the stationary distribution $\mathbf{Po}(3)$, so X_t has the same distribution for all $t \geq 0$. Let $(X_t)_{t\geq 0}$ walk independently of $(Y_t)_{t\geq 0}$ as generated by the graph process $(G_t)_{t\geq 0}$. We aim to estimate the total variation distance between Y_t and X_t. By the definition of the total variation distance (2.1.2),

$$d_{TV}(X_t, Y_t) = \frac{1}{2}\sum_{i=0}^{\infty} |\mathbf{P}(X_t = i) - \mathbf{P}(Y_t = i)|\,.$$

Define $\delta_t = \mathbf{P}(X_t = 0) - \mathbf{P}(Y_t = 0)$. Then

$$d_{TV}(X_t, Y_t) \geq |\delta_t|.$$

From the definition of δ_t, we have

$$\begin{aligned}\delta_{t+1} \quad = \quad & \mathbf{P}(X_t = 0)\mathbf{P}(X_{t+1} = 0 \mid X_t = 0) - \mathbf{P}(Y_t = 0)\mathbf{P}(Y_{t+1} = 0 \mid Y_t = 0) \\ & +\mathbf{P}(X_t \neq 0)\mathbf{P}(X_{t+1} = 0 \mid X_t \neq 0) - \mathbf{P}(Y_t \neq 0)\mathbf{P}(Y_{t+1} = 0 \mid Y_t \neq 0). \quad (3.3.7)\end{aligned}$$

Without loss of generality, we may assume that $n_0 \geq 9$. Then from the transition probability of X_t we have

$$\mathbf{P}(X_{t+1} \neq 0 \mid X_t = 0) = \frac{9}{n_t} \quad \text{for all } t \geq 0. \qquad (3.3.8)$$

Now we estimate $\mathbf{P}(Y_{t+1} \neq 0 \mid Y_t = 0)$. We consider the creation of a new triangle. Given an edge e of G_t, a new triangle is created containing e if and only if the two pegged edges e_1 and e_2 are both adjacent to e. Of course, in a view of the definition of pegging, they must be incident with different end-vertices of e. Since G_t is 4-regular, the number of ways to choose such e_1 and e_2 is precisely 9 conditional on $Y_t = 0$. It follows that the expected number of new triangles created is $9 \cdot 2n_t/N_t$. By (3.2.3),

$$\mathbf{E}(Y_{t+1} \mid Y_t = 0) = \frac{9 \cdot 2n_t}{n_t(2n_t - 7)} = \frac{9}{n_t} + \frac{63}{2n_t^2} + O(n_t^{-3}).$$

Conditional on $Y_t = 0$, there is no chord in any 4-cycle. Then it is impossible to create more than two triangles in a single step. Hence $\mathbf{P}(Y_{t+1} \geq 3 \mid Y_t = 0) = 0$. Hence we obtain

$$\mathbf{P}(Y_{t+1} = 1 \mid Y_t = 0) + 2\mathbf{P}(Y_{t+1} = 2 \mid Y_t = 0) = \frac{9}{n_t} + \frac{63}{2n_t^2} + O(n_t^{-3}). \qquad (3.3.9)$$

To create two triangles in a single step, it is required to peg two non-adjacent edges both contained in a 4-cycle. For any 4-cycle, there are precisely two ways to choose two nonadjacent edges, so

$$\mathbf{P}(Y_{t+1} = 2 \mid Y_t = 0, Y_{t,4} = j) = \frac{2j}{N_t} = \frac{j(1+o(1))}{n_t^2},$$

and thus

$$\mathbf{P}(Y_{t+1} = 2 \mid Y_t = 0) \quad = \quad \sum_{j=0}^{\infty} \frac{j(1+o(1))}{n_t^2} \mathbf{P}(Y_{t,4} = j \mid Y_t = 0). \tag{3.3.10}$$

By Theorem 3.2.1 and 3.3.1, Y_t and $Y_{t,4}$ are asymptotically independent Poisson, with means 3 and 9 respectively, and the total variation distance between the joint distribution of $(Y_t, Y_{t,4})$ and its limit is at most $O(n_t^{-1})$. So $\mathbf{P}(Y_{t,4} = j \mid Y_t = 0) = e^{-9}9^j/j! + O(n_t^{-1})$. Hence

$$\sum_{j \le \log n_t} \frac{j(1+o(1))}{n_t^2} \mathbf{P}(Y_{t,4} = j \mid Y_t = 0) \quad = \quad \frac{9}{n_t^2} + o(n_t^{-2}). \tag{3.3.11}$$

It was shown in the proof of Theorem 3.2.1 that $\mathbf{E}Y_{t,4}^3 = O(1)$. By Theorem 3.3.1, the total variation distance between the distribution of Y_t and its limit $\mathbf{Po}(3)$ is $O(n_t^{-1})$. So $\mathbf{P}(Y_t = 0) = e^{-3} + O(n_t^{-1})$. Then by the Markov inequality,

$$\mathbf{P}(Y_{t,4} \ge j \mid Y_t = 0) = \mathbf{P}(Y_{t,4}^3 \ge j^3 \mid Y_t = 0) \le \frac{1}{j^3}\mathbf{E}(Y_{t,4}^3 \mid Y_t = 0) = O(1/j^3).$$

Thus

$$\sum_{j > \log n_t} \frac{j(1+o(1))}{n_t^2} \mathbf{P}(Y_{t,4} = j \mid Y_t = 0) = o(n_t^{-2}). \tag{3.3.12}$$

By (3.3.9)–(3.3.12),

$$\mathbf{P}(Y_{t+1} = 2 \mid Y_t = 0) \quad = \quad \frac{9}{n_t^2} + o(n_t^{-2}), \tag{3.3.13}$$

$$\mathbf{P}(Y_{t+1} = 1 \mid Y_t = 0) \quad = \quad \frac{9}{n_t} + \frac{27}{2n_t^2} + o(n_t^{-2}), \tag{3.3.14}$$

$$\mathbf{P}(Y_{t+1} \neq 0 \mid Y_t = 0) \quad = \quad \frac{9}{n_t} + \frac{45}{2n_t^2} + o(n_t^{-2}). \tag{3.3.15}$$

From (3.3.7), (3.3.8) and (3.3.15),

$$\begin{aligned}
\delta_{t+1} &= \mathbf{P}(X_t = 0)\left(1 - \frac{9}{n_t}\right) - (\mathbf{P}(X_t = 0) - \delta_t)\left(1 - \frac{9}{n_t} - \frac{45}{2n_t^2} + o(n_t^{-2})\right) \\
&\quad + \mathbf{P}(X_t \neq 0)\mathbf{P}(X_{t+1} = 0 \mid X_t \neq 0) - (\mathbf{P}(X_t \neq 0) + \delta_t)\mathbf{P}(Y_{t+1} = 0 \mid Y_t \neq 0) \\
&= \delta_t\left(1 - \frac{9}{n_t} + O(n_t^{-2}) - \mathbf{P}(Y_{t+1} = 0 \mid Y_t \neq 0)\right) \\
&\quad + \mathbf{P}(X_t \neq 0)\left(\mathbf{P}(X_{t+1} = 0 \mid X_t \neq 0) - \mathbf{P}(Y_{t+1} = 0 \mid Y_t \neq 0)\right) \\
&\quad + \mathbf{P}(X_t = 0)\left(\frac{45}{2n_t^2} + o(n_t^{-2})\right). \qquad (3.3.16)
\end{aligned}$$

It only remains to estimate $\mathbf{P}(X_{t+1} = 0 \mid X_t \neq 0)$ and $\mathbf{P}(Y_{t+1} = 0 \mid Y_t \neq 0)$. From the definition of the random walk of $(X_t)_{t \geq 0}$,

$$\begin{aligned}
\mathbf{P}(X_{t+1} = 0 \mid X_t \neq 0) &= \frac{\mathbf{P}(X_t = 1)\mathbf{P}(X_{t+1} = 0 \mid X_t = 1)}{\mathbf{P}(X_t \neq 0)} \\
&= \frac{3}{n_t}\frac{\mathbf{P}(X_t = 1)}{\mathbf{P}(X_t \neq 0)}. \qquad (3.3.17)
\end{aligned}$$

The calculation of $\mathbf{P}(Y_{t+1} = 0 \mid Y_t \neq 0)$ is not so straightforward. Given any two distinct edges e_i and e_j, we can define a walk $e_i, e_{l_1}, e_{l_2}, \ldots, e_{l_k}, e_j$, such that every two consecutive edges appearing in the walk are adjacent. The *distance* of e_i and e_j is defined to be the length of the shortest walk between e_i and e_j. For instance, if e_i and e_j are adjacent, then their distance is 1. Conditional on $Y_t = 1$, i.e. the number of triangles in G_t being 1, if this triangle is destroyed without creating any new triangles, then one of the edges contained in the triangle must be pegged. Call it e_1. The other edge e_2 being pegged must be chosen from those whose distance from e_1 is at least 3. Let $\mathcal{R}$ be the rare event that at least one 4-cycle shares a common edge with this triangle, and $\overline{\mathcal{R}}$ be the complement of $\mathcal{R}$. There are 3 ways to choose e_1 and 21 edges within distance 2 from e_1, including e_1 itself, if $\overline{\mathcal{R}}$ occurs. Otherwise, there are in any case $O(1)$ edges within distance 2 from e_1. Hence

$$\mathbf{P}(Y_{t+1} = 0 \mid Y_t = 1) = \frac{3(2n_t - 21)}{N_t}\mathbf{P}(\overline{\mathcal{R}} \mid Y_t = 1) + \frac{3(2n_t - O(1))}{N_t}\mathbf{P}(\mathcal{R} \mid Y_t = 1).$$

Note that the occurrence of $\mathcal{R}$ implies that $W_{t,5,1} \geq 1$. So by Lemma 3.2.4,

$$\mathbf{P}(\mathcal{R} \mid Y_t = 1) \leq \frac{\mathbf{P}(\mathcal{R})}{\mathbf{P}(Y_t = 1)} = O(n_t^{-1}).$$

Noting that (3.2.3) implies $1/N_t = 1/(2n_t^2)(1 + 7/2n_t + O(n_t^{-2}))$,

$$\mathbf{P}(Y_{t+1} = 0 \mid Y_t = 1) = \frac{3}{n_t} - \frac{21}{n_t^2} + O(n_t^{-3}). \qquad (3.3.18)$$

Given $Y_t = j$ for any $j \geq 3$, to destroy all j triangles in a single step, it is required either to peg an edge contained in j triangles, and hence a small subgraph with excess at least 2, or to peg two edges such that one edge is contained in at least one triangle, and the other edge contained in at least two triangles. The latter is a small subgraph with excess at least 1. Both cases imply that for $j \geq 3$,

$$\mathbf{P}(Y_{t+1} = 0 \mid Y_t = j) = O(n_t^{-3}). \tag{3.3.19}$$

Now we only need to compute $\mathbf{P}(Y_{t+1} = 0 \mid Y_t = 2)$. To destroy two triangles in a single step, either the two triangles are isolated and the algorithm pegs two edges which are contained in two triangles, or the two triangles share a common edge and the algorithm pegs the common edge, i.e. the chord of a C_4^*. Conditional on $Y_t = 2$, the number of C_4^* can be either 0 or 1. Let W_t denote the number of C_4^* as before. If $W_t = 0$, the two triangles are isolated, and then two edges contained in different triangles are pegged, so $\mathbf{P}(Y_{t+1} = 0 \mid Y_t = 2, W_t = 0) = 9/N_t$. If $W_t = 1$, then the algorithm pegs the chord of the C_4^*. So $\mathbf{P}(Y_{t+1} = 0 \mid Y_t = 2, W_t = 1) = (2n_t - 7)/N_t$. Thus

$$\begin{aligned}
&\mathbf{P}(Y_{t+1} = 0 \mid Y_t = 2) \\
&= \frac{9}{N_t}\left(1 - \mathbf{P}(W_t = 1 \mid Y_t = 2)\right) + \frac{2n_t - 7}{N_t}\mathbf{P}(W_t = 1 \mid Y_t = 2) \\
&= \frac{9}{N_t} + \frac{2n_t - 16}{N_t}\mathbf{P}(W_t = 1 \mid Y_t = 2).
\end{aligned} \tag{3.3.20}$$

By Lemma 3.3.3, $\mathbf{P}(Y_t = 2 \mid W_t = 1) = e^{-3} + o(1)$ and therefore using (3.3.4),

$$\mathbf{P}(W_t = 1 \mid Y_t = 2) = \frac{\mathbf{P}(Y_t = 2 \mid W_t = 1)\mathbf{P}(W_t = 1)}{\mathbf{P}(Y_t = 2)} = \frac{3 + o(1)}{2n_t} + O(n_t^{-2}).$$

Combining this with (3.3.20) and (3.2.3), we have

$$\mathbf{P}(Y_{t+1} = 0 \mid Y_t = 2) = \frac{6 + o(1)}{n_t^2} + O(n_t^{-3}). \tag{3.3.21}$$

From (3.3.18), (3.3.19) and (3.3.21) we have

$$\begin{aligned}
\mathbf{P}(Y_{t+1} = 0 \mid Y_t \neq 0) \;\; = \;\; & \left(\frac{3}{n_t} - \frac{21}{n_t^2} + O(n_t^{-3})\right)\frac{\mathbf{P}(Y_t = 1)}{\mathbf{P}(Y_t \neq 0)} + \frac{6 + o(1)}{n_t^2}\frac{\mathbf{P}(Y_t = 2)}{\mathbf{P}(Y_t \neq 0)} \\
& + O(n_t^{-3}).
\end{aligned} \tag{3.3.22}$$

By Theorem 3.3.1, $d_{TV}(X_t, Y_t) = O(n_t^{-1})$, and so (3.3.17) gives

$$\begin{aligned}
&\mathbf{P}(X_{t+1} = 0 \mid X_t \neq 0) - \mathbf{P}(Y_{t+1} = 0 \mid Y_t \neq 0) \\
&\quad = \frac{3}{n_t}\left(\frac{\mathbf{P}(X_t = 1)}{\mathbf{P}(X_t \neq 0)} - \frac{\mathbf{P}(Y_t = 1)}{\mathbf{P}(Y_t \neq 0)}\right) + \frac{21}{n_t^2}\frac{\mathbf{P}(X_t = 1)}{\mathbf{P}(X_t \neq 0)} - \frac{6 + o(1)}{n_t^2}\frac{\mathbf{P}(X_t = 2)}{\mathbf{P}(X_t \neq 0)} + O(n_t^{-3}) \\
&\quad = \frac{3}{n_t}O(d_{TV}(X_t, Y_t)) + \frac{36e^{-3}}{(1 - e^{-3})n_t^2} + o(n_t^{-2}).
\end{aligned}$$

Combining this with (3.3.16) and (3.3.22) gives

$$\delta_{t+1} \geq \delta_t(1-\gamma(t)) + \frac{3(1-e^{-3})}{n_t} O(d_{TV}(X_t, Y_t)) + \frac{117e^{-3}}{2n_t^2} + o(n_t^{-2}), \tag{3.3.23}$$

where $\gamma(t) = 9/n_t + \mathbf{P}(Y_{t+1} = 0 \mid Y_t \neq 0) + O(n_t^{-2}) \geq 9/n_t + O(n_t^{-2})$. For a contradiction, assume that $d_{TV}(X_t, Y_t) = o(n_t^{-1})$. Then (3.3.23) gives

$$\delta_{t+1} \geq \delta_t(1-\gamma(t)) + \frac{117e^{-3}}{2n_t^2} + w(n_t),$$

for some function $w(n_t)$ such that $w(n_t) = o(n_t^{-2})$.

Let $(a_t)_{t\geq 0}$ be defined as $a_0 = \delta_0$ and for all $t \geq 0$,

$$a_{t+1} = a_t(1-\gamma(t)) + \frac{117e^{-3}}{2n_t^2} + w(n_t).$$

Clearly $\delta_0 \geq a_0$. Assume $\delta_t \geq a_t$ for some $t \geq 0$. Then

$$\delta_{t+1} \geq \delta_t(1-\gamma(t)) + \frac{117e^{-3}}{2n_t^2} + w(n_t) \geq a_t(1-\gamma(t)) + \frac{117e^{-3}}{2n_t^2} + w(n_t) = a_{t+1}.$$

Hence $\delta_t \geq a_t$ for all $t \geq 0$. By Lemma 3.2.2, $a_t = \Theta(n_t^{-1})$. Hence $\delta_t = \Omega(n_t^{-1})$, which contradicts the assumption that $d_{TV}(X_t, Y_t) = o(n_t^{-1})$. So $d_{TV}(Y_t, \mathbf{Po}(3))$ is not $o(n_t^{-1})$. Clearly

$$d_{TV}(\mathbf{Y}_{t,k}, \mathbf{Po}(\mu_3, \dots, \mu_k)) \geq d_{TV}(Y_t, \mathbf{Po}(3)),$$

where $\mathbf{Po}(\mu_3, \dots, \mu_k)$ is the joint independent Poisson distribution with means $\mu_3, \dots, \mu_k$, and μ_i is as stated in Theorem 3.2.1, for all $3 \leq i \leq k$. So $d_{TV}(\mathbf{Y}_{t,k}, \mathbf{Po}(\mu_3, \dots, \mu_k))$ is not $o(n_t^{-1})$. ∎

3.4 Connectivity

In this section, we study the connectivity of graphs generated from the pegging algorithm. The simulations in [34] indicate that the SWAN network has nice connectivity properties whereas our results indicate the connectivity properties of the SWAN network under long-term growth. It is well known [10, 59] that the random d-regular graphs are a.a.s. d-connected in the uniform model for any fixed constant $d \geq 3$. We show that the random d-regular graphs generated by the pegging algorithm, for any arbitrary even integer $d \geq 4$, are a.a.s. d-connected.

We call a vertex cut of size k a *k-cut*, and an edge cut of size k a *k-edge-cut*. The vertex cuts in G_t are closely related to the edge cuts. We say an edge cut A is *generated by* a vertex cut S in a graph G if for some component S' of $G - S$, A is the set of edges with one end vertex in S and the other end vertex in S'. Figure 3.5 is an example for the simple case of a 3-cut when $d = 4$. These are edge cuts of size 6, 5 and 7 generated by the 3-cuts in Figure 3.5. If S is a k-cut, then there are at least two components of $G - S$ and hence S generates at least two edge cuts. The number of edges with exactly one end vertex in S is at most dk. Hence there exists an edge cut generated by S with size at most $dk/2$. In particular, when $d = 4$, a 3-cut generates at least one edge cut of size at most 6. Similarly, a 2-cut generates some edge cut of size at most 4, and a 1-cut generates some edge cut of size at most 2. So the study of edge cuts of size at most 6 will be helpful for the discussion of vertex cuts of size at most 3 in the case $d = 4$.

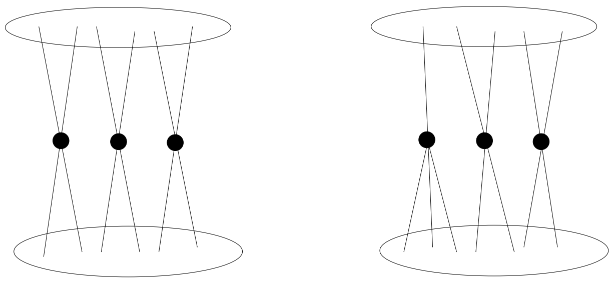

Figure 3.5: *a 3-cut which generates an edge-cut of size 6 or smaller*

We call an edge cut A *proper* if no proper subset of A is an edge cut of G. Unless otherwise specified, all edge cuts discussed in this section are proper edge cuts. The graph $G - A$ has two components if A is a proper edge cut. In this section, we call these two components the *side-components* of the edge cut A.

Definition 3.4.1 *We call an edge cut A of a graph G trivial if it is of the form $A = E(S, \overline{S})$ for some S that induces a tree in G, where $\overline{S} = V(G) - S$. The notation $E(S, \overline{S})$ denotes the set of edges with one end in S and the other end in $\overline{S}$.*

For any S and A specified as in the above definition, there are $|S| - 1$ edges between vertices in S. Since G_t is d-regular, $|A| = d|S| - 2(|S| - 1)$, and hence a trivial edge cut is always of size $dl - 2(l - 1)$ for some integer $l \geq 1$. Next we show that given an integer $l \geq 1$, the number of trivial edge cuts in G_t with size $dl - 2(l - 1)$ is a.a.s. $\Theta(n_t)$. Any trivial edge cut of this size has a side-component whose size is l. For any vertex $v \in G_t$, there are at most $\binom{d(d-1)^{l-2}}{l-1} = O(1)$ trees of size l that contain v. So the number of induced trees of size l is at most $O(n_t)$, which gives an upper bound on the number of trivial edge cuts with size $dl - 2(l-1)$. The expected number of cycles of length at most l is $O(1)$ by Theorem 3.2.1, and so there are a.a.s. $\Theta(n_t)$ induced $(l-1)$-paths. Since each induced $(l-1)$-path forms a trivial edge cut, this gives a lower bound on the number of trivial edge cuts with size $dl - 2(l-1)$. Thus, a.a.s. there are $\Theta(n_t)$ trivial edge cuts with size $dl - 2(l-1)$.

Among all edge cuts other than trivial ones, we define the *semi-trivial* edge cut, which acts as a transition from trivial edge cuts to the rest.

Definition 3.4.2 *An edge cut is called semi-trivial if it is of the form $A = E(S, \overline{S})$ for some S that induces a connected unicyclic subgraph. Edge cuts that are neither trivial nor semi-trivial are called non-trivial.*

By definition, one of the side-components of a semi-trivial edge cut is connected and contains one cycle. If that side-component is of size l, then it contains l edges, and therefore the semi-trivial edge cut is of size $dl - 2l$. By Theorem 3.2.1, the expected number of semi-trivial edge cuts of any given size is $O(1)$.

By examining the neighbours of a trivial or semi-trivial edge cut, we prove the following lemma, indicating that, to determine the vertex connectivity, it is sufficient to study the non-trivial edge cuts.

Lemma 3.4.3 *For any graph in $\mathcal{P}(G_0, d)$ with $d \geq 3$, the probability that G_t contains an edge cut that is trivial or semi-trivial and is generated by some vertex cut of size at most $d - 1$ is $O(n_t^{-1})$.*

Proof Assume that S is a vertex cut of size at most $d-1$, and S generates a trivial or semi-trivial edge cut A. Let S_1 denote the smaller side-component of A and let l denote $|S_1|$. Then $|A| \geq dl - 2l$ since S_1 contains at most one cycle. We also have $|A| < d(d-1)$ since A is generated by S with $|S| \leq d-1$. This implies $l < d^2$. If A is a trivial edge cut, then $|A| = dl - 2(l-1)$ and the subgraph of G_t induced by $S_1 \cup S$ contains at most $l + d - 1$ vertices, and at least $(l-1) + |A| = dl - l + 1$ edges. Then the excess of this subgraph is at least $dl - l + 1 - (l + d - 1) = (d-2)(l-1)$. Since $d \geq 3$, the excess can be 0 only if $l = 1$, which means that the vertex cut is the set of neighbours of some vertex, which is of size d. This contradicts $|S| \leq d - 1$. Hence the subgraph has excess at least 1. Similarly, if A is semi-trivial, then $|A| = dl - 2l$ and the subgraph induced by $S \cup S_1$ contains at most $l + d - 1$ vertices, and at least $l + |A| = dl - l$ edges. So the excess is at least $dl - l - (l + d - 1) = (d-2)(l-1) - 1$. Since S_1 contains a cycle, $l \geq 3$. So the excess of the subgraph is at least 1. We have shown that if A is generated by a vertex cut of size at most $d-1$ and A is trivial or semi-trivial, there exists a subgraph of G_t with size at most d^2 and excess at least 1. By Lemma 3.2.4, this occurs with probability $O(n_t^{-1})$. Thus, a.a.s. there exists no trivial or semi-trivial edge cuts generated by vertex cuts of size at most $d-1$. ∎

We first study the connectivity of random d-regular graphs generated by the pegging algorithm when d is even. In this way we prove Theorem 3.4.5 for even degrees.

The graph $G_t \in \mathcal{P}(G_0, d)$ contains $n_t = n_0 + t$ vertices and $m_t = dn_t/2$ edges for any even $d \geq 4$. Let N_t be the number of ways to choose $d/2$ non-adjacent edges. Then N_t is the number of ways to perform a pegging operation at step t, and N_t is asymptotically $\binom{m_t}{d/2}$, since the number of ways to choose $d/2$ edges that are not pairwise non-adjacent is $O\big(m_t^{d/2-1}\big)$. Let $Y_{t,k}$ be the number of non-trivial edge cuts of size k in G_t. Correspondingly, let $Y^*_{t,k}$, $\widetilde{Y}_{t,k}$ be the number of trivial and semi-trivial edge cuts of size k. So $Y^*_{t,k} = \Theta(n_t)$ if $k = dl - 2(l-1)$ for some integer $l \geq 1$, and $Y^*_{t,k} = 0$ for other values of k. By Theorem 3.2.1 the expected number of semi-trivial edge cuts of size $dl - 2l$ for any fixed integer l is $O(1)$, i.e. $\mathbf{E}\widetilde{Y}_{t,k} = O(1)$ for any fixed k. By Lemma 3.4.3 we know that the study of the behavior of $Y_{t,k}$ is enough, but in some sense it relates to $\widetilde{Y}_{t,k}$ as we will see later.

We study the random process $(Y_{t,k})_{t\geq 0}$ for any fixed k, or more precisely, we estimate the expected changes of the value of $Y_{t,k}$ in a single step. Let A be a k-edge-cut in G_t. We say that A is *destroyed* if either some edge in A is pegged or A is no longer an edge cut in G_{t+1}. Therefore to destroy A simply requires that either at least one edge in A is pegged, or two edges from different side-components of A are pegged. We say that an edge cut A' in G_{t+1} is a *new edge cut created from* A if A is destroyed and A' contains at least one new edge created at step t. Note that whenever A is destroyed, there are always new edge cuts being created at the same time. The number and the size of new edge cuts depend on the way that the former edge cut is destroyed. For any given k-edge-cut A, there are three ways to destroy it, according to the relative positions of the edges that are pegged. Let $e_1, \ldots, e_{d/2}$ be the $d/2$ pegged edges, and v the peg vertex. Of course the d new edges added form a trivial d-edge-cut themselves, but we do not count this case since $Y_{t,k}$ counts only the non-trivial edge cuts.

Type 1(i): A contains e_1 and another $k-1$ edges as shown in Figure 3.6. The other $d/2-1$ edges pegged other than e_1 are all in the same side-component of A. Figure 3.6 is an example of $d=4$. In this case, a new k-edge-cut and a new $(k+d-2)$-edge cut are created.

Type 1(ii): A contains only one edge that is pegged, and the other $d/2-1$ edges pegged are not contained in the same side-component of A. In this case, a new $(k+i)$-edge-cut and a new $(k+d-2-i)$-edge-cut are created for some $2 \leq i \leq d-4$.

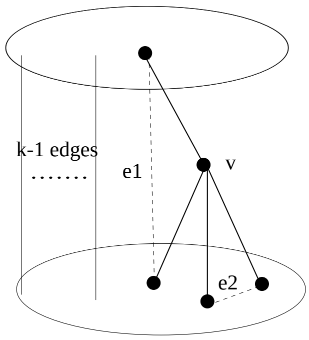

Figure 3.6: *only e_1 contained in the edge cut*

Type 2: A contains e_1, e_2 and another $k-2$ edges. Figure 3.7 is an example with $d = 4$. The probability that A is destroyed this way is $O(n_t^{-2})$.

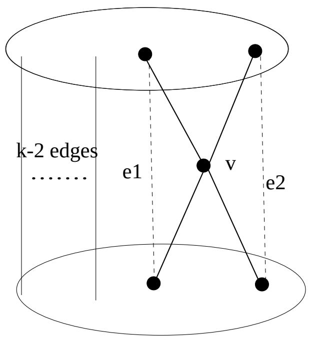

Figure 3.7: *e_1 and e_2 are both contained in the edge-cut*

Type 3(i): None of the edges in A are pegged, and one of the pegged edges lies in one side-component of A, while the rest lie on the other side-component. See Figure 3.9 as an example with $d = 4$. In this case, at most one new $(k + 2)$-edge-cut and one new $(k+d-2)$-edge-cut are created. A slight difference from Type 1 and 2 is that there might be other edge-cuts created besides the above two new edge-cuts, when the edge pegged is a bridge of some side-component. For example, let's consider $d = 4$. Let S_1 and S_2 be the two side-components. If e_1 is contained in a cycle of S_1 and e_2 is contained in a cycle of S_2, then two new $(k + 2)$-edge cuts are created. This is illustrated in Figure 3.9. Otherwise, assume e_1 is a bridge of S_1. Then we have created a new $(i + 1)$-edge cut and a new $(j+1)$-edge cut and a $(k+2)$-edge cut, where $i+j=k$. This is shown in the right hand side of Figure 3.8. Note that this implies the existence of an $(i+1)$-edge cut and a $(j+1)$-edge cut in G_t, and we will count the new $(i+1)$-edge cut and $(j+1)$-edge cut when the existing $(i+1)$-edge cut and $(j+1)$-edge cut are destroyed with Type 1. Without over counting, we only count the creation of the new $(k+2)$-edge cuts for the destruction of A. For the same reason, if both e_1 and e_2 are bridges in S_1 and S_2, then no creation of new edge-cuts is counted for the destruction of A. In conclusion, at most two $(k+2)$-edge-cuts are created for the Type 3(i) destruction of A.

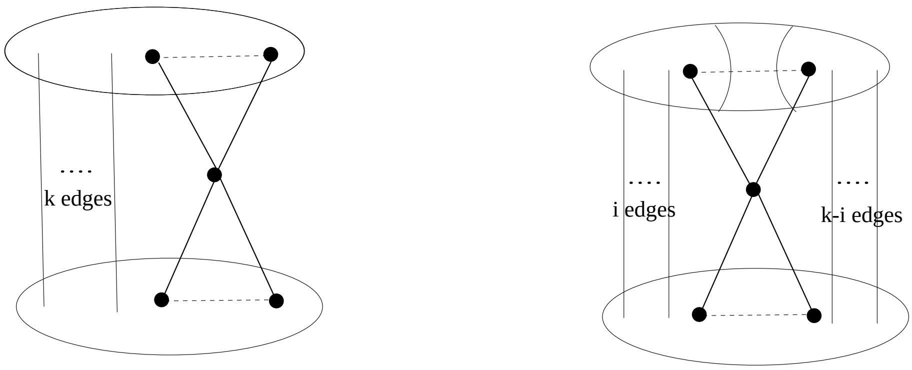

Figure 3.8: *neither e_1 nor e_2 is contained in the edge-cut*

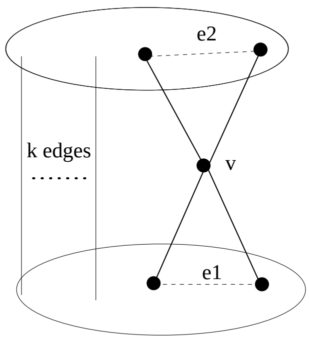

Figure 3.9: *neither e_1 nor e_2 is contained in the edge-cut*

Type 3(ii): A contains none of the edges being pegged, and both side-components of A contain at least two pegged edges. In this case, a new $(k+i)$-edge cut and a new

$(k+d-i)$-edge cut are created, for some $4 \le i \le d-4$.

Given any constant integer $M > 0$, define $\widehat{C}(M,t)$ to be the set of all non-trivial edge cuts with size at most M in graph G_t, and for any $t \ge 0$. Hence $|\widehat{C}(M,t)| = \sum_{i \le M} Y_{t,i}$. We can partition all edge cuts in $\widehat{C}(M,t)$ into three types.

- Edge cuts which are created from destruction of some edge cut in $\widehat{C}(M, t-1)$.
- Edge cuts that are in $\widehat{C}(M, t-1)$ and remain from G_{t-1} to G_t.
- Edge cuts created from some semi-trivial edge cut in G_{t-1}.

The following lemma shows that $\widehat{C}(M,t)$ is essentially empty.

Lemma 3.4.4 *Let $G_t \in \mathcal{P}(G_0, d)$, $M > 0$ be any given integer, then as $t \to \infty$,*

$$\mathbf{E}(|\widehat{C}(M,t)|) = o(1).$$

To prove this lemma we estimate the expected changes of $|\widehat{C}(M,t)|$ going from G_t to G_{t+1}. The contribution to changes comes from edge cuts of the first and the third types. The truth of this lemma is that the sizes of the edge cuts created are always at least that of the destroyed one, and the expected number of non-trivial edge cuts coming from semi-trivial edge cuts is very small. The outline of the proof is as follows.

S1. In each step, the destruction of any edge cut does not create any edge cut smaller than the one destroyed, and the number of new edge cuts created is bounded.
S2. In each step, the destruction of any edge cut creates at most one new edge cut that is of the same size of the one destroyed.
S3. There is a significant probability ($\Theta(n_t^{-1})$), that all new edge cuts created are of strictly larger size than that of the destroyed one.
S4. The probability of creating a non-trivial edge cut of size at most M from some semi-trivial edge cut is $O(n_t^{-2})$.

We are going to prove the statements S1-S4, and then show how these statements lead to the lemma.

Proof of Lemma 3.4.4. Note that $Y_{t,k}$ is the number of k-edge-cuts in $\widehat{C}(M,t)$. Define the weight of a k-edge-cut to be $1/k!$. Hence the weight of $\widehat{C}(M,t)$ is

$$W_t = \sum_{k=1}^{M} \frac{1}{k!} \widehat{Y}_{t,k}.$$

We estimate the expected value of the one step change $W_{t+1} - W_t$ by first estimating the expected change $W_{t+1} - W_t$ conditional on the value of $\widehat{C}(M,t)$. We first estimate the change caused by destruction of edge cuts in $\widehat{C}(M,t)$. For a given k-edge cut A, we analyse the different ways that it is destroyed.

Type 1(i): A new k-edge cut and a new $(k+d-2)$-edge cut are created, so the weight change is $1/k! + 1/(k+d-2)! - 1/k! \le 1/(k+2)!$. There are k ways to choose an edge in A, and at most $\binom{m_t}{d/2-1}$ ways to choose the rest $d/2-1$ edges, which lie in the same side-component of A. So the probability for this to occur is at most

$$\frac{k\binom{m_t}{d/2-1}}{N_t} \sim \frac{k}{n_t}.$$

The expected increase of weight is at most

$$\frac{1}{(k+2)!}\frac{kY_{t,k}}{n_t}.$$

Type 1(ii): A new $(k+i)$-edge cut and a new $(k+d-2-i)$-edge cut are created. The weight change is $1/(k+i)! + 1/(k+d-2-i)! - 1/k! \le 2/(k+2)! - 1/k! < 0$.

Type 2: Two new edge cuts are created, with size at least k, and thus the contribution to the weight change is at most $O(1/k!)$. The probability of this to occur is $O(n_t^{-2})$ as shown before. So the expected increase of weight is bounded by

$$O\left(\frac{1}{k!}\frac{Y_{t,k}}{n_t^2}\right).$$

Type 3(i): A new $(k+2)$-edge cut and a new $(k+d-2)$-edge cut are created, and the weight change is $1/(k+2)! + 1/(k+d-2)! - 1/k! \le 2/(k+2)! - 1/k!$. Let p denote the probability of this occurrence. Then p depends on the size of each side-component. Let S_1 denote the side-component containing less edges. If S_1 contains at least $\log m_t$ edges, then the number of ways of choosing an edge in S_1 is at least $\log m_t$ and the number of ways of choosing another non-adjacent $d/2-1$ edges in $G_t \setminus S_1$ is at least asymptotically $\binom{m_t/2}{d/2-1} = \Theta(m_t^{d/2-1})$. So $p \ge \Theta((\log m_t)m_t^{d/2-1}/N_t) = \Theta(\log m_t/n_t)$. If S_1 contains less than $\log m_t$ edges, there is at least one edge in S_1 that can be pegged and the number of ways to choose $d/2-1$ non-adjacent edges from $G_t \setminus S_1$ is asymptotically $\binom{m_t - O(\log m_t)}{d/2-1} \sim \binom{m_t}{d/2-1}$. Hence $p \ge n_t^{-1}$ in this case. Combining with the previous case, the probability for this event to occur is at least n_t^{-1}. Thus, the expected decrease of the weight is at least

$$\left(\frac{1}{k!} - \frac{2}{(k+2)!}\right)\frac{Y_{t,k}}{n_t}.$$

Type 3(ii): A new $(k+i)$-edge cut and a new $(k+d-i)$-edge cut are created, and the weight change is $1/(k+i)! + 1/(k+d-i)! - 1/k! \leq 2/(k+2)! - 1/k! < 0$.

Another contribution to the change of W_t comes from the non-trivial edge cuts newly created from semi-trivial edge cuts. Recall that the expected number of semi-trivial edge cuts of any given size in G_t is $O(1)$. Let A be a semi-trivial edge cut of size k. All the ways of creating a non-trivial edge cut from A are listed as follows.

(1) Destroy A of Type 1, with two of the edges e_1, e_2 being pegged such that they are adjacent to some common edge. Hence the common edge together with two of the new added edges will form a new triangle, which creates a non-trivial edge cut. There are only $O(1)$ ways to choose e_1 and e_2. So the probability of this to occur is $O(n_t^{-2})$.

(2) Peg at least two edges in A, i.e. destroy A of Type 2, which occurs with probability $O(n_t^{-2})$.

(3) Destroy A of Type 3, with e_1 and e_2 both adjacent to some edge in A, hence that edge together with two of the new edges form a new triangle and a new non-trivial edge cut appears. The number of choices of e_1 and e_2 is bounded and hence the probability for this to occur is $O(n_t^{-2})$.

Thus we have

$$\begin{aligned}\mathbf{E}(W_{t+1}-W_t \mid \widehat{C}(M,t)) \;\leq\; & \sum_{k=1}^{M-2}\frac{1}{(k+2)!}\frac{kY_{t,k}}{n_t} + \sum_{k=1}^{M} O\left(\frac{Y_{t,k}}{n_t^2}\right) + \sum_{k=1}^{M-2}\left(\frac{2}{(k+2)!}-\frac{1}{k!}\right)\frac{Y_{t,k}}{n_t} \\ & - \sum_{k=M-1}^{M}\frac{1}{k!}\frac{Y_{t,k}}{n_t} + \sum_{k=1}^{M} O(\widetilde{Y}_{t,k} n_t^{-2}).\end{aligned}$$

The first term comes from Type 1(i), second term comes from Type 2, and the third and fourth terms from Type 3(i). The contributions from Type 1(ii) and Type 3(ii) are ignored since they are negative. The fifth term comes from non-trivial edge cuts created from semi-trivial ones. So we have

$$\begin{aligned}&\mathbf{E}(W_{t+1}-W_t \mid \widehat{C}(M,t)) \\ &\quad\leq \sum_{k=1}^{M-2} -\left(\frac{1}{k!}-\frac{2}{(k+2)!}-\frac{k}{(k+2)!}\right)\frac{Y_{t,k}}{n_t} - \sum_{k=M-1}^{M}\frac{1}{k!}\frac{Y_{t,k}}{n_t} + \sum_{k=1}^{M} O\left(\frac{1+Y_{t,k}+\widetilde{Y}_{t,k}}{n_t^2}\right) \\ &\quad\leq \sum_{k=1}^{M-2} -\frac{1}{(k+1)!}\frac{Y_{t,k}}{n_t} - \sum_{k=M-1}^{M}\frac{1}{k!}\frac{Y_{t,k}}{n_t} + \sum_{k=1}^{M} O\left(\frac{1+Y_{t,k}+\widetilde{Y}_{t,k}}{n_t^2}\right) \\ &\quad\leq \left(-\frac{1/(M+1)}{n_t} + O(n_t^{-2})\right) W_t + \sum_{k=1}^{M} O(\widetilde{Y}_{t,k} n_t^{-2}).\end{aligned}$$

Taking the expectation of both side of the above inequality gives

$$\mathbf{E}\left(W_{t+1}\right) \leq \left(1 - \frac{1/(M+1)}{n_t} + O(n_t^{-2})\right) \mathbf{E}(W_t) + \omega(n_t),$$

where $\omega(n_t)$ is some function of n_t such that $\omega(n_t) = O(n_t^{-2})$. Define $(a_t)_{t\geq 0}$ to be $a_0 = W_0$, and

$$a_{t+1} = \left(1 - \frac{1/(M+1)}{n_t}\right) a_t + \omega(n_t), \quad \text{for all } t \geq 0.$$

Assume $\mathbf{E}(W_t) \leq a_t$ for some $t \geq 0$, then

$$\mathbf{E}\left(W_{t+1}\right) \leq \left(1 - \frac{1/(M+1)}{n_t}\right) \mathbf{E}(W_t) + \omega(n_t) \leq \left(1 - \frac{1/(M+1)}{n_t}\right) a_t + \omega(n_t) = a_{t+1}.$$

Hence $\mathbf{E}(W_t) \leq a_t$ for all $t \geq 0$. By Lemma 3.2.2 $a_t = O\left(n_t^{-1/(M+1)}\right)$, and therefore $\mathbf{E}(W_t) = O\left(n_t^{-1/(M+1)}\right)$. Hence Lemma 3.4.4 follows. ∎

Theorem 3.4.5 *Let $G_t \in \mathcal{P}(G_0, d)$ for any even $d \geq 4$, then G_t is a.a.s. d-connected.*

Proof Clearly any vertex cut of size at most $d-1$ generates an edge cut of size at most $d(d-1)/2$. By putting $M = d(d-1)/2$, the theorem follows directly from Lemma 3.4.3 and Lemma 3.4.4. More precisely, by Markov inequality, we have

$$\mathbf{P}(G_t \text{ is not } d\text{-connected}) = \mathbf{P}(W_t \geq 1) + O(n_t^{-1}) = O\left(n_t^{-1/(M+1)}\right). \quad ∎$$

The pegging algorithm for odd $d \geq 3$ is defined in Chapter 6. We can follow the same routine to prove the connectivity result when $d \geq 3$ is odd. In each step, two new vertices and $d-1$ new edges are added. The only difference from the even degree case is that, for any given k-edge cut that is destroyed, there can be up to four new edge cuts created instead of two. So there are more complicate transitions to obtaining a new k-edge cut for any k. It is straightforward but tedious to check the statements $S1$-$S4$ stated before the proof of Lemma 3.4.4. We checked the case $d = 3$ and believe that $G_t \in \mathcal{P}(G_0, d)$ are d-connected for any arbitrary integer $d \geq 3$.

Chapter 4

Orientability thresholds of random hypergraphs and the load balancing problem

4.1 Introduction

In this chapter we study the orientation of random hypergraphs. Recall that a hyperedge is said to be w*-oriented* if exactly w distinct vertices in it are marked with positive signs with respect to the hyperedge and a (w, k)-orientation of an h-hypergraph is a w-orientation of all hyperedges such that each vertex has indegree at most k. Note that a sufficiently sparse hypergraph is easily (w, k)-orientable. On the other hand, a trivial requirement for (w, k)-orientability is $m \leq kn/w$, since any w-oriented h-hypergraph with m edges has average indegree mw/n. In our work, we show the existence and determine the value of the sharp threshold at which the random h-uniform hypergraph $\mathcal{G}_{n,m,h}$ fails to be (w, k)-orientable, provided k is a sufficiently large constant. To be precise, this threshold is a number $c_{h,w,k}$ such that $\mathcal{G}_{n,m,h}$ is a.a.s. (w, k)-orientable for $m < (c_{h,w,k} - \epsilon)n$, and a.a.s. *not* (w, k)-orientable for $m > (c_{h,w,k} + \epsilon)n$, for any fixed $\epsilon > 0$. We show that the threshold is the same as the threshold at which a certain type of subhypergraph achieves a critical density.

We give precise statements of our results, including definition of the $(w, k + 1)$-core, in Section 4.2. In Section 4.3 we study the $(w, k + 1)$-core. In Section 4.4, we formulate the appropriate network flow problem, determine a canonical minimum cut for a network corresponding to a non-(w, k)-orientable hypergraph, and give conditions under which such a minimum cut can exist. Finally, in Section 4.5, we show that for k is sufficiently large, such a cut a.a.s. does not exist when the density of the core is below a certain threshold.

4.2 Main results

Let $h > w > 0$ and $k \geq 2$ be given constants. For any h-hypergraph H, we examine whether a (w, k)-orientation exists. We call a vertex *light* if the degree of the vertex is at most k. For any light vertex v, we can give v the positive sign respect to any hyperedge x that is incident to v (we call this *partially orienting* x towards to v), without violating the condition that each vertex has indegree at most k. Remove v from H, and for each hyperedge x incident to v, simply update x by removing v. Then the size of x decreases by 1, and it has one less vertex that needs to be given the positive signs. If a hyperedge becomes of size $h - w$, we can simply remove that hyperedge from the hypergraph. Repeating this until no light vertex exists, we call the remaining hypergraph $\widehat{H}$ the $(w, k+1)$*-core* of H. Every vertex in $\widehat{H}$ has degree at least $k+1$, and every hyperedge in $\widehat{H}$ of size $h-j$ requires a $(w - j)$-orientation in order to obtain a w-orientation of the original hyperedge in H.

In order to simplify the notation, we use $\bar{n}$, $\bar{m}$ and $\bar{\mu}$ to denote the number of vertices, the number of hyperedges and the average degree of $H \in \mathcal{G}_{\bar{n},\bar{m},h}$. Correspondingly we use n, m_{h-j} and μ to denote the number of vertices, the number of hyperedges of size $h - j$ and the average degree of $\widehat{H}$. Define

$$f_k(\mu) = \sum_{i \geq k} e^{-\mu} \cdot \frac{\mu^i}{i!} = 1 - \sum_{i=0}^{k-1} e^{-\mu} \cdot \frac{\mu^i}{i!}, \tag{4.2.1}$$

for any integer $k \geq 0$. By convention, define $f_k(\mu) = 1$ for any $k < 0$. The following theorem shows that the size and the number of hyperedges of $\widehat{H}$ are highly concentrated around the solution of a system of differential equations. The theorem covers the cases for any arbitrary $h > w \geq 2$ and holds for all sufficiently large k. The special case $w = 1$ has been studied by various authors and the concentration results can be found in [15, Theorem 3] which hold for all $k \geq 0$.

Theorem 4.2.1 *Let $h > w \geq 2$ be two given constant integers. Let H be a random h-hypergraph with $\bar{n}$ vertices and $\bar{m}$ hyperedges and let $\widehat{H}$ be its $(w, k+1)$-core. Let n be the number of vertices and m_{h-j} the number of hyperedges of size $h - j$ of $\widehat{H}$. Let $\bar{\mu}$ be the average degree of H. Assume $\bar{\mu} \geq ck$ for some constant $c > 1$. Then, provided k is sufficiently large, a.s.s. $n \sim \alpha\bar{n}$ and $m_{h-i} \sim \beta_{h-i}\bar{n}$ for $0 \leq i \leq w-1$, for some constants $\alpha > 0$ and $\beta_{h-i} > 0$, which are determined by the solution of the differential equation system (4.3.5)–(4.3.18) that depends only on $\bar{\mu}$, k, w and h. Furthermore, $\widehat{H}$ is uniformly random conditional on its number of vertices and hyperedges of size $h - j$ for any $j = 0, \ldots, w-1$.*

We note that our proof actually produces a lot of information on the random $(w, k+1)$-core $\widehat{H}$ such as on degree distribution (similar to that obtained for k-cores of hypergraphs in [16] for example).

Let $\mathcal{P}$ be a hypergraph property and let $\mathcal{G}_{m,n,h} \in \mathcal{P}$ denote the event that a random hypergraph from $\mathcal{G}_{m,n,h}$ has the property $\mathcal{P}$.

Definition 4.2.2 *A hypergraph property $\mathcal{P}$ has a* sharp threshold function $f(n)$ *if for any constant $\epsilon > 0$,*

$$\begin{aligned} \mathbf{P}(\mathcal{G}_{m,n,h} \in \mathcal{P}) &\to 1, \quad \textit{when} \;\; m \le (1-\epsilon)f(n) \\ \mathbf{P}(\mathcal{G}_{m,n,h} \in \mathcal{P}) &\to 0, \quad \textit{when} \;\; m \ge (1+\epsilon)f(n). \end{aligned}$$

Recall that n denotes the number of vertices and m_{h-j} denotes the number of hyperedges of size $h-j$ of $\widehat{H}$. Let $\kappa(\widehat{H})$ denote $\sum_{j=0}^{w-1}(w-j)m_{h-j}/n$. Then $\kappa(\widehat{H})$ defines a certain type of density of $\widehat{H}$. We say that a hypergraph H has *property $\mathcal{T}$* if its $(w, k+1)$-core $\widehat{H}$ satisfies the condition that $\kappa(\widehat{H})$ is at most k. (Since w and k are fixed, we often drop them from the notation.) The following theorem shows that there is a sharp threshold function of the property $\mathcal{T}$.

Theorem 4.2.3 *There exists a sharp threshold function $f(n)$ for the graph property $\mathcal{T}$ provided k is sufficiently large.*

Note that the threshold function is determined by the solution of the differential equation system (4.3.5)–(4.3.18) as referred to in Thm 2.1. Theorem 4.2.3 is approached by analysing this differential equation system.

By counting the positive signs in orientations, we see that if property $\mathcal{T}$ fails, there is no (w, k)-orientation of $\widehat{H}$ and hence there is no (w, k)-orientation of H.

We say that there exists a (w, k)-orientation of $\widehat{H}$ if there is an $(w-j)$-orientation of each hyperedge of size $h-j$ such that every vertex has indegree at most k. By counting the positive signs in orientations, we see that if property $\mathcal{T}$ fails, $\widehat{H}$ cannot have any (w, k)-orientation and hence there is no (w, k)-orientation of H.

Let $\mathbf{m} := (m_{h-w+1}, \ldots, m_h)$ be a nonnegative integer vector. Let $\mathcal{H}(n, \mathbf{m}, k+1)$ be defined as the probability space of all hypergraphs on n vertices with the following constraints and with the uniform distribution:

(a) each vertex has degree at least $k+1$;

(b) the size of each hyperedge is between $h-w+1$ and h, and the number of hyperedges of size $h-j$ for $0 \le j \le w-1$, is m_{h-j}.

By Theorem 4.2.1, $\mathcal{H}(n, \mathbf{m}, k+1)$ is a uniformly random $(w, k+1)$-core conditioned on the number of vertices and the number of hyperedges of each size.

Given a vertex set S, we say a hyperedge x is *partially contained in* S if $|x \cap S| \geq 2$.

Definition 4.2.4 *Let* $0 < \gamma < 1$. *We say that a hypergraph* G *has property* $\mathcal{A}(\gamma)$ *if for all* $S \subset V(G)$ *with* $|S| < \gamma|V(G)|$ *the number of hyperedges partially contained in* S *is less than* $k|S|/2w$.

In the following theorem, $\mathbf{m} = \mathbf{m}(n)$ denotes an integer vector for each n.

Theorem 4.2.5 *Let* γ *be any constant between* 0 *and* 1. *Then there exists a constant* $N > 0$ *depending only on* γ, *such that for all* $k > N$ *and any* $\epsilon > 0$, *if* $\mathbf{m}(n)$ *satisfies* $\sum_{j=0}^{w-1}(w-j)m_{h-j}(n) \leq kn - \epsilon n$ *for all* n, *then* $G \in \mathcal{H}(n, \mathbf{m}(n), k+1)$ *a.a.s. either has a* (w, k)*-orientation or does not have property* $\mathcal{A}(\gamma)$.

Let $f(\bar{n})$ be the threshold of property $\mathcal{T}$ given in Theorem 4.2.3. We show in Corollary 4.4.3 that a.a.s. $\widehat{H}$ has property $\mathcal{A}(\gamma)$ for a certain γ if the average degree of H is at most hk/w. Combining Theorem 4.2.3 with Theorem 4.2.5, we obtain that $f(\bar{n})$ is a sharp threshold for the existence of a (w, k)-orientation of H, when k is large enough.

Corollary 4.2.6 *Let* $h > w > 0$ *be two given integers and* k *be a sufficiently large constant. Let* $f(\bar{n})$ *be the threshold function of property* $\mathcal{T}$ *given in Theorem 4.2.3. Then* $f(\bar{n})$ *is the sharp threshold for the* (w, k)*-orientability in* $\mathcal{G}_{\bar{n},\bar{m},h}$.

For any vertex set $S \subset V(H)$, define *the subgraph induced by* S *with parameter* w to be the subgraph of G on vertex set S with the set of hyperedges $\{x' = x \cap S : x \in H, \ s.t. \ |x'| \geq h - w + 1\}$. Call this hypergraph H_S. Let $d(H_S)$ denote the degree sum of vertices in the hypergraph H_S and let $e(H_S)$ denote the number of hyperedges in H_S.

Corollary 4.2.7 *The following three graph properties of* $H \in \mathcal{G}_{\bar{n},\bar{m},h}$ *have the same sharp threshold.*

(i) H *is* (w, k)*-orientable.*

(i) H *has property* $\mathcal{T}$.

(i) There exists no $H' \subset H$ *as an induced subgraph with parameter* w *such that* $d(H') - (h-w)e(H') > ks$.

4.3 Analysing the size and density of the $(w, k+1)$-core

A model of generating random graphs via multigraphs, used by Bollobás and Frieze [13] and Chvatál [18], is described as follows. Let $\mathcal{P}_{\bar{n},\bar{m}}$ be the probability space of functions $g : [\bar{m}] \times [2] \to [\bar{n}]$ with the uniform distribution. Then a probability space of random multigraphs can be obtained by taking $\{g(i, 1), g(i, 2)\}$ as an edge for each i. Let $\mathbf{m} = (m_2, \ldots, m_h)$. This model can easily be extended to generate non-uniformly random multihypergraphs by taking $\mathcal{P}_{\bar{n},\mathbf{m}} = \{g : \cup_{i=2}^{h}[m_i] \times [i] \to [\bar{n}]\}$. Let $\mathcal{M}_{\bar{n},\mathbf{m}}$ be the probability space of random multihypergraphs obtained by taking each $\{g(j, 1), \ldots, g(j, i)\}$ as a hyperedge, where $j \in [m_i]$ and $2 \le i \le h$. (Loops and multiple edges are possible.) It was shown in [18] that $\mathcal{G}_{\bar{n},\bar{m},2}$ is equal to $\mathcal{M}_{\bar{n},\mathbf{m}}$, where $\mathbf{m} = (0, \ldots, 0, \bar{m})$, conditioned on the multihypergraph being simple, and that the probability of a multihypergraph in $\mathcal{M}_{\bar{n},\mathbf{m}}$ being simple is $\Omega(1)$ if $\bar{m} = O(\bar{n})$. This result is easily extended to the general case for any fixed $h \ge 2$ and vector $\mathbf{m}$, using the same method of proof.

Cain and Wormald [16] recently introduced a new model to analyse the k-core of a random (multi)graph or (multi)hypergraph, including its size and degree distribution. This model is called the *pairing-allocation* model. The *partition-allocation* model, as defined below, is a refinement of the pairing-allocation model, and analyses cores of multihypergraphs with given numbers of hyperedges of various sizes. We will use this model to prove Theorem 4.2.5 and to analyse a randomized algorithm called the RanCore algorithm, defined later in this section, which outputs the $(w, k + 1)$-core of an input h-hypergraph.

Given $h \ge 2$, n, $\mathbf{m} = (m_2, \ldots, m_h)$, $\mathbf{L} = (l_2, \ldots, l_h)$ and a nonnegative integer k such that $D - \ell \ge kn$, where $D = \sum_{i=2}^{h} im_i$ and $\ell = \sum_{i=2}^{h} l_i$, let V be a set of size n, and $\mathbf{M}$ a collection of pairwise disjoint sets $\{M_1, \ldots, M_h\}$, each disjoint from V, where M_i contains im_i elements partitioned into parts, each of size i, for all $2 \le i \le h$. The *partition-allocation* model defines a probability space $\mathcal{P}(V, \mathbf{M}, \mathbf{L}, k)$ as the set of mappings

$$\left\{g : \bigcup_{i=2}^{h} M_i \to \{\mathbb{L}\} \cup V\right\} \tag{4.3.1}$$

such that $|g^{-1}(\mathbb{L})| = \ell$, $|g^{-1}(j)| \ge k$ for all $j \in V$, and for all $2 \le i \le h$,

$$\left|\{j \in M_i : g(j) = \mathbb{L}\}\right| = l_i,$$

and each mapping has equal probability.

The probability space $\mathcal{P}(V, \mathbf{M}, \mathbf{L}, k)$ can equivalently be defined by way of the following algorithm. Let $\mathcal{C} = \{c_2, \ldots, c_h\}$ be a set of colours. Represent V as a set of n bins and represent each element in M_i as a ball. Colour balls in M_i with c_i. (The function of the colours is only to denote the size of the part a ball lies in.) Let $\mathbb{L}$ be a bin such that $\{\mathbb{L}\}$

is disjoint from V. Then allocate the D balls uniformly at random (u.a.r.) into the bins in $V \cup \{\mathbb{L}\}$, such that the following constraints are satisfied:

(i) $\mathbb{L}$ contains exactly ℓ balls;

(ii) each bin in V contains at least k balls;

(iii) for any $2 \le i \le h$, the number of balls with colour c_i that are contained in $\mathbb{L}$ is l_i.

Note that the mappings (4.3.1) correspond to the ways of partitioning and allocating the balls into the bins.

We call $\mathbb{L}$ a *light* bin and all bins in V heavy. As we will see later, the probability space $\mathcal{P}(V, \mathbf{M}, \mathbf{L}, k)$ corresponds to the probability space $\mathcal{M}_{\bar{n},\mathbf{m}}$ conditional on the event that the number of heavy vertices is n and the contribution to the sum of degrees of light vertices (vertices with degree at most $k-1$) from hyperedges of size i is l_i.

To make the explanation easier, the probability space $\mathcal{P}_{\bar{n},\mathbf{m}}$ can be described using the bins and balls model similarly. We may consider each element in $\cup_{i=2}^{h}[m_i]$ as a ball and each element in $[\bar{n}]$ as a bin and say a ball u is dropped into a bin v if u is mapped to v by g.

The following alternative algorithm generates the same probability space. First, allocate D balls randomly into bins $\{\mathbb{L}\} \cup V$ with the restriction that $\mathbb{L}$ contains exactly ℓ balls and each bin in V contains at least k balls. Then colour the balls u.a.r. with the following constraints:

(i) exactly im_i balls are coloured with c_i;

(ii) for any $i = 2, \ldots, h$, the number of balls with colour c_i contained in $\mathbb{L}$ is exactly l_i.

Then take u.a.r. a partition of the balls such that for any $i = 2, \ldots, h$, all balls with colour c_i are partitioned into parts of size i. We call this alternative algorithm the *allocation-partition* algorithm since it allocates before partitioning the balls. This algorithm assists with analysis in some situations.

Given $\bar{n}$, $\mathbf{L}$, V, $\mathbf{M}$, k and $g \in \mathcal{P}(V, \mathbf{M}, \mathbf{L}, k)$, define $H(g)$ to be the probability space of multihypergraphs g' such that g' is obtained from g by reallocating balls in $\mathbb{L}$ into the set of bins $V' = [\bar{n}] \setminus V$ such that no bin in V' receives more than $k-1$ balls. Moreover, all g' in $H(g)$ occur with equal probability. This immediately leads to the following proposition.

Proposition 4.3.1 *For any $\bar{n}$, n, $\mathbf{m}$, $\mathbf{L}$, k such that $D - \ell \ge kn$ and $\ell \le (k-1)(\bar{n}-n)$, where $D = \sum_{i=2}^{h} im_i$ and $\ell = \sum_{i=2}^{h} l_i$ as defined before, the distribution of $H(g)$ where*

$g \in \mathcal{P}(V, \mathbf{M}, \mathbf{L}, k)$ *is identical to* $g' \in \mathcal{P}_{\bar{n},\mathbf{m}}$ *conditional on the event that the set of heavy vertices is* V*, and the contribution to the sum of degrees of light vertices from hyperedges of size* i *is* l_i *for every* $2 \le i \le h$*.*

Given the values of n, $\mathbf{m} = (m_{h-w+1,\dots,m_h})$, let $M = (M_{h-w+1}, \dots, M_h)$, where $M_i = [m_i] \times [i]$. We consider $\mathcal{M}(n, \mathbf{m}, \mathbf{0}, k)$, the probability space of random multihypergraphs with minimum degree at least $k+1$ generated by $\mathcal{P}([n], \mathbf{M}, \mathbf{0}, k)$. For convenience, we denote it by $\mathcal{M}(n, \mathbf{m}, k)$. We will show in the proof of Theorem 4.2.1 that conditional on the event that the number of vertices is n and the numbers of hyperedges of size i are m_i for $h-w+1 \le i \le h$ in $\widehat{H}$, the $(w, k+1)$-core of $H \in \mathcal{M}_{\bar{n},(0,\dots,\bar{m})}$, the distribution of $\widehat{H}$ is identical to $\mathcal{M}(n, \mathbf{m}, k+1)$. We will use $\mathcal{P}([n], \mathbf{M}, \mathbf{0}, k+1)$ in Section 4.5 to prove that the result of Theorem 4.2.5 holds also for the probability space $\mathcal{M}(n, \mathbf{m}, k+1)$. The following lemma shows the relation between $\mathcal{M}(n, \mathbf{m}, k+1)$ and $\mathcal{H}(n, \mathbf{m}, k+1)$.

Lemma 4.3.2 *The probability space* $\mathcal{H}(n, \mathbf{m}, k+1)$ *is equal to* $\mathcal{M}(n, \mathbf{m}, k+1)$ *restricted to simple hypergraphs.*

Proof Let G be a hypergraph from $\mathcal{H}(n, \mathbf{m}, k+1)$. Let $\mathbf{d} = (d_1, \dots, d_n)$ be its degree sequence. Consider the allocation-partition algorithm which generates $\mathcal{P}([n], \mathbf{M}, \mathbf{0}, k+1)$. The number of ways to allocate balls into bins such that the degree sequence is $\mathbf{d}$ is

$$\binom{D}{d_1, d_2, \dots, d_n} = \frac{D!}{d_1! \cdots d_n!}.$$

For every such allocation, the number of partitions resulting in G after contracting bins into vertices is $\prod_{i=1}^{n} d_i!$. Hence the number of $g \in \mathcal{P}([n], \mathbf{M}, \mathbf{0}, k+1)$ that correspond to any simple hypergraph G is

$$\frac{D!}{d_1! \cdots d_n!} \cdot \prod_{i=1}^{n} d_i! = D!.$$

Since this does not depend on G and also due to the relation between $\mathcal{P}([n], \mathbf{M}, \mathbf{0}, k+1)$ and $\mathcal{M}(n, \mathbf{m}, k+1)$, the simple hypergraphs in $\mathcal{M}(n, \mathbf{m}, k+1)$ are uniformly distributed. Then Lemma 4.3.2 follows. ∎

A deletion algorithm producing the k-core of a random multigraph was analysed in [16]. The differential equation method [60] was used to analyse the size and the number of hyperedges of the final k-core. The degree distribution of the k-core was shown to be a truncated multinomial. We extend the deletion algorithm to analyse the $(w, k+1)$-core of $H \in \mathcal{G}_{\bar{n},\bar{m},h}$ using the partition-allocation model, but the analysis is more complicated. It is clear that the degree distribution of the $(w, k+1)$-core, conditional on the number of hyperedges of each size, is truncated multinomial by considering the allocation-partition algorithm which generates $\mathcal{P}([n], \mathbf{M}, \mathbf{0}, k+1)$.

We apply the following randomized algorithm to find the $(w, k+1)$-core of an input hypergraph H. Recalling the setting of representing multihypergraphs using bins and balls, for any hyperedge x, let $h(x)$ denote the set of $|x|$ balls in x. Initially let LV be the set of all light vertices, and let $\overline{LV} = V(H) \setminus LV$ be the set of heavy vertices. Let the balls contained in LV be called light balls.

RanCore Algorithm to obtain the $(w, k+1)$-core
Input: an h-hypergraph H. Set $t := 0$.
While neither LV nor $\overline{LV}$ is empty,
$t := t + 1$;
Remove all empty bins;
U.a.r. choose a light ball u. Let x be the hyperedge that contains u and let v be the vertex that contains u;
If $|x| \geq h - w + 2$, update x with $x \setminus \{u\}$,
otherwise, remove this hyperedge x from the current hypergraph. If any vertex $v' \in \overline{LV}$ becomes light, move v' to LV together with all balls in it;
If LV is empty, ouput the remaining hypergraph, otherwise, output the empty graph.

Let the function $f_k(x)$ be defined as in (4.2.1). Let $Z_{(\geq k)}$ be a truncated Poisson random variable with the parameter λ defined as follows.

$$\mathbf{P}(Z_{(\geq k)} = j) = \frac{e^{-\lambda}}{f_k(\lambda)} \cdot \frac{\lambda^j}{j!}, \quad \text{for any } j \geq k. \tag{4.3.2}$$

Note that it follows that $\mathbf{P}(Z_{(\geq k)} = j) = 0$ whenever $j < k$. The following proposition will be used in the proof of Theorem 4.2.1 and in Section 4.5.

Proposition 4.3.3 *Given any integer $k \geq -1$, let λ be a positive real number satisfying $\lambda f_k(\lambda) = \mu f_{k+1}(\lambda)$, where $\mu \geq ck$ for some $c > 1$. Then $\mu \geq \lambda$, and $\mu - \lambda \to 0$, $f_k(\lambda) \to 1$ as $k \to \infty$.*

Proof Since $f_k(\lambda) \geq f_{k+1}(\lambda)$ for all $k \geq -1$ by the definition of the function $f_k(x)$ in (4.2.1), it follows directly that $\mu \geq \lambda$. There is a unique λ satisfying $\mu = \lambda f_k(\lambda)/f_{k+1}(\lambda)$ since $\lambda f_k(\lambda)/f_{k+1}(\lambda)$ is an increasing function of λ (see a short proof of the monotonicity in [47, Lemma 1]). Let $p_j(\lambda) = e^{-\lambda}\lambda^j/j!$. Then

$$\mu - \lambda = \lambda \left(\frac{f_k(\lambda)}{f_{k+1}(\lambda)} - 1 \right) = \frac{\lambda p_k(\lambda)}{f_{k+1}(\lambda)}. \tag{4.3.3}$$

We first show that $\lambda > \mu - 1$ provided k is sufficiently large by showing that $\lambda + \frac{\lambda p_k(\lambda)}{f_{k+1}(\lambda)} < \mu$ by taking $\lambda = \mu - 1$. Since

$$\begin{aligned}(\mu - 1) \cdot p_k(\mu - 1) &= (\mu - 1)e^{-\mu+1}\frac{(\mu - 1)^k}{k!} \leq \mu e^{-\mu+1}\left(\frac{e\mu}{k}\right)^k \\ &\leq \exp\Big(1 - \mu + k(1 + \ln(\mu/k)) + \ln\mu\Big) \qquad (4.3.4)\end{aligned}$$

and (4.3.4) is a decreasing function of μ, we have

$$(\mu - 1) \cdot p_k(\mu - 1) \leq \exp\Big(1 - ck + k(1 + \ln c) + \ln(ck)\Big).$$

Since $c > 1$, we have $(\mu - 1) \cdot p_k(\mu - 1) \to 0$, and $f_{k+1}(\mu - 1) \to 1$ as $k \to \infty$. Hence

$$(\mu - 1) + \frac{(\mu - 1) \cdot p_k(\mu - 1)}{f_{k+1}(\mu - 1)} < (\mu - 1) + 1/2 < \mu$$

provided k is sufficiently large. Thus, $\lambda > \mu - 1$. Then $f_{k+1}(\lambda) \to 1$ as $k \to \infty$, and

$$\mu - \lambda = 2\lambda p_k(\lambda) \leq 2\lambda e^{-\lambda}\left(\frac{e\lambda}{k}\right)^k \leq \mu e^{-\mu+1}\left(\frac{e\mu}{k}\right)^k,$$

which goes to 0 as $k \to \infty$ as proved before. ∎

Lemma 4.3.4 *Let $c > 0$, δ be constants. Let $(Y_t)_{t\geq 1}$ be independent random variables such that $|Y_t| \leq c$ always and $\mathbf{E}Y_t \leq \delta$ for all $t \geq 1$. Let $X_0 = 0$ and $X_t = \sum_{i\leq t} Y_i$ for all $t \geq 1$. Then for any $\epsilon > 0$, a.a.s. $X_n \leq \delta n + \epsilon|\delta|n$. Moreover, $\mathbf{P}(X_n \geq \delta n + \epsilon|\delta|n) \leq \exp(-\Theta(\epsilon^2 n))$.*

Proof Clearly $(X_t - \delta t)_{t\geq 0}$ is a supermartingle. Let $n \to \infty$. Then the Azuma-Hoeffding inequality, originally given in [33, Theorem 3] in a form of a martingale inequality and also introduced in many textbooks (see [3]), gives

$$\mathbf{P}(X_n - \delta n \geq \epsilon|\delta|n) \leq \exp\left(-\frac{\epsilon^2\delta^2 n^2}{cn}\right) = \exp(-\Theta(\epsilon^2 n)) = o(1).$$

Thus, a.a.s. $X_n \leq \delta n + \epsilon|\delta|n$. ∎

Lemma 4.3.5 *Let $c > 0$, δ be constants. Let $(Y_t)_{t\geq 1}$ be independent random variables such that $|Y_t| \leq c$ always and $\mathbf{E}Y_t \geq \delta$ for all $t \geq 1$. Let $X_0 = 0$ and $X_t = \sum_{i\leq t} Y_i$ for all $t \geq 1$. Then for any $\epsilon > 0$, a.a.s. $X_n \geq \delta n - \epsilon|\delta|n$. Moreover, $\mathbf{P}(X_n \leq \delta n - \epsilon|\delta|n) \leq \exp(-\Theta(\epsilon^2 n))$.*

Proof This follows by applying Lemma 4.3.4, taking Y_t as $-Y_t$ and X_t as $-X_t$. ∎

The following proof of Theorem 4.2.1 is rather long and complicated. To assist the reading, we first briefly explain the structure of the proof. We first prove the uniformity of the $(w, k+1)$-core $\widehat{H}$. That is to say, we show that conditional on the number of vertices and hyperedges of each size in $\widehat{H}$, $\widehat{H}$ is distributed as $\mathcal{M}(n, \mathbf{m}, k+1)$ restricted to the simple hypergraphs. This part of the proof proceeds by first adapting the RanCore algorithm to be run on the partition-allocation model, and then proving uniformity of the partition-allocations in each step by induction.

After the proof of this part, we use the DE method to analyse asymptotic values of random variables under consideration in the random process generated by the RanCore algorithm. The difficulty in this part of the proof is caused by the fact that it is not possible to define an open domain for which we can apply the DE method from the beginning of the process. Thus, we define a domain that does not contain the origin and we need to show that the random variables under consideration "enter" the domain not too long after the start of the algorithm. Lastly we can enlarge the domain gradually so that it gets arbitrarily close to the origin and the end of the process.

Once it is shown that the asymptotic values inside the domain are approximated by a system of differential equations, we analyse the random variables when they leave the domain. We show that provided k is sufficiently large, the algorithm terminates quickly and then we estimate the size and density of $\widehat{H}$.

Proof of Theorem 4.2.1. Given the values of $HV_0 \geq 0$ and $L_0 \geq 0$ satisfying $h\bar{m} \geq L_0 + (k+1)HV_0$, $L_0 \leq k(\bar{n} - HV_0)$, and V_0, $\mathbf{M}_0$ such that $|V_0| = HV_0$ and $|M_i| = 0$ for $i \leq h-1$ and $|M_h| = h\bar{m}$, the random multigraph $H \in \mathcal{M}_{\bar{n},(0,\ldots,0,\bar{m})}$ conditional on the event that the heavy vertex set is V_0 and the degree sum of light vertices is L_0, is distributed identically to $H(g)$ with $g \in \mathcal{P}(V_0, \mathbf{M}_0, \mathbf{L}_0, k+1)$, where $\mathbf{L}_0 = (0, \ldots, 0, L_0)$, by Proposition 4.3.1. The RanCore algorithm can be adapted in an obvious way, which we explain next, to be run on $\mathcal{P}(V_0, \mathbf{M}_0, \mathbf{L}_0, k+1)$. Let $g_0 \in \mathcal{P}(V_0, \mathbf{M}_0, \mathbf{L}_0, k+1)$ and let g_t denote the partition-allocation obtained after t steps of the RanCore algorithm. In each step t, the algorithm removes a ball, denoted by u, u.a.r. chosen from all balls in $\mathbb{L}$. If the colour of u is c_{h-i} for $i < w-1$, the algorithm recolours the balls in the same part as u with the new colour c_{h-i-1}. If the colour of u is c_{h-w+1}, the algorithm removes all balls contained in the same part as u, and if any heavy bin becomes light (i.e. the number of balls contained in it becomes at most k) because of the removal of balls, the bin is removed and the balls remaining in it are put into $\mathbb{L}$.

Let V_t denote the set of heavy bins after step t and let $\mathbf{M}_t$ denote the class of sets $\{M_{t,h-w+1}, \ldots, M_{t,h}\}$ such that $M_{t,h-j}$ denotes the set of partitioned balls with colour c_{h-j} after step t for $0 \leq j \leq w-1$. Let $HV_t = |V_t|$, $m_{t,h-j} = |M_{t,h-j}|/(h-j)$ and

$\mathbf{m} = (m_{t,h-w+1}, \ldots, m_{t,h})$. Let $L_{t,h-j}$ denote the number of balls with colour c_{h-j} in $\mathbb{L}$ and let $\mathbf{L} = (L_{t,h-w+1}, \ldots, L_{t,h})$. Let $L_t = \sum_{j=0}^{w-1} L_{t,h-j}$.

We first show, by induction on t, that the partition-allocation g_t, conditional on V_t, $\mathbf{M}_t$ and $\mathbf{L}_t$, is distributed as $\mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1)$. We have already shown that $g_0 \in \mathcal{P}(V_0, \mathbf{M}_0, \mathbf{L}_0, k+1)$, conditional on V_0 and L_0, by Proposition 4.3.1. For the inductive step, it suffices to show that for any $t \geq 0$, conditional on $V_t, \mathbf{M}_t, \mathbf{L}_t$ and $V_{t+1}, \mathbf{M}_{t+1}, \mathbf{L}_{t+1}$, g_{t+1} is distributed as $\mathcal{P}(V_{t+1}, \mathbf{M}_{t+1}, \mathbf{L}_{t+1}, k+1)$. The operation of the algorithm in step $t+1$ is determined by the values of $HV_t, \mathbf{m}_t, \mathbf{L}_t$ and $HV_{t+1}, \mathbf{m}_{t+1}, \mathbf{L}_{t+1}$. For instance, if $\sum_{j=0}^{w-1} m_{t,h-j} = \sum_{j=0}^{w-1} m_{t+1,h-j}$, it implies that the algorithm removes a ball u with colour c_{h-j} from $\mathbb{L}$ for some $j \neq w-1$, and then recolours $h-j-1$ balls that are in the same part as u with the new colour c_{h-j-1}. This corresponds to updating a hyperedge x incident with a light vertex v by $x \setminus \{v\}$. The value j is determined by the difference between $\mathbf{L}_t$ and $\mathbf{L}_{t+1}$. The difference of $M_{t,h-j}$ and $M_{t,h-j}$ determines the balls that are in the same part as u. Hence there is a unique g_t that can lead to g_{t+1} by applying the RanCore algorithm. Since $g_t \in \mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1)$ by the inductive hypothesis, it then follows that g_{t+1} is distributed as $\mathcal{P}(V_{t+1}, \mathbf{M}_{t+1}, \mathbf{L}_{t+1}, k+1)$ in this case.

Similarly we analyse the other two cases. If $\sum_{j=0}^{w-1} m_{t+1,h-j} < \sum_{j=0}^{w-1} m_{t,h-j}$ and $HV_{t+1} = HV_t$, then this implies that the algorithm removes a ball u with colour c_{h-w+1} from $\mathbb{L}$ together with all balls in the same part as u, and no heavy bins become light because of the removal of balls. For any g_{t+1}, we count g_t that can lead to g_{t+1} by one step of the algorithm. The difference of $M_{t,h-w+1}$ and $M_{t+1,h-w+1}$ determines the set of $h-w+1$ balls being removed. The difference of $L_{t,h-w+1}$ and $L_{t+1,h-w+1}$ determines how many balls from $\mathbb{L}$ are removed at step $t+1$. Hence the number of g_t that can lead to g_{t+1} equals the number of ways to choose $L_{t,h-w+1} - L_{t+1,h-w+1}$ from $h-w+1$ balls and drop the remaining balls into HV_t bins. This number depends only on $L_{t,h-w+1}$, $L_{t+1,h-w+1}$, $h-w+1$ and HV_t and so is independent of the choice of g_{t+1}. So g_{t+1} is distributed as $\mathcal{P}(V_{t+1}, \mathbf{M}_{t+1}, \mathbf{L}_{t+1}, k+1)$ in this case. The last and most complicated case involves $\sum_{j=0}^{w-1} m_{t+1,h-j} < \sum_{j=0}^{w-1} m_{t,h-j}$ and $HV_{t+1} < HV_t$, which implies that the algorithm removes a ball u with colour c_{h-w+1} from $\mathbb{L}$ together with all balls in the same part as u and the removal of balls causes some heavy bins in g_t to become light. First, the difference of $M_{t,h-w+1}$ and $M_{t+1,h-w+1}$ determines the set of $h-w+1$ balls that are removed. Let i denote the number of balls removed from $\mathbb{L}$. Then $1 \leq i \leq h-w$. Let V' denote the set of bins that become light after the removal of balls. Then the difference of V_t and V_{t+1} determines V'. The difference between $\mathbf{L}_t$ and $\mathbf{L}_{t+1}$ then determines the number of balls remain in V' before the bins in V' are removed by the algorithm. Then for any given i, the number of g_t that can lead to g_{t+1} equals the number of ways to choose the right set of balls determined by the difference between $\mathbf{L}_t$ and $\mathbf{L}_{t+1}$, and reallocate them to bins in V' such that none of bins in V' receives more then k balls, and reallocate $h-w+1$ balls of colour c_{h-w+1} such that i of them are dropped into $\mathbb{L}$ and the rest are dropped into $V' \cup V_{t+1}$ such that all bins in V' contain at least $k+1$ balls. This number, denoted by $\phi(i)$, depends only on i, $h-w+1$, $\mathbf{L}_t$, $\mathbf{L}_{t+1}$, $M_{t,h-w+1}$, $M_{t+1,h-w+1}$, V_t and V_{t+1}. Then the total number of g_t that can lead to a given g_{t+1} by one step of the algorithm is $\sum_{i=1}^{h-w}\phi(i)$, which is independent of the choice of g_{t+1}. Thus, we have shown that in all cases, conditional on V_t, $\mathbf{M}_t$, $\mathbf{L}_t$ and V_{t+1}, $\mathbf{M}_{t+1}$, $\mathbf{L}_{t+1}$, there are a fixed number of partition-allocations g_t that lead to any given g_{t+1}. The inductive hypothesis gives that $g_t \in \mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1)$. Thus, g_{t+1} is distributed as $\mathcal{P}(V_{t+1}, \mathbf{M}_{t+1}, \mathbf{L}_{t+1}, k+1)$. Therefore, by induction, $g_t \in \mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1)$, conditional on V_t, $\mathbf{M}_t$, $\mathbf{L}_t$, for all $t \geq 0$. In particular, let τ be the stopping time of the RanCore algorithm. If the $(w, k+1)$-core is not empty, then conditional on V_τ, $\mathbf{M}_\tau$, $g_\tau \in \mathcal{P}(V_\tau, \mathbf{M}_\tau, \mathbf{0}, k+1)$.

Now we analyse the size of the $(w, k+1)$-core. After step t of the algorithm, let B_t denote the total number of balls remaining and let $B_{t,h-j}$ denote the number of balls coloured c_{h-j}. Let $H_{t,h-j}$ denote the number of balls contained in heavy bins that are coloured c_{h-j}, and let $A_{t,i}$ be the number of bins containing exactly i balls. We have just shown that g_t is distributed as $\mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1)$ conditional on V_t, $\mathbf{M}_t$ and $\mathbf{L}_t$. Recall that n and m_{h-j} denote the number of vertices and the number of hyperedges of size $h-j$ in $\widehat{H}$, the $(w, k+1)$-core of H and $\bar{\mu}$ denotes the average degree of H. We show, using the DE method introduced in Section 2.2.3, that if $\bar{\mu} \geq ck$ for some $c > 1$ and k is sufficiently large, a.a.s. $n \sim \alpha\bar{n}$ and $m_{h-j} \sim \beta_{h-j}\bar{n}$ for some constants $\alpha > 0$, $\beta_{h-j} > 0$ which are determined by the solution of the following differential equation system. We will show that $\alpha = HV(x^*)$ and $\beta_{h-j} = H_{h-j}(x^*)/(h-j)$ where x^* is the smallest root of $L(x) = 0$. We use the same symbols for random variables and the real valued functions that are associated to the random variables by scaling (explained below). For instance, the real function $L_{h-j}(x)$ is associated to the random variable $L_{t,h-j}$.

For $x > 0$,

$$
\begin{aligned}
L'_{h-i}(x) &= \frac{L_{h-i}(x)}{L(x)}\left(-1 - \frac{(h-i-1)L_{h-i}(x)}{B_{h-i}(x)}\right) \\
&\quad + \frac{L_{h-w+1}(x)}{L(x)}\left(\frac{(h-w)H_{h-w+1}(x)}{B_{h-w+1}(x)} \cdot \frac{(k+1)A(x)}{B(x)-L(x)} \cdot k \cdot \frac{H_{h-i}(x)}{B(x)-L(x)}\right) \\
&\quad + \frac{L_{h-i+1}(x)}{L(x)}\frac{(h-i)L_{h-i+1}(x)}{B_{h-i+1}(x)}, \quad i = 1, \ldots, w-1, &(4.3.5) \\
H'_{h-i}(x) &= \frac{L_{h-i}(x)}{L(x)}\left(-\frac{(h-i-1)H_{h-i}(x)}{B_{h-i}(x)}\right) \\
&\quad - \frac{L_{h-w+1}(x)}{L(x)}\left(\frac{(h-w)H_{h-w+1}(x)}{B_{h-w+1}(x)} \cdot \frac{(k+1)A(x)}{B(x)-L(x)} \cdot k \cdot \frac{H_{h-i}(x)}{B(x)-L(x)}\right) \\
&\quad + \frac{L_{h-i+1}(x)}{L(x)}\frac{(h-i)H_{h-i+1}(x)}{B_{h-i+1}(x)}, \quad i = 1, \ldots, w-1, &(4.3.6) \\
L'(x) &= -1 + \frac{L_{h-w+1}(x)}{L(x)}\Bigg(-\frac{(h-w)L_{h-w+1}(x)}{B_{h-w+1}(x)} \\
&\quad + (h-w)k \cdot \frac{H_{h-w+1}(x)}{B_{h-w+1}(x)} \cdot \frac{(k+1)A(x)}{B(x)-L(x)}\Bigg), &(4.3.7) \\
B'(x) &= -1 - \frac{(h-w)L_{h-w+1}(x)}{L(x)}, &(4.3.8) \\
HV'(x) &= -\frac{L_{h-w+1}(x)}{L(x)}\frac{(h-w)H_{h-w+1}(x)}{B_{h-w+1}(x)} \cdot \frac{(k+1)A(x)}{B(x)-L(x)}, &(4.3.9) \\
\lambda'(x) &= \frac{((B'(x)-L'(x))HV(x) - (B(x)-L(x))HV'(x))f_{k+1}(\lambda(x))}{HV(x)^2(f_k(\lambda(x)) + \lambda(x)e^{-\lambda(x)} \cdot \frac{\lambda(x)^{k-1}}{(k-1)!} - \frac{B(x)-L(x)}{HV(x)} \cdot e^{-\lambda(x)} \cdot \frac{\lambda(x)^k}{k!})}, &(4.3.10)
\end{aligned}
$$

and

$$
\begin{aligned}
L'_{h-i}(0) &= 0, \quad H'_{h-i}(0) = 0, \quad i = 2, \ldots, w-1, &(4.3.11) \\
L'_{h-1}(0) &= \frac{(h-1)L_h(0)}{L_h(0)+H_h(0)}, \quad H'_{h-1}(0) = \frac{(h-1)H_h(0)}{L_h(0)+H_h(0)}, &(4.3.12) \\
L'(0) &= -1, \quad B'(0) = -1, \quad HV'(0) = 0, \quad \lambda'(0) = 0, &(4.3.13)
\end{aligned}
$$

and for all $x \geq 0$,

$$L_h(x) = L(x) - \sum_{i=1}^{w-1} L_{h-i}(x), \quad H_h(x) = B(x) - L(x) - \sum_{i=1}^{w-1} H_{h-i}(x), \quad (4.3.14)$$

$$B_{h-i}(x) = L_{h-i}(x) + H_{h-i}(x), \quad \text{for every } 0 \leq i \leq w-1, \quad (4.3.15)$$

$$A(x) = \frac{\lambda(x)^{k+1}}{e^{\lambda(x)}(k+1)! f_{k+1}(\lambda(x))} HV(x). \quad (4.3.16)$$

The initial conditions are

$$B(0) = \bar{\mu}, \; L_{h-i}(0) = 0, \; H_{h-i}(0) = 0, \; \text{for all } 1 \leq i \leq w-1, \quad (4.3.17)$$

$$L(0) = \bar{\mu}(1 - f_k(\bar{\mu})), \; HV(0) = 1 - \exp(-\bar{\mu}) \sum_{i=0}^{k} \bar{\mu}^i / i!, \; \lambda(0) = \bar{\mu}. \quad (4.3.18)$$

Let $x = t/\bar{n}$ and let $L_{h-j}(x) = L_{t,h-j}/\bar{n}$. We call this the scaling of the variable $L_{t,h-j}$. Do the same scaling on all the other variables B_t, L_t, $H_{t,h-j}$, etc. We will apply the DE method to analyse the asymptotic values of these random variables in the process generated by the RanCore algorithm. First we fix some small constant $\epsilon > 0$, and define $\mathcal{D}(\epsilon)$ to be a connected open set depending on the value of ϵ, in which we apply the DE method. Let $\Gamma(t)$ denote the vector $(t, z_{L,h-w+1}, \ldots, z_{L,h-1}, z_{B,h-w+1}, \ldots, z_{B,h-1}, z_L, z_B, z_{HV})$. The subscript of z indicates which scaled random variable z is associated to. For instance, $z_{L,h-j}$ is a real number that is associated to $L_{t,h-j}/\bar{n}$, etc. Hence the right hand side of the differential equations (4.3.5)–(4.3.13) are multi-variable functions of $\mathbf{z}$, where $\mathbf{z}$ denotes the set of variables $z_{L,h-j}$, etc. The domain $\mathcal{D}(\epsilon)$ is defined by the set $\{\Gamma(t), t \geq 0\}$ such that $z_L > \epsilon$, $0 < z_{L,h-j} < z_{B,h-j}$, $z_{L,h-j} < z_L$ for all $j \geq 1$ and $z_B - z_L > \epsilon$. The proof is sketched as follows. We first show that the scaled random variables under consideration a.a.s. enter the domain $\mathcal{D}(\epsilon)$ before $t_0 = \epsilon' \bar{n}$ steps from the beginning of the process. From this point onwards, we analyse the asymptotic values of these variables using the DE method. In order to apply the DE method, we need to check that the three hypotheses in Theorem 2.2.1 hold. The first hypothesis clearly holds. To verify the second hypothesis, we need to show that the expected one step change of $L_{t,h-i}$ can be approximated within an error of $o(1)$ as some function f, such that f is a continuous multi-variable function of $t/\bar{n}$, $L_{t,h-i}/\bar{n}$, and other random variables under consideration at step t, also divided by $\bar{n}$, then do the same to the other random variables under consideration. We need to show that these f correspond to the multi-variable functions at the right hand side of (4.3.5)–(4.3.13). Lastly, to verify the third hypothesis, we first shift the process by letting $t := t - \epsilon' \bar{n}$ and choose the open set D in Theorem 2.2.1 (iii) to be the domian $\mathcal{D}(\epsilon)$ in which $t_0 = \epsilon' \bar{n}$ is shifted to 0. Then we show that these f are continuous and Lipschitz inside the shifted domain $\mathcal{D}(\epsilon)$. After we apply the DE method to analyse the random variables from step

t_0 up to the step when the scaled variables reach the boundary of the domain $\mathcal{D}(\epsilon)$, we continue the analysis of the process after the scaled random variables leave the domain.

We will use a separate argument to analyse the initial segment of the process since the real functions f under consideration are not Lipschitz at $\Gamma(0)$. For instance, let $f_{L,h-j}$ denote the multi-variable function on the right hand side of (4.3.5). Clearly, $f_{L,h-j}$ is discontinuous and non-Lipschitz at $z_{B,h-j} = 0$. However, as we will show later, the continuity and Lipschitz condition hold inside the domain $\mathcal{D}(\epsilon)$ for any given $\epsilon > 0$. We will also show that for any $\epsilon' > 0$, a.a.s. $\Gamma(\epsilon' \bar{n})$ is in $\mathcal{D}(\epsilon)$. We also show that $\Gamma(t)$ remains in $\mathcal{D}(\epsilon)$ until $L_t/\bar{n}$ reaches ϵ. We postpone these details of justification and first estimate the expected one step change of each variable which leads to the continuous functions f. We show that these changes lead to (4.3.5)–(4.3.18).

Let $\tau = \min\{t : HV_t = 0 \text{ or } L_t = 0\}$ and consider any $0 \le t < \tau$. At step $t+1$, a partition-allocation g_{t+1} is to be obtained by applying the RanCore algorithm to g_t. Let v be the ball randomly chosen by the algorithm from $\mathbb{L}$. Let $C(v) = c_{h-j}$ be the colour of v. If $j < w-1$, the algorithm removes another $h-j-1$ balls that are uniformly distributed among all balls with colour c_{h-j} since $g_t \in \mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1)$ as proved. If $j = w-1$, then the algorithm removes v together with another $h-w$ balls which are chosen u.a.r. from all balls of colour c_{h-w+1}. If the removal of the $h-w$ balls results in some heavy bins turning into light bins, these bins are removed and the balls remaining in these bins are put into $\mathbb{L}$.

Now we estimate the expected value of $L_{t+1,h-j} - L_{t,h-j}$ for any $1 \le j \le w-1$ and for any $0 \le t < \tau$ conditional on V_t, $\mathbf{M}_t$, $\mathbf{L}_t$ and the event $g_t \in \mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1)$. The probability that $C(v) = c_{h-j}$ is $L_{t,h-j}/L_t$. If $C(v) = c_{h-j}$, one ball of colour c_{h-j} contained in $\mathbb{L}$ is removed by the algorithm, and another $h-j-1$ balls of colour c_{h-j} are recoloured with c_{h-j-1} (or removed if $j = w-1$). The expected number of those balls that are contained in $\mathbb{L}$ is

$$\frac{(h-j-1)L_{t,h-i}}{B_{t,h-j}},$$

if $B_{t,h-j} > 0$, or 0 otherwise. By convention, let

$$\frac{(h-j-1)L_{t,h-j}}{B_{t,h-j}} = 0$$

if $B_{t,h-j} = 0$. Hence

$$\frac{L_{t,h-j}}{L_t}\left(-1 - \frac{(h-j-1)L_{t,h-j}}{B_{t,h-j}}\right)$$

is the negative contribution to $\mathbf{E}(L_{t+1,h-j} - L_{t,h-j} \mid V_t, \mathbf{M}_t, \mathbf{L}_t, g_t \in \mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1))$. The positive contribution to $\mathbf{E}(L_{t+1,h-j} - L_{t,h-j} \mid V_t, \mathbf{M}_t, \mathbf{L}_t, g_t \in \mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1))$ comes from the following two cases.

Case 1: $C(v) = c_{h-w+1}$. Then the algorithm removes v and another $h-w$ balls of colour c_{h-w+1}. $\mathbf{P}(C(v) = h-w+1) = L_{t,h-w+1}/L_t$. We first show that the contribution from the case that i of the $h-w$ removed balls lie in a bin containing at most $k+i$ balls for some $2 \le i \le h-w$ is negligible. For any i balls, the probability that they are in the same bin containing at most $k+i$ balls is

$$\left(\frac{H_{t,h-w+1}}{B_{t,h-w+1}}\right)^i O\left(\left(\frac{(k+i)}{B_t - L_t}\right)^i \sum_{j=1}^{i} A_{t,k+j}\right) = O(\bar{n}^{-i+1}) = o(1),$$

since for any particular ball removed by the algorithm among the $h-w$ balls, the probability that it is contained in a heavy bin is $H_{t,h-w+1}/B_{t,h-w+1}$, and $((k+i)/(B_t - L_t))^i$ is the probability that i balls are in a particular bin containing at most $k+i$ balls conditional on that it is contained in a heavy bin l, whereas $\sum_{j=1}^{i} A_{t,k+j}$ is the total number of heavy bins with at most $k+i$ balls.

It only remains to consider the contribution from the case that some ball removed by the algorithm is contained in a bin containing exactly $k+1$ balls. For any of these balls that the algorithm removes, the probability that it is in a bin containing exactly $k+1$ balls is

$$\frac{H_{t,h-w+1}}{B_{t,h-w+1}} \cdot \frac{(k+1)A_{t,k+1}}{B_t - L_t}.$$

The removal of a ball from a bin l containing $k+1$ balls causes l to become light. Since the balls of each colour are uniformly distributed among all balls in the heavy bins because $g_t \in \mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1)$, the expected number of balls of colour c_{h-j}, for $0 \le j \le w-1$, among the remaining k balls in the bin l is

$$k \cdot \frac{H_{t,h-j}}{B_t - L_t}.$$

There are in total $h-w$ balls of colour c_{h-w+1} other than v that are removed by the algorithm. Hence the expected contribution to $\mathbf{E}(L_{t+1,h-j} - L_{t,h-j} \mid V_t, \mathbf{M}_t, \mathbf{L}_t, g_t \in \mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1))$ is

$$(h-w) \cdot \frac{L_{t,h-w+1}}{L_t} \cdot \frac{H_{t,h-w+1}}{B_{t,h-w+1}} \cdot \frac{(k+1)A_{t,k+1}}{B_t - L_t} \cdot k \cdot \frac{H_{t,h-j}}{B_t - L_t}.$$

Case 2: $C(v) = c_{h-j+1}$. The algorithm then removes v and chooses another $h-j$ balls u.a.r. from those of colour c_{h-j+1}, and recolours them with c_{h-j}. Since $\mathbf{P}(C(v) = c_{h-j+1}) = L_{t,h-j+1}/L_t$, and conditional on $C(v) = c_{h-j+1}$, the expected number of balls of colour c_{h-j+1} that are in the light bins and are recoloured by the algorithm is

$$(h-j) \cdot \frac{L_{t,h-j+1}}{B_{t,h-j+1}}.$$

Hence the positive contribution to $\mathbf{E}(L_{t+1,h-j} - L_{t,h-j} \mid V_t, \mathbf{M}_t, \mathbf{L}_t, g_t \in \mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1))$ is

$$\frac{L_{t,h-j+1}}{L_t} \cdot (h-j) \cdot \frac{L_{t,h-j+1}}{B_{t,h-j+1}}.$$

Therefore

$$\begin{aligned}
&\mathbf{E}(L_{t+1,h-j} - L_{t,h-j} \mid V_t, \mathbf{M}_t, \mathbf{L}_t, g_t \in \mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1)) \\
&= \frac{L_{t,h-j}}{L_t}\left(-1 - \frac{(h-j-1)L_{t,h-j}}{B_{t,h-j}}\right) + \frac{L_{t,h-j+1}}{L_t} \cdot \frac{(h-j)L_{t,h-j+1}}{B_{t,h-j+1}} \\
&+ \frac{L_{t,h-w+1}}{L_t}\left(\frac{(h-w)H_{t,h-w+1}}{B_{t,h-w+1}} \cdot \frac{(k+1)A_{t,k+1}}{B_t - L_t} \cdot k \cdot \frac{H_{t,h-j}}{B_t - L_t}\right) + o(1), \quad (4.3.19)
\end{aligned}$$

for $j = 1, \ldots, w-1$. Then the right hand side of (4.3.19) gives the continuous multi-variable function f in $\mathcal{D}(\epsilon)$ which leads to (4.3.5). Similarly, by computing the expected changes of $H_{t,h-i}$, B_t, D_t, HV_t, conditional on V_t, $\mathbf{M}_t$, $\mathbf{L}_t$ and the event $g_t \in \mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1)$, and doing the same scaling of these variables, we obtain (4.3.6)–(4.3.9).

The following obvious equations

$$L_{t,h} = L_t - \sum_{i=1}^{w-1} L_{t,h-i}, \quad H_{t,h} = B_t - L_t - \sum_{i=1}^{w-1} H_{t,h-i}, \quad (4.3.20)$$

$$B_{t,h-i} = L_{t,h-i} + H_{t,h-i}, \quad \text{for every } h-w+1 \leq i \leq h \quad (4.3.21)$$

lead to (4.3.14) and (4.3.15). Since $g_t \in \mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1)$ for every t, by considering the allocation-partition algorithm that generates $\mathcal{P}(V_t, \mathbf{M}_t, \mathbf{L}_t, k+1)$, the degree sequence of the heavy vertices obeys the truncated multinomial distribution, which asymptotically converges in distribution to the truncated Poisson distribution as $HV_t \to \infty$ by [16, Lemma 1]. Let λ_t be the parameter of the truncated Poisson distribution. Then the average number of balls in heavy bins is $\lambda_t f_k(\lambda_t)/f_{k+1}(\lambda_t)$. So λ_t is the root of

$$\frac{\lambda_t f_k(\lambda_t)}{f_{k+1}(\lambda_t)} - \frac{B_t - L_t}{HV_t} = 0, \quad \text{for all } t = 0, 1, 2, \ldots, \tau.$$

Note that the uniqueness of the root follows by the monotonicity of $\lambda_t f_k(\lambda_t)/f_{k+1}(\lambda_t)$ discussed in the proof of Proposition 4.3.3. Thus,

$$A_{t,k+1} \sim \frac{e^{-\lambda_t}\lambda_t^{k+1}}{(k+1)! f_{k+1}(\lambda_t)} HV_t, \quad (4.3.22)$$

which gives (4.3.18). Let $x = t/\bar{n}$, and let $B(x) = B_t/\bar{n}$, $L_{h-i}(x) = L_{t,h-i}/\bar{n}$, $H_{h-i}(x) = H_{t,h-i}/\bar{n}$, $L(x) = L_t/\bar{n}$, $A(x) = A_{t,k+1}/\bar{n}$ and $HV(x) = HV_t/\bar{n}$. Let $\mu(x)$ denote $(B(x) - L(x))/HV(x)$. Then $\lambda(x)$ is an implicit function of x satisfying

$$\lambda(x) f_k(\lambda(x)) = \mu(x) f_{k+1}(\lambda(x)), \quad \text{for all valid } x. \quad (4.3.23)$$

So (4.3.10) follows by first taking the derivative of both sides of (4.3.23),

$$\lambda'(x)\cdot f_k(\lambda(x))+\lambda(x)\cdot\frac{d\,f_k(\lambda)}{d\,\lambda}\bigg|_{\lambda=\lambda(x)}\cdot\lambda'(x)=\mu'(x)\cdot f_{k+1}(\lambda(x))+\mu(x)\cdot\frac{d\,f_{k+1}(\lambda)}{d\,\lambda}\bigg|_{\lambda=\lambda(x)}\cdot\lambda'(x),$$

together with

$$\frac{d\,f_k(\lambda)}{d\,\lambda}=e^{-\lambda}\cdot\frac{\lambda^{k-1}}{(k-1)!}\quad\text{and}\quad\mu'(x)=\frac{(B'(x)-L'(x))HV(x)-(B(x)-L(x))HV'(x)}{HV(x)^2}.$$

Now we justify the third hypothesis in Theorem 2.2.1 for the shifted process. Recall that the domain $\mathcal{D}(\epsilon)$ in which the DE method is applied is defined by the set of $\{\Gamma(t) : t \geq 0\}$ such that $z_L > \epsilon$, $0 < z_{L,h-j} < z_{B,h-j}$, $z_{L,h-j} < z_L$ for all $j \geq 1$, and $z_B - z_L > \epsilon$. Let $f_{L,h-j}$ denote the multi-variable function on the right hand side of (4.3.5). Clearly $f_{L,h-j}$ (and other real functions under consideration) is continuous in $\mathcal{D}(\epsilon)$ since $z_L > 0$ and $z_{B,h-j} > 0$. Next we show that these functions are Lipschitz in $\mathcal{D}(\epsilon)$. It is easy to see that all these continuous functions (e.g. $f_{L,h-j}$) are expressed as sums of products of ratios of real numbers such as z_L, etc. We only need to prove that each summand is Lipschitz. Note that the Lipschitz condition of this type of summand can only be violated if the denominator is not bounded away from 0. Hence we only need to check summands of the following form

$$f(z_{L,h-j}, z_L, z_{B,h-j}) = \frac{z_{L,h-j}}{z_L}\cdot\frac{z_{L,h-j}}{z_{B,h-j}},$$

where $z_L > \epsilon$, $z_{B,h-j} > 0$, $0 < z_{L,h-j} < z_{B,h-j}$, $z_{L,h-j} < z_L$ by the definition of $\mathcal{D}(\epsilon)$. Thus,

$$\left|\frac{\partial\, f}{\partial\, z_{L,h-j}}(z_{L,h-j}, z_L, z_{B,h-j})\right| = \frac{2z_{L,h-j}}{z_L z_{B,h-j}} < \frac{2}{z_L} < \frac{2}{\epsilon}, \tag{4.3.24}$$

$$\left|\frac{\partial\, f}{\partial\, z_L}(z_{L,h-j}, z_L, z_{B,h-j})\right| = \frac{z_{L,h-j}^2}{z_L^2 z_{B,h-j}} < \frac{1}{\epsilon}, \tag{4.3.25}$$

$$\left|\frac{\partial\, f}{\partial\, z_{B,h-j}}(z_{L,h-j}, z_L, z_{B,h-j})\right| = \frac{z_{L,h-j}^2}{z_L z_{B,h-j}^2} < \frac{1}{\epsilon}. \tag{4.3.26}$$

Therefore, the Lipschitz condition is satisfied.

We first show that for any small enough $\epsilon' > 0$, a.a.s. $\Gamma(\epsilon'\bar{n}) \in \mathcal{D}(\epsilon)$ by taking $z_L = L_{\epsilon'\bar{n}}/\bar{n}$, etc.

In order to show that for any small enough $\epsilon' > 0$, a.a.s. $\Gamma(\epsilon'\bar{n}) \in \mathcal{D}(\epsilon)$, it is enough to show that for any given small enough $\epsilon' > 0$, a.a.s. $L_{\epsilon'\bar{n},h-j}/\bar{n} > \alpha_1$ and $H_{\epsilon'\bar{n},h-j}/\bar{n} > \alpha_2$ for some $\alpha_1, \alpha_2 > 0$, for any $0 \leq j \leq w-1$. Initially, all balls are coloured with c_h. A special case of [16, Lemma 1] shows that there exists a constant $\beta' = e^{-\Theta(k)} > 0$ depending only on k and $\bar{\mu}$ such that the proportion of balls that are initially light is a.a.s. at least β'.

The number of light balls removed in each step is $O(1)$, so there exists a constant $\rho > 0$, depending only on k and $\bar{\mu}$, such that $L_t \geq L_0/2$ a.a.s. and the proportion of balls coloured with c_h at step t among all light balls is at least $1/2$ a.a.s. for all $t \leq \rho\bar{n}$. The function of ρ is to restrict the choice of ϵ' so that it is not too large.

In the process, if a ball coloured with c_{h-j} is recoloured by the algorithm with c_{h-j-1}, we say that a ball of colour c_{h-j} is destroyed and a ball of colour c_{h-j-1} is created. First we know that initially the proportion of balls that are light is a.a.s. at least $\beta' > 0$. Since $L_t \geq L_0/2$ for all $t \leq \rho$, there exists $\beta > 0$, such that a.a.s. in all steps up to ρ, the proportion of balls that are light is at least β. Now we show that for any $0 < \epsilon' < \rho$, and any $j \geq 0$, there exist some $\alpha_1, \alpha_2 > 0$ such that a.a.s. $L_{t,h-j} > \alpha_1\bar{n}$ and $H_{t,h-j} > \alpha_2\bar{n}$ for all $\epsilon'\bar{n} \leq t \leq \rho\bar{n}$, by induction on j. For $j = 0$ this is trivially true. Now assume this holds for some $0 \leq j < w - 1$.

By the inductive hypothesis, there exists $\alpha_1, \alpha_2 > 0$ such that a.a.s. $L_{t,h-j} \geq \alpha_1\bar{n}$ $H_{t,h-j} \geq \alpha_2\bar{n}$ for any $\epsilon'\bar{n}/2 \leq t \leq \rho\bar{n}$. We also have that a.a.s. the proportion of light balls in each step up to ρ is at least β. Define the stopping time T to be the minimum $t \geq \epsilon'\bar{n}/2$ such that either $L_{t,h-j} < \alpha_1\bar{n}$ or $H_{t,h-j} < \alpha_2\bar{n}$ or the proportion of balls that are light becomes less than β. Then a.a.s. $T \geq \rho\bar{n}$. Let $t \wedge T$ denote $\min\{t, T\}$. Next we show that there exists $\gamma > 0$ such that $L_{t,h-j}/L_t \geq \gamma$ for all $t \in [\epsilon'\bar{n}/2, \rho\bar{n} \wedge T]$. This is trivially true if $j = 0$, since $L_{t,h}/L_t \geq 1/2$ for any $\epsilon'\bar{n}/2 \leq t \leq \rho\bar{n}$. Assume $j \geq 1$. Since $L_{t,h-j} \geq \alpha_1\bar{n}$ for all $\epsilon'\bar{n}/2 \leq t \leq \rho\bar{n} \wedge T$, there exists $\gamma > 0$ such that $L_{t,h-j}/L_t \geq \gamma$ for all $\epsilon'\bar{n}/2 \leq t \leq \rho\bar{n} \wedge T$. Thus, the probability that the light ball chosen by the algorithm is coloured with c_{h-j} is at least γ in every step from $\epsilon'\bar{n}/2$ to $\rho\bar{n} \wedge T$. Conditional on $T \geq \rho\bar{n}$, for any t with $\epsilon'\bar{n} \leq t \leq \rho\bar{n}$, the probability that the number of balls with colour c_{h-j-1} created from step $\epsilon'\bar{n}/2$ to t is less than $\gamma(h-j-1)\epsilon'\bar{n}/4$ is at most $\exp(-\Theta(\epsilon'^2\bar{n}))$ by Lemma 4.3.5. A proportion of those balls would be destroyed or not counted by $H_{t,h-j-1}$. These include:

(a) those that are light balls at the time they were created;

(b) those that were heavy when they were created but are turned into light by the algorithm;

(c) those that are removed or recoloured because a light ball with colour c_{h-j-1} was chosen and removed by the algorithm.

It is easy to check that the proportion from case (a) and (b) is a.a.s. $O(1/k)$, since the proportion of balls that are light is $\beta = O(\beta') = O(e^{-\Theta(k)}) = O(1/k)$ in each step, and the proportion from case (c) is a.a.s. at most $1/2$, since the proportion of balls with colour c_h among all light light balls is at least $1/2$. So for sufficiently large k, conditional on $T \geq \rho$, the probability that $H_{t,h-j-1}/\bar{n} < \gamma(h-j-1)\epsilon'\bar{n}/12$ is at most $\exp(-\Theta(\epsilon'^2\bar{n}))$ for all $\epsilon'\bar{n} \leq t \leq \rho$. Since a.a.s. $T \geq \rho$, we have that the probability that $T < \rho$ or that there exists some t with $\epsilon'\bar{n} \leq t \leq \rho\bar{n}$ such that $H_{t,h-j-1}/\bar{n} < \gamma(h-j-1)\epsilon'\bar{n}/12$ is at most

$o(1) + \rho\bar{n}\exp(-\Theta(\epsilon'^2\bar{n})) = o(1)$. This proves that there exists $\alpha_2' > 0$ such that a.a.s. for all $\epsilon'\bar{n} \le t \le \rho\bar{n}$, $H_{t,h-j-1}/\bar{n} > \alpha_2'$. Similarly, we can show that there exists $\alpha_1' > 0$ such that a.a.s. for all $\epsilon'\bar{n} \le t \le \rho\bar{n}$, $L_{t,h-j-1}/\bar{n} > \alpha_1'$. Hence by induction, for any sufficiently small $\epsilon' > 0$, there exists $\alpha_1 > 0$ and $\alpha_2 > 0$ such that $L_{\epsilon'\bar{n},h-j}/\bar{n} > \alpha_1$ and $H_{\epsilon'\bar{n},h-j}/\bar{n} > \alpha_2$ for all $0 \le j \le w-1$.

Now we have shown that for any $\epsilon' > 0$, a.a.s. $L_{\epsilon'\bar{n}}/\bar{n} > \alpha_1$ and $H_{\epsilon'\bar{n}}/\bar{n} > \alpha_2$ for some constants $\alpha_1, \alpha_2 > 0$. Recall that $z_{L,h-j}$ corresponds to the scaled variable $L_{\epsilon'\bar{n},h-j}/\bar{n}$. Thus we have for some $\alpha > 0$, $\alpha < z_{L,h-j} < z_{B,h-j}$ and $z_{L,h-j} < z_L$ for all $j \ge 1$. The constraints $z_L > \epsilon$ and $z_B - z_L > \epsilon$ are clearly satisfied when ϵ' is sufficiently small. Hence a.a.s. the shifted domain $\mathcal{D}(\epsilon)$ contains some open set D such that there is no $\mathbf{z}$ that are on the boundaries of both D and $\mathcal{D}(\epsilon)$, and a.a.s. every $\Gamma(0)$ (shifted from $\Gamma(\epsilon'\bar{n})$) for which $\mathbf{P}(L_0 = z_L\bar{n}, z_L \in \Gamma(0)) \ne 0$, etc. is contained in D. Thus we are ready to apply Theorem 2.2.1 to analyse the time-shifted process from step 0, or equivalently, to analyse the original process $(g_t)_{t \ge \epsilon'\bar{n}}$. Note that Theorem 2.2.1 (iii) requires the shifted $\mathcal{D}(\epsilon)$ to contain some neighbourhood of such $\Gamma(0)$ deterministically instead of asymptotically almost surely. However, this does not cause any problem since we can restrict our discussion to those $g_{\epsilon'\bar{n}}$ so that the shifted $\Gamma(0)$ is in D, and apply the DE method only to these $g_{\epsilon'\bar{n}}$, whereas the probability that $g_{\epsilon'\bar{n}}$ is not in this category converges to 0 as $\bar{n} \to \infty$. Hence the a.a.s. properties obtained by applying the DE method hold in our case, too.

Let $t_1(\epsilon)$ denote the first time the scaled random variables reaches the boundary of the domain $\mathcal{D}(\epsilon)$ after they entered the domain. Our proof will be almost complete after we show the following statements.

(S1) The initial derivatives (4.3.11)–(4.3.13) approximate the changes of the random variables at the beginning of the process.

(S2) There is a unique solution to the differential equation system (4.3.5)–(4.3.18) and the solution is in the domain $\mathcal{D}(\epsilon)$ for all $0 < x \le x(\epsilon)$, where $x(\epsilon)$ is the point at which the solution reaches the boundary of $\mathcal{D}(\epsilon)$ for the first time after it entered $\mathcal{D}(\epsilon)$.

(S3) In all steps $0 \le t \le t_1(\epsilon)$, the deviation of each random variable from what is given by the solution to the differential equation system is a.a.s. bounded by $o(\bar{n})$.

(S4) The time $t_1(\epsilon)$ coincides with the time at which $L(x)$ or $B(x) - L(x)$ decreases to ϵ, where $L(x)$ and $B(x)$ are solution functions of the differential equation system.

Note that (S3) is proved later by applying Theorem 2.2.1 and the uniform convergence for any $\epsilon' > 0$; (S4) shows in which way the solution functions to the differential equation system leave the domain $\mathcal{D}(\epsilon)$.

We need the following two theorems in our proof. The first theorem gives sufficient conditions under which a solution to a given differential equation system with initial conditions exists.

Theorem (Peano's Existence Theorem [46]) Let $n \geq 1$ be a fixed integer and D an open subset of $\mathbb{R}^{n+1}$. Assume $\varphi_i : D \mapsto \mathbb{R}$ for $i = 1, \ldots, n$ are continuous functions on D. Consider the initial value problem

$$\frac{d\,x_i}{d\,t} \;=\; \varphi_i(t, x_1, \ldots, x_n), \quad \text{for all } i = 1, \ldots, n,$$

with $x_i(b) = a_i$ for $i = 1, \ldots, n$, where $(b, a_1, \ldots, a_n) \in D$. Then there exists $\delta > 0$ and n functions $x_1, \ldots, x_n$ such that for every $i = 1, \ldots, n$, $x_i : (b - \delta, b + \delta) \mapsto \mathbb{R}$ is a continuous function that satisfies $x_i'(t) = \varphi_i(t, x_1(t), \ldots, x_n(t))$ for all $t \in (b - \delta, b + \delta)$.

The next theorem gives sufficiently conditions under which there is a unique solution to a given differential equation system with initial conditions.

Theorem (Picard-Lindelöf Theorem) Assume the differential equation system with initial conditions is defined as in the Peano's Existence Theorem. Assume further that φ_i for $i = 1, \ldots, n$ are continuous in t and Lipschitz continuous in x_i for all $i = 1, \ldots, n$. Then for some $\delta > 0$ and for every $i = 1, \ldots, n$, there exists a unique function x_i such that $x_i : (b - \delta, b + \delta) \mapsto \mathbb{R}$ is a continuous function that satisfies $x_i'(t) = \varphi_i(t, x_1(t), \ldots, x_n(t))$ for all $t \in (b - \delta, b + \delta)$.

A well-known proof of the Picard-Lindelöf Theorem uses the Picard iteration. By the method of Picard iteration, it immediately gives the following lemma.

Lemma (Stability of ODE systems) Assume all the hypotheses in Picard-Lindelöf Theorem are satisfied. Let $x_i^{(1)}$ be the unique solutions extended in a Lipschitz continuous domain containing $(b, a_1^{(1)}, \ldots, a_n^{(1)})$ and $x_i^{(2)}$ be the unique solutions extended in a Lipschitz continuous domain containing $(b, a_1^{(2)}, \ldots, a_n^{(2)})$. Let I be any bounded connected open subset of $\mathbb{R}$ containing b such that for any $i = 1, \ldots, n$, both solutions $x_i^{(1)}(t)$ and $x_i^{(2)}(t)$ can be uniquely extended to any $t \in I$. Let $\mathbf{a}^{(i)} = (a_1^{(i)}, \ldots, a_n^{(i)})$ for $i = 1, 2$. Then if for some $\epsilon > 0$, $||\mathbf{a}^{(1)} - \mathbf{a}^{(2)}|| \leq \epsilon$, then there exists $C > 0$ such that for any $t \in I$ and any $i = 1, \ldots, n$, $|x_i^{(1)}(t) - x_i^{(2)}(t)| \leq C\epsilon$.

We first prove (S3), assuming (S1) and (S2). Recall that $\bar{\mu} = h\bar{m}/\bar{n}$. By [16, Lemma 1], the number of vertices with degree d is a.a.s. asymptotically $(e^{-\bar{\mu}}\bar{\mu}^d/d!)\bar{n}$. So a.a.s. $L_0 \sim (1 - f_k(\bar{\mu}))\bar{\mu}\bar{n}$, $B_0 \sim \bar{\mu}\bar{n}$ and $HV_0 \sim f_{k+1}(\bar{\mu})\bar{n}$. We also have $L_{h-i,0} = 0$, $H_{h-i,0} = 0$ for all $1 \leq i \leq w - 1$, $L_{h,0} = L_0$ and $H_{h,0} = B_0 - L_0$. By scaling these variables we obtain that the initial deviation of each variable from the initial conditions (4.3.17) and (4.3.18) of the differential equation system is a.a.s. $o(\bar{n})$. For any random variable under consideration, its one step change is bounded by some absolute constant, for instance, the one step change of L_t is bounded by h, etc. We have already shown that the initial deviation of each variable at $t = 0$ is a.a.s. bounded by $o(\bar{n})$. Hence the (accumulative) deviation of $L_{\epsilon'\bar{n}}$ (and the other random variables) is $O(\epsilon'\bar{n})$, where $O()$ is independent of ϵ' and $\bar{n}$. Let $y(x) = (L_{h-w+1}(x), \ldots, L_{h-1}(x), B_{h-w+1}(x), \ldots, B_{h-1}(x), L(x), B(x), HV(x))$ be the solution to the differential equations (4.3.5)–(4.3.18). (Note that the other functions, e.g. $L_h(x)$, $\lambda(x)$, can be determined by $L(x)$, $HV(x)$, etc.). By Theorem 2.2.1 (b), we deduce that the values of sequences $(L_{t,h-i})_{\epsilon'\bar{n}\le t\le t_1(\epsilon)}$, $(H_{t,h-i})_{\epsilon'\bar{n}\le t\le t_1(\epsilon)}$, etc., are approximated by $y(x)$, in the sense that a.a.s. $L_{t,h-i} = \bar{n}L_{h-i}(t/\bar{n}) + o(\bar{n}) + O(\epsilon'\bar{n})$, etc., for all $\epsilon'\bar{n} \le t \le t_1(\epsilon)$. The error term $o(\bar{n})$ follows by Theorem 2.2.1 (b) and $O(\epsilon'\bar{n})$ is obtained by the following three facts. Firstly, as we have shown, the deviation of each random variable from what is given by the initial conditions (4.3.17) and (4.3.18) is a.a.s. $o(\bar{n})$. Secondly, the accumulative changes of each variable from step 0 to step $\epsilon'\bar{n}$ is $O(\epsilon'\bar{n})$. Lastly, the difference of values of each solution function (e.g. $L(x)$, etc.) to the differential equations system at $x = 0$ and $x = \epsilon'$ is at most $O(\epsilon')$ since $|f_L|$, $|f_{L,h-j}|$, etc. are bounded for all sufficiently small $x \ge 0$. Note that the boundedness follows because the solution functions to the differential equation system are in the domain $\mathcal{D}(\epsilon)$ for every $x > 0$ by (S2). By the above three facts, the error $O(\epsilon'\bar{n})$ follows by the Lemma of the stability of ordinary differential equation systems. Since the error $O(\epsilon'\bar{n})$ holds uniformly for any small enough $\epsilon' > 0$, we have the uniform convergence by letting $\epsilon' \to 0$. Hence we have $L_{\lfloor x\bar{n}\rfloor,h-i} = \bar{n}L_{h-i}(x) + o(\bar{n})$, etc., for all $0 < x < t_1(\epsilon)/\bar{n}$, whereas $L_0 = \bar{n}L(0) + o(\bar{n})$. This proves (S3). Theorem 2.2.1 also tells that $t_1(\epsilon)$ coincides with the time at which $y(t/\bar{n})$ leaves the domain $\mathcal{D}(\epsilon)$.

Now we prove (S1). Let the RanCore algorithm run for $t = \epsilon'\bar{n}$ steps from the initial partition-allocation. Then obviously $L_{t,h-j} = O(\epsilon'\bar{n})$ for all $j \ge 1$. Thus the probability that the algorithm chooses any light ball with colour c_{h-j} for $j \ge 1$ is $O(\epsilon')$ at each step up to step t and hence for any $j \ge 2$, the total number of balls with colour c_{h-j} that have ever been created up to step t is a.a.s. $O(\epsilon' t)$. Hence $L_{t,h-j} = O(\epsilon' t)$ and $H_{t,h-j} = O(\epsilon' t)$ for all $j \ge 2$. Recall that $L_{h-j}(x)$ is the function scaled from $L_{t,h-j}$ by taking $x = t/\bar{n}$ and $L_{h-j}(x) = L_{t,h-j}/\bar{n}$. Hence for all $j \ge 2$,

$$
\begin{aligned}
L'_{h-j}(0) &= \lim_{\epsilon'\to 0} \frac{L_{h-j}(\epsilon') - L_{h-j}(0)}{\epsilon'} = \lim_{\epsilon'\to 0} \frac{L_{\epsilon'\bar{n},h-j}/\bar{n}}{\epsilon'} = \lim_{\epsilon'\to 0} \frac{O(\epsilon'^2)}{\epsilon'} = 0,\\
H'_{h-j}(0) &= \lim_{\epsilon'\to 0} \frac{H_{\epsilon'\bar{n},h-j}/\bar{n}}{\epsilon'} = 0.
\end{aligned}
$$

This verifies (4.3.11). The total number of light balls with colour c_h chosen by the algorithm is then a.a.s. $(1 - O(\epsilon'))t$ up to step t. Every such choice results in removing one light ball coloured c_h and choosing a set V' of $h-1$ balls u.a.r. from those with colour c_h and recolouring them with c_{h-1}. The number of light and heavy balls coloured c_h in each step up to step t is $L_{0,h} - O(\epsilon'\bar{n})$ and $H_{0,h} - O(\epsilon'\bar{n})$ respectively. Hence in each step, the probability that a particular ball in V' lies in $\mathbb{L}$ is $L_{0,h}/B_{0,h} + O(\epsilon')$. Then a.a.s.

$L_{t,h-1} = tL_{0,h}/B_{0,h} + O(\epsilon' t)$. Similarly, we can show that $H_{t,h-1} = tH_{0,h}/B_{0,h} + O(\epsilon' t)$. This verifies (4.3.12). The verification of the first three derivatives of (4.3.13) follows immediately from the above argument that verifies (4.3.12). The third derivative follows by noticing that (4.3.10) holds for $x = 0$ as well by the argument below (4.3.21). This proves (S1).

Next we show (S2). Intuitively this must be true since the functions $f_{L,h-j}$, etc. are continuous and uniformly Lipschitz inside the domain $\mathcal{D}(\epsilon)$, and the initial derivatives (4.3.11)–(4.3.13) "force" the solution to enter the domain. Then the solution will be extended uniquely to the boundary of $\mathcal{D}(\epsilon)$. In our proof, we will first extend the function $f_{L,h-j}$ into a larger domain and then prove the existence of solutions to the "extended" differential equation system. Then we show that all these solutions enter the domain $\mathcal{D}(\epsilon)$ for sufficiently small $x > 0$ because of the initial derivatives. Eventually this leads to prove the uniqueness of a solution before the solution leaves the domain $\mathcal{D}(\epsilon)$.

We have already shown that the functions $f_{L,h-j}$, etc. are continuous and Lipschitz in the domain $\mathcal{D}(\epsilon)$ for any $\epsilon > 0$. We first extend the functions $f_{L,h-j}$, etc. to continuous functions in the domain $D := \{(z_{L,h-w+1}, \ldots, z_{L,h-1}, z_{B,h-w+1}, \ldots, z_{B,h-1}, z_L, z_B, z_{HV}) : z_L > \epsilon, z_B - z_L > \epsilon\}$. Let $\mathbf{z}_0$ denote the vector in D with $z_L = \bar{\mu}(1 - f_k(\bar{\mu}))$, $z_B = \bar{\mu}$, $z_{HV} = 1 - \exp(-\bar{\mu}) \sum_{i=0}^{k} \bar{\mu}^i / i!$, $z_{L,h-j} = z_{B,h-j} = 0$ for all $1 \leq j \leq w-1$. Clearly $\mathbf{z}_0$ is an interior point in D but is on the boundary of $\mathcal{D}(\epsilon)$. The reason that we extend f to the domain D is because we can prove the existence of solutions to the extended differential equation system using the Peano's Existence Theorem. However, as we will show later, for growing x, all the solutions enter the domain $\mathcal{D}(\epsilon)$ right away and hence all these solutions are solutions to the original differential equation system.

To extended the functions such as $f_{L,h-j}$ to continuous functions in D, it is enough to extend the function

$$f(z_{L,h-j}, z_{B,h-j}, z_L) = \frac{z_{L,h-j}^2}{z_L z_{B,h-j}}$$

to the domain D, for the same reason as the argument above (4.3.24). Define

$$f(z_{L,h-j}, z_{B,h-j}, z_L) = \begin{cases} z_{L,h-j}^2 / z_L z_{B,h-j} & \text{if } 0 \leq z_{L,h-j} \leq z_{B,h-j}, z_{B,h-j} > 0 \\ z_{B,h-j}/z_L & \text{if } z_{L,h-j} > z_{B,h-j} \geq 0 \\ |z_{L,h-j}|^2 / z_L |z_{B,h-j}| & \text{otherwise.} \end{cases} \tag{4.3.27}$$

Clearly, f is continuous inside the boundary $0 \leq z_{L,h-j} \leq z_{B,h-j}$. By symmetry and the fact that f takes same values for points outside $|z_{L,h-j}| \leq |z_{B,h-j}|$ as those on the boundary, f is a continuous function in D. We extend the differential equation system (4.3.5)–(4.3.18) by extending the functions on the right hand side of (4.3.5)–(4.3.9) as (4.3.27). We take (4.3.5) as an example, and the other functions are extended similarly. Let

$$\mathbf{z} = (z_{L,h-w+1}, \ldots, z_{L,h}, z_{H,h-w+1}, \ldots, z_{H,h}, z_{B,h-w+1}, \ldots, z_{B,h}, z_A, z_B, z_L, z_{HV}, z_\lambda)$$

denote a point in $\mathbb{R}^{3w+2}$ and let $\mathbf{z}_0$ denote the point $\mathbf{z}$ with $z_{L,h-j} = 0$, $z_{H,h-j} = 0$ and $z_{B,h-j} = 0$ for all $j \geq 1$ and $z_{L,h} = z_L = \bar{\mu}(1 - f_k(\bar{\mu}))$, $z_{H,h} = z_B - z_L = \bar{\mu} f_k(\bar{\mu})$, $z_{HV} = f_{k+1}(\bar{\mu})$, $z_\lambda = \bar{\mu}$ and $z_A = e^{-\bar{\mu}}\bar{\mu}^{k+1}/(k+1)!$. Note that the definition of $\mathbf{z}_0$ is actually equivalent to the $\mathbf{z}_0$ defined above (4.3.27), because the other components of $\mathbf{z}_0$, e.g. $z_{H,h-j}$, z_λ and z_A are restricted in the (extended and original) differential equation system by the constraints (4.3.14)–(4.3.16).

Let $f(z_{L,h-j}, z_{B,h-j}, z_L)$ be defined as in (4.3.27). Then for any $i = 1, \ldots, w-1$, define the extension of $f_{L,h-i}$ as follows.

$$
\begin{aligned}
f_{L,h-i}(\mathbf{z}) = \quad = \quad & -\frac{z_{L,h-i}}{z_L} - (h-i-1) f(z_{L,h-i}, z_{B,h-i}, z_L) \\
& + \left((h-w)\frac{(k+1)z_A}{z_B - z_L} \cdot k \cdot \frac{z_{H,h-i}}{z_B - z_L}\right)\left(\frac{z_{L,h-w+1}}{z_L} - f(z_{L,h-w+1}, z_{B,h-w+1}, z_L)\right) \\
& + (h-i) f(z_{L,h-i+1}, z_{B,h-i+1}, z_L).
\end{aligned}
$$

Extend (4.3.5) to $L'_{h-i}(x) =$

$$f_{L,h-i}(L_{h-w+1}(x), \ldots, L_h(x), H_{h-w+1}(x), \ldots, H_h(x), B_{h-w+1}(x), \ldots, B_h(x), A(x), B(x), L(x)),$$

and extend (4.3.6)–(4.3.9) analogously. Call the resulting differential equation system the *extended differential equation system*. We have shown that the extended functions $f_{L,h-i}$, etc. are continuous at a small neighbourhood of $\mathbf{z}_0$. By the Peano's Existence Theorem, there exists a solution to the extended differential equation system in a small neighbourhood of $\mathbf{z}_0$.

Next we show that there is a unique solution to the original differential equation system. Given $y(x) = (L_{h-w+1}(x), \ldots, L_{h-1}(x), B_{h-w+1}(x), \ldots, B_{h-1}(x), L(x), B(x), HV(x))$ as a solution to the extended differential equation system, let $\Gamma_y(x)$ denote the vector

$$(\lfloor x\bar{n} \rfloor, L_{h-w+1}(x), \ldots, L_{h-1}(x), B_{h-w+1}(x), \ldots, B_{h-1}(x), L(x), B(x), HV(x)).$$

We show that for any solution $y(x)$, for any sufficiently small $x > 0$, $\Gamma_y(x) \in \mathcal{D}(\epsilon)$.

The partial derivatives (4.3.24)–(4.3.26) imply that there is a uniform Lipschitz constant at a small neighbourhood of $\mathbf{z}_0$ (by neighbourhood of $\mathbf{z}_0$ we mean $B(\mathbf{z}_0, \delta) \setminus \{\mathbf{z}_0\}$ for some sufficiently small $\delta > 0$). Hence for any $x_1 > 0$, and any

$$(\lfloor x_1 \bar{n} \rfloor, z_{L,h-w+1}, \ldots, z_{L,h-1}, z_{B,h-w+1}, \ldots, z_{B,h-1}, z_L, z_B, z_{HV}) \in \mathcal{D}(\epsilon),$$

there is a unique solution to the differential equation system with initial conditions $L_{h-j}(x_1) = z_{L,h-j}$, etc., and the solution can be extended arbitrarily close to the boundary of $\mathcal{D}(\epsilon)$.

For all sufficiently small $x > 0$, we have $L(x) > 0$, $A(x) > 0$, $B(x) - L(x) > 0$, and $L'_{h-1}(x) > 0$, $H'_{h-1}(x) > 0$ since $L'_{h-1}(0) > 0$, $H'_{h-1}(0) > 0$ by (4.3.12). Hence $L_{h-1}(x) > 0$

and $H_{h-1}(x) > 0$ for all sufficiently small $x > 0$. We show by induction that $L_{h-j}(x) > 0$ and $H_{h-j}(x) > 0$ for all $j \geq 1$ and all sufficiently small $x > 0$. We have shown this is true for $j = 1$. Assume it is true for some $1 \leq j \leq w - 2$. Then the derivative (4.3.6) implies that for any sufficiently small $x > 0$,

$$\begin{aligned} H'_{h-j-1}(x) \quad \geq \quad & \frac{L_{h-j-1}(x)}{L(x)} \left(-\frac{(h-j-2)H_{h-j-1}(x)}{B_{h-j-1}(x)} \right) \\ & -\frac{L_{h-w+1}(x)}{L(x)} \left(\frac{(h-w)H_{h-w+1}(x)}{B_{h-w+1}(x)} \cdot \frac{(k+1)A(x)}{B(x)-L(x)} \cdot k \cdot \frac{H_{h-j-1}(x)}{B(x)-L(x)} \right) + c, \end{aligned}$$

for some constant $c > 0$, since $L_{h-j}(x) > 0$ and $H_{h-j}(x) > 0$. We call a function $c(x)$ bounded if $|c(x)| < M$ for some constant $M > 0$. Thus, the derivative of $H_{h-j-1}(x)$ is at least $c + c(x)H_{h-j-1}(x)$ for some $c > 0$ and bounded $c(x)$. We also have $H_{h-j-1}(0) = 0$. Thus, $H_{h-j-1}(x) > 0$ for all sufficiently small $x > 0$.

Since $L_{h-j}(x) > 0$ for all sufficiently small $x > 0$, then by (4.3.5), $L'_{h-j-1}(x) > 0$ for all sufficiently small $x > 0$. This is because $L_{h-w+1}(x)$ and $H_{h-w+1}(x)$ can only turn negative if $L_{h-w+2}(x)$ become negative, which never happens as long as $L_{h-w+1}(x) \geq 0$ by (4.3.5). Since $L_{h-j}(0) \geq 0$ for any $j \geq 0$, we have $L_{h-w+1}(x) \geq 0$ for all sufficiently small $x > 0$. Then (4.3.5) shows that the derivative of $L_{h-j-1}(x)$ is at least $c + c(x)L_{h-j-1}(x)$ for some $c > 0$ and bounded $c(x)$. We also have $L_{h-j-1}(0) = 0$. These imply that $L_{h-j-1}(x) > 0$ for all sufficiently small $x > 0$.

Now we have shown that for all sufficiently small $x > 0$, for any $j \geq 1$, $L_{h-j}(x) > 0$ and $H_{h-j}(x) > 0$. Thus, $L_{h-j}(x) < B_{h-j}(x)$ and $L_{h-j}(x) < L(x)$ by (4.3.14). Hence for any sufficiently small $x > 0$, $\Gamma_y(x) \in \mathcal{D}(\epsilon)$.

Next we show that there is a unique solution to the extended differential equation system. Assume not. Let $y^{(i)}$ for $i = 1, 2$ be two distinct solutions to the extended differential equation system. Then there exists $x_1 > 0$ such that $y^{(1)}(x_1) \neq y^{(2)}(x_1)$. If both $\Gamma_{y^{(1)}}(x_1)$ and $\Gamma_{y^{(2)}}(x_1)$ are in $\mathcal{D}(\epsilon)$, then there must exists $0 < x_2 < x_1$, such that both $\Gamma_{y^{(1)}}(x_2)$ and $\Gamma_{y^{(2)}}(x_2)$ are in $\mathcal{D}(\epsilon)$ and $y^{(1)}(x_2) \neq y^{(2)}(x_2)$, since otherwise, there is unique solution to the differential equations with initial conditions $y^{(1)}(x_2)$ at $x = x_2$, which contradicts $y^{(1)}(x_1) \neq y^{(2)}(x_1)$. Thus, we can assume that $\Gamma_{y^{(1)}}(x) \notin \mathcal{D}(\epsilon)$ for all $0 \leq x \leq x_1$. However, this contradicts that for any solution $y(x)$ to the extended differential equation system, $\Gamma_y(x) \in \mathcal{D}(\epsilon)$ for all sufficiently small $x > 0$. Thus, we cannot find any solution $y(x)$ for which there exists $x_1 > 0$ such that $\Gamma_y(x) \notin \mathcal{D}(\epsilon)$ for all $0 \leq x \leq x_1$. This shows that there is a unique solution $y(x)$, for all $0 \leq x \leq x(\epsilon)$, to the extended differential equation system, where $y(x) \in \mathcal{D}(\epsilon)$ for all $0 < x \leq x(\epsilon)$. By the definition of the extension of $f(z_L, z_{L,h-j}, z_{B,h-j})$ in (4.3.27), for any $x > 0$, the continuous functions $f_{L,h-j}$, etc. are the same as their extention whenever $y(x)$ is in the domain $\mathcal{D}(\epsilon)$. It is also clear that these functions are the same as their extention at $x = 0$, by (4.3.11)–(4.3.12) and (4.3.27). Hence $y(x)$ for $0 \le x \le x(\epsilon)$ is the unique solution to the original differential equation system.

It only remains to prove (S4). The proof uses a similar argument as part of the proof of (S2) that shows the uniqueness of the solution. Let $\mathcal{U}(\epsilon) = \{x \ge 0 : \ L(x) > \epsilon$ and $B(x) - L(x) > \epsilon\}$. For all $x \in \mathcal{U}(\epsilon)$, the derivative of $H_h(x)$ equals $c(x)H_h(x)$, for some bounded $c(x)$. (Note that there is an upper bound of $|c(x)|$ that depends only on ϵ.) We also have $H_h(0) > 0$. Thus, $H_h(x) > 0$ for all $x \in \mathcal{U}(\epsilon)$. Then the derivative of $L_h(x)$ is at least c plus $c(x)L_h(x)$ for some $c > 0$ and bounded $c(x)$. This implies that $L_h(x) > 0$ for all $x \in \mathcal{U}(\epsilon)$. Then following the same argument in the proof of (S2), we can show that for each $j \ge 1$, the derivative of $L_{h-j}(x)$ is at least $c + c(x)L_{h-j}(x)$ and the derivative of $H_{h-j}(x)$ is at least $c + c(x)H_{h-j}(x)$ for some $c > 0$ and bounded $c(x)$. It follows then that $L_{h-j}(x) > 0$ and $H_{h-j}(x) > 0$ for all $x \in \mathcal{U}(\epsilon)$. It then follows that $\Gamma_y(x) \in \mathcal{D}(\epsilon)$ for all $x \in \mathcal{U}(\epsilon)$. This finishes the proof of (S4).

By the Picard-Lindelöf theorem, the solution to the differential equation system (4.3.5)–(4.3.18) can be extended arbitrarily close to the boundary inside which some Lipschitz condition is satisfied for the continuous functions f_L, etc. We have shown that for any $\epsilon > 0$, some Lipschitz condition is satisfied in $\mathcal{D}(\epsilon)$. Together with (S4), we have that the solution to the differential equations can be extended uniquely to all $x > 0$ such that $L(x) > 0$ and $B(x) - L(x) > 0$. Let x^* denote the real number at which $L(x)$ or $B(x) - L(x)$ reaches 0. Then there is a unique solution to the differential equation for all $0 \le x < x^*$. For notational convenience, let Y_t denote the set of random variables under consideration in step t and let $y(x)$ denote the set of solution functions to the differential equation system. By (S2) and (S3), $Y_{\lfloor x\bar{n} \rfloor} = \bar{n}y(x) + o(\bar{n})$, for all $0 \le x \le x(\epsilon)$ and by $Y_{\lfloor x\bar{n} \rfloor} = \bar{n}y(x) + o(\bar{n})$ we mean $L_{\lfloor x\bar{n} \rfloor} = \bar{n}L(x) + o(\bar{n})$, and the same to all the other random variables under consideration. We also have $|Y_{\lfloor x\bar{n} \rfloor} - Y_{\lfloor x(\epsilon)\bar{n} \rfloor}| = O((x - x(\epsilon))\bar{n})$ for all $x(\epsilon) \le x \le x^*$ since for each variable the one step change is bounded by $O(1)$. On the other hand, we have $|y(x) - y(x(\epsilon))| = O(x - x(\epsilon))$ since the right hand side of (4.3.5)–(4.3.9) are uniformly bounded for all $0 \le x < x^*$. Hence for all $x(\epsilon) \le x < x^*$, a.a.s.

$$\begin{aligned} |Y_{\lfloor x\bar{n} \rfloor} - \bar{n}y(x)| &\le |Y_{\lfloor x\bar{n} \rfloor} - Y_{\lfloor x(\epsilon)\bar{n} \rfloor}| + |Y_{\lfloor x(\epsilon)\bar{n} \rfloor} - \bar{n}y(x(\epsilon))| + |\bar{n}y(x(\epsilon)) - \bar{n}y(x)| \\ &= O((x^* - x(\epsilon))\bar{n}) + o(\bar{n}), \end{aligned}$$

where $O()$ is independent of ϵ and x^*. It then follows that for all $0 \le x < x^*$, a.a.s. $Y_{\lfloor x\bar{n} \rfloor} = \bar{n}y(x) + O((x^* - x(\epsilon))\bar{n}) + o(\bar{n})$. Since $x(\epsilon) \to x^*$ as $\epsilon \to 0$ by the definition of x^*, we have $Y_{\lfloor x\bar{n} \rfloor} = \bar{n}y(x) + o(\bar{n})$ uniformly for all $0 \le x < x^*$ by letting $\epsilon \to 0$. Thus we have that the solution $y(x)$ to the differential equations (4.3.5)–(4.3.10) approximates the asymptotic values of all random variables under consideration for all $t = \lfloor x\bar{n} \rfloor$, $0 \le x < x^*$, where x^* is the smallest real number such that $L(x^*) = 0$ or $B(x^*) - L(x^*) = 0$.

Now we show that if $\bar{\mu} \ge ck$ for some $c > 1$, then for sufficiently large k (depending on the value of c), the function $L(x)$ reaches 0 before $B(x) - L(x)$ reaches 0 and we estimate an upper bound of the value of x^*. Let $H(x) = B(x) - L(x)$. Clearly $L(x) = \sum_{i=0}^{w-1} L_{h-i}(x)$ and $H(x) = \sum_{i=0}^{w-1} H_{h-i}(x)$. Then (4.3.7) and (4.3.8) immediately lead to

$$L'(x) \leq -1 + \frac{(h-1)k(k+1)A(x)}{H(x)}, \quad H'(x) \geq -(h-1) - \frac{(h-1)k(k+1)A(x)}{H(x)} \tag{4.3.28}$$

Let $\delta = (1 - f_k(\bar{\mu}))\bar{\mu}$. Then the initial conditions give $L(0) = \delta$ and $H(0) = \bar{\mu} - \delta$. Since $\bar{\mu} \geq ck$ for some $c > 1$, $\delta = \exp(-\Theta_c(k))$. Let $\lambda(x)$ be defined as in (4.3.23). By the argument in the proof of Proposition 4.3.3, we may assume $|\mu(x) - \lambda(x)| \leq 1$ as k sufficiently large. Then $A(x)/H(x) = \exp(-\Theta_{c'}(k))$ as long as $\mu(x) \geq c'k$ for some $c' > 1$. Choose k sufficiently large (depending only on the value of c) such that $\delta \leq 1$, and $|\lambda(x) - \mu(x)| \leq 1$ and $hk(k+1)A(x)/H(x) \leq 1/2$ as long as $\mu(x) \geq \bar{\mu} - 3h$. Since $HV(x) \leq 1$ for any $x < x^*$, $\mu(x) = H(x)/HV(x) \geq H(x)$. Then $L'(x) \leq -1/2$ and $H'(x) \geq -(h-1) - 1/2$ for all x smaller than the value at which $H(x)$ reaches $\bar{\mu} - 3h$. Then it is easy to check that for any $0 \leq x < \min\{x^*, 3\delta\}$,

$$\begin{aligned} H(x) &\geq \bar{\mu} - \delta + 3\delta\left(-(h-1) - \frac{1}{2}\right) \geq \bar{\mu} - 3h, \\ L(x) &\leq \delta - \frac{x}{2}. \end{aligned} \tag{4.3.29}$$

Thus we have $x^* < 3\delta$ since otherwise $3\delta \leq x^*$ and $L(3\delta) \leq \delta - 3\delta/2 < 0$, contradicting the definition of x^*. This shows that $L(x)$ reaches 0 before $H(x)$ reaches 0. We also have that $L'(x) \leq -1/2$ for all $x < x^*$ since $x^* < 3\delta$. Let $t^* = \lfloor x^* \bar{n} \rfloor$. We have shown that for any $x < x^*$, a.a.s. $H_{\lfloor x\bar{n} \rfloor, h-j} = \bar{n} H_{h-j}(x) + o(\bar{n})$. We also have $|H_{\lfloor x^*\bar{n} \rfloor, h-j} - H_{\lfloor x\bar{n} \rfloor, h-j}| = O((x^* - x)\bar{n})$ since the one step change of $H_{t,h-j}$ is $O(1)$. Hence $H_{t^*,h-j} = \bar{n} \lim_{x \to x^*} H_{h-j}(x) + o(\bar{n})$, where the limit is taking by letting x approach to x^* from $x < x^*$. For notational convenience, let $H_{h-j}(x^*) := \lim_{x \to x^*} H_{h-j}(x)$ and $HV(x^*) := \lim_{x \to x^*} HV(x)$.

Let δ_1 denote the number of light balls remaining at step $t^* = x^* \bar{n}$. Then $\delta_1 = o(\bar{n})$ since $L(x^*)$ reaches 0. Let $\Delta = \max\{4\delta_1, \log \bar{n}\}$. Then $\Delta = o(\bar{n})$ and $\Delta \to \infty$ as $n \to \infty$. Applying Lemma 4.3.4 with $X_0 = L_{t^*}$, $X_n = L_{t^*+\Delta}$, $n = \Delta$, $\delta = -1/2$ and $c = h$, we have a.a.s. $L_{t^*+\Delta} \leq \delta_1 - \Delta/4 \leq 0$. Hence, a.a.s. the RanCore algorithm terminates before step $t^* + \Delta$. Recall the the algorithm terminates at step τ. So $\tau \leq t^* + \Delta$. In addition, $|H_{\tau,h-j} - H_{t^*,h-j}| \leq h\Delta = o(\bar{n})$. Thus, we have that a.a.s. H has a non-empty $(w, k+1)$-core $\widehat{H}$. Recall that n and m_{h-j} denote the number of vertices and hyperedges of size $h-j$ in $\widehat{H}$. Then a.a.s. the number of vertices in $\widehat{H}$ is $\bar{n}HV(x^*) + o(\bar{n})$, and the number of hyperedges of size $h - j$ in $\widehat{H}$ is $\bar{n}H_{h-j}(x^*)/(h-j) + o(\bar{n})$. Since $HV(x^*) > 0$ and $H_{h-j}(x^*) > 0$, we have a.a.s. $n \sim \alpha\bar{n}$ and $m_{h-j} \sim \beta_{h-j}\bar{n}$, where $\alpha = HV(x^*)$ and $\beta_{h-j} = H_{h-j}(x^*)/(h-j)$, which are determined by the solution to (4.3.5)–(4.3.18). ∎

We have shown that the partition-allocation g_τ output by the RanCore algorithm, if it is nonempty, is distributed as $\mathcal{P}(V_\tau, \mathbf{M}_\tau, \mathbf{0}, k+1)$ conditional on V_τ and $\mathbf{M}_\tau$. Let n denote $|V_\tau|$ and m_{h-j} denote $|M_{\tau,h-j}|/(h-j)$ for all $0 \le j \le w-1$. Then we can relabel elements in V by $[n]$ and M_{h-j} by $\cup_{j=0}^{w-1}[m_{h-j}] \times [h-j]$ in a canonical way. Then we can simplify the notation of the partition-allocation model and write $\mathcal{P}(n, \mathbf{m}, \mathbf{0}, k+1)$ by convenience.

Lemma 4.3.6 *Assume $c_1 k < h\bar{m}/\bar{n} < c_2 k$ for some constants $c_2 > c_1 > 1$. Let H be a random multihypergraph in $\mathcal{M}_{\bar{n},\bar{m},h}$. Then a.a.s. H has a nonempty $(w, k+1)$-core with average degree $O(k)$ provided k is sufficiently large.*

Proof Let $\bar{\mu} = h\bar{m}/\bar{n}$. Since $\bar{\mu} > c_1 k$ for some $c_1 > 1$, the existence of a non-empty $(w, k+1)$-core follows directly by the argument below (4.3.28). Let x^* be as defined in the statement of Theorem 4.2.1 and let δ be the proportion of vertices in H with degree at most k. We have shown that $\delta = O(e^{-\Theta(k)})$ below (4.3.28) and $x^* \le 3\delta$ below (4.3.29). Let $B(x)$ and $HV(x)$ be defined the same as those functions in (4.3.5)–(4.3.18) for $0 \le x \le x^*$. Then clearly $B(x^*) \le B(0)$ since $B'(x) \le -1$ for all $0 \le x \le x^*$. We also have $HV'(x) \ge -1$ for all $0 \le x \le x^*$ when k is large enough, by following the similar argument as deriving $L'(x) \le -1/2$ for sufficiently large k below (4.3.28). So $HV(x^*) \ge HV(0) - x^*$ for sufficiently large k. Since $HV(0) = f_{k+1}(\bar{\mu}) = 1 - O(e^{-\Theta(k)})$ and $x^* = O(e^{-\Theta(k)})$, we have $HV(x^*) = 1 - O(e^{-\Theta(k)})$. Let $\mu(x) = B(x)/HV(x)$. Then clearly, $\mu(0) = (1+O(e^{-\Theta(k)}))h\bar{m}/\bar{n} = O(k)$ since $h\bar{m}/\bar{n} < c_2 k$, and

$$\mu(x^*) \le \frac{B(0)}{HV(x^*)} = \frac{B(0)}{HV(0)}(1 + O(e^{-\Theta(k)})) = O(k).$$

By Theorem 4.2.1, the average degree of the $(w, k+1)$-core of H is asymptotically $\mu(x^*)$, which is bounded by $O(k)$. ∎

The following lemma gives a lower bound on the size of the $(w, k+1)$-core of a random h-multihypergraph.

Lemma 4.3.7 *Assume $c_1 k < h\bar{m}/\bar{n} < c_2 k$ for some constants $c_2 > c_1 > 1$. Let H be a random multihypergraph in $\mathcal{M}_{\bar{n},\bar{m},h}$. Then a.a.s. the number of vertices in the $(w, k+1)$-core of H is $(1 - O(e^{-\Theta(k)}))\bar{n}$.*

Proof Let n denote the number of vertices in the $(w, k+1)$-core of H. We have shown, by the argument below (4.3.29), that RanCore algorithm terminates within $O(e^{-\Theta(k)}\bar{n})$ steps when applied to H, and in each step at most h heavy bins can disappear. So a.a.s. $n = (1 - O(e^{-\Theta(k)}))\bar{n}$. ∎

We need the following lemma before proving Theorem 4.2.3, the existence of a sharp threshold of property $\mathcal{T}$.

Lemma 4.3.8 *Assume $c_1k < h\bar{m}/\bar{n} < c_2k$ for some constants $c_2 > c_1 > 1$. Let $\epsilon > 0$ be fixed. Let H_1 be a random multihypergraph in $\mathcal{M}_{\bar{n},\bar{m},h}$ and $H_2 \in \mathcal{M}_{\bar{n},\bar{m}+\epsilon\bar{n},h}$. Let n_1 and n_2 be the number of vertices in the $(w, k+1)$-core of H_1 and H_2 respectively. Then a.a.s. we have $|n_1 - n_2| = O(e^{-\Theta(k)}\epsilon\bar{n})$.*

Proof Since $c_1k < h\bar{m}/\bar{n} < c_2k$, by Lemma 4.3.6, the $(w, k+1)$-core $\widehat{H}_1$ of H_1 exists and the average degree of $\widehat{H}_1$ is $O(k)$. Let H_2 be a random uniform multihypergraph obtained from $H_1 \cup \mathcal{E}$, where $\mathcal{E}$ is a set of $\epsilon\bar{n}$ hyperedges, each of which is a multiset of h vertices, each of which u.a.r. chosen from $[\bar{n}]$. Then $H_2 \in \mathcal{M}_{\bar{n},\bar{m}+\epsilon\bar{n},h}$. We say that the hyperedges in $\mathcal{E}$ are *marked*, and the other hyperedges in H_2 are *unmarked*. Define a random process $(H_t^{(1)}, H_t^{(2)})_{t\geq 0}$ as follows.

(i) The process starts with $(H_0^{(1)}, H_0^{(2)}) = (H_1, H_2)$.

(ii) The RanCore algorithm is applied to $H_t^{(2)}$ for every $t \geq 0$. The process $(H_t^{(1)}, H_t^{(2)})_{t\geq 0}$ stops when the RanCore algorithm running on $(H_t^{(2)})_{t\geq 0}$ terminates.

(iii) For every $t \geq 0$, if a marked hyperedge x in $H_{t-1}^{(2)}$ is updated to x', then x' remains marked in $H_t^{(2)}$ and $H_t^{(1)}$ is defined as $H_{t-1}^{(1)}$; if a marked hyperedge x is removed, let $H_t^{(1)} = H_{t-1}^{(1)}$.

(iv) For every $t \geq 0$, if an unmarked hyperedge x in $H_{t-1}^{(2)}$ is updated or removed, do the same operation to x in $H_{t-1}^{(1)}$ and define $H_t^{(1)}$ to be the resulting hypergraph.

We call the random process $(H_t^{(i)})_{t\geq 0}$ for $i = 1, 2$ generated by $(H_t^{(1)}, H_t^{(2)})_{t\geq 0}$ the H_i-*process*. Note that the H_1-process is not equivalent to running the RanCore algorithm on H_1, since the light balls are not chosen u.a.r. in each step.

Instead of analysing $(H_t^{(1)}, H_t^{(2)})_{t\geq 0}$ directly, we consider $(g_t^{(1)}, g_t^{(2)})_{t\geq 0}$, the corresponding process obtained by considering the pairing-allocation model. Recall that H_1 can be represented as dropping $h\bar{m}$ unmarked balls u.a.r. into $\bar{n}$ bins with balls evenly partitioned into $\bar{m}$ groups randomly and H_2 can be represented as dropping $h\epsilon\bar{n}$ partitioned marked balls into H_1. The partition-allocations $g_0^{(1)}$ and $g_0^{(2)}$ are obtained by splitting all balls contained in light bins of H_1 and H_2 respectively into bins containing exactly one ball. Conditional on $L_0^{(i)}$, $V_0^{(i)}$, $\mathbf{M}_0^{(i)}$ and $\mathbf{L}_0^{(i)}$, $g_0^{(i)}$ is distributed as $\mathcal{P}(V_0^{(i)}, \mathbf{M}_0^{(i)}, \mathbf{L}_0^{(i)}, k+1)$ for $i = 1, 2$ and all balls in $g_0^{(1)}$ are unmarked.

Let $\bar{\mu}$ denote the average degree of H_1. Define $L_t^{(i)}$, $HV_t^{(i)}$, $\mathbf{m}_t^{(i)}$ and $\mathbf{L}_t^{(i)}$, etc., for $i = 1, 2$, the same way as in the proof of Theorem 4.2.1, for the H_i-process. Let τ be the time the H_2-process terminates. It is easy to show that $g_\tau^{(1)}$ is distributed as $\mathcal{P}(HV_\tau^{(1)}, \mathbf{m}_\tau^{(1)}, \mathbf{l}_\tau^{(1)}, k+1)$

conditional on the values of $HV_\tau^{(1)}$, $\mathbf{m}_\tau^{(1)}$ and $\mathbf{l}_\tau^{(1)}$, since whenever a light ball is chosen, even not uniformly at random, it results in recolouring or removal of heavy balls that are uniformly chosen at random. We will later let the RanCore algorithm be run on $g_\tau^{(1)}$ in the following steps and apply the DE method to analyse the asymptotic behavior of this process.

First we show that $\tau = O(e^{-k}\bar{n})$. The solution of the differential equation system (4.3.5)–(4.3.18) tells the asymptotic value of $L_t^{(2)}$ in every step t. Let $x_{(2)}^*$ be the smallest root of $L^{(2)}(x) = 0$. Since $L^{(2)}(0) = O(e^{-\Theta(k)})$ and $L'^{(2)}(x) < -1/2$ for all $0 \le x < x_{(2)}^*$, provided k sufficiently large, by the argument below (4.3.28), we have $x_{(2)}^* = O(e^{-\Theta(k)})$ and so $\tau = O(e^{-\Theta(k)}\bar{n})$.

Next we show that $n_2 - n_1 = O(e^{-\Theta(k)}\epsilon\bar{n})$, assuming the following three statements.

(S1) The number of balls that are unmarked and light in $g_0^{(1)}$ but not in $g_0^{(2)}$ is bounded by $O(e^{-\Theta(k)}\epsilon\bar{n})$.

(S2) The number of bins that begin heavy in the H_1-process and become light in that process but remain heavy in the H_2-process up to step τ is $O(e^{-\Theta(k)}\epsilon\bar{n})$.

(S3) $L_\tau^{(1)} = O(e^{-\Theta(k)}\epsilon\bar{n})$.

Run the Rancore algorithm on $g_\tau^{(1)}$. The differential equation system (4.3.5)–(4.3.18) tells the asymptotic values of the various random variables in $g_t^{(1)}$ for all $t \ge \tau$. Let $x_{(1)}^*$ be the smallest root of $L^{(1)}(x) = 0$. Since $L^{(1)}(\tau/\bar{n}) = O(e^{-k}\epsilon)$ by (S3) and $L'^{(1)}(x) \le -1/2$ for all $\tau/\bar{n} \le x < x_{(1)}^*$, provided k sufficiently large by the argument below (4.3.28), we have $x_{(1)}^* - \tau/\bar{n} = O(e^{-k}\epsilon)$. We also have $HV'(x) \ge -1$ for sufficiently large k for all $\tau/\bar{n} \le x < x_{(1)}^*$ as explained in Lemma 4.3.6. So $HV_\tau^{(1)} - n_1 = O(e^{-k}\epsilon\bar{n})$. Since $n_2 - HV_\tau^{(1)}$ counts the number of bins that are, or become light in the H_1-process but stay heavy in the H_2-process, it follows from (S1) and (S2) that $n_2 - HV_\tau^{(1)} = O(e^{-k}\epsilon\bar{n})$. So $n_1 - n_2 = O(e^{-k}\epsilon\bar{n})$.

It only remains to prove (S1)–(S3). We first show that (S3) follows directly from (S1) and (S2). $L_\tau^{(1)}$ counts two types of light balls. The first type comes from balls that are unmarked and light in $g_0^{(1)}$ but not in $g_0^{(2)}$. By (S1), the number of these balls is a.a.s. $O(e^{-\Theta(k)}\epsilon\bar{n})$. The second type comes from balls that begin heavy and become light in the H_1-process but stay heavy in the H_2-process. By (S2), the number of these balls is a.a.s. $k \cdot O(e^{-\Theta(k)}\epsilon\bar{n}) = O(e^{-\Theta(k)}\epsilon\bar{n})$. Thereby (S3) follows.

Next we show (S1). At step 0, clearly the set of unmarked light balls in $g_0^{(2)}$ is a subset of those in $g_0^{(1)}$. The number of light balls in $g_0^{(1)}$ is a.a.s. $(1 - f_k(\bar{\mu}))\bar{\mu}\bar{n} = O(e^{-\Theta(k)}\bar{n})$ as shown in the proof of Theorem 4.2.1 and hence the number of light vertices of H_1 is a.a.s. $O(e^{-\Theta(k)}\bar{n})$. Since each multihyperedges in $\mathcal{E}$ is a random multihyperedges, the expected number of those which contains a light vertex in H_1 is $O(e^{-\Theta(k)}\epsilon\bar{n})$, hence the number of light vertex in H_1 that become heavy after the hyperedges in $\mathcal{E}$ being dropped is a.a.s. $O(e^{-\Theta(k)}\epsilon\bar{n})$ and each of these vertex/bin contains at most k unmarked balls. Thus (S1) follows.

Now we show (S2). Recall that H_1 is represented as dropping $h\bar{m}$ unmarked balls u.a.r. into $\bar{n}$ bins and H_2 is obtained by dropping $h\epsilon\bar{n}$ extra marked balls u.a.r. into the $\bar{n}$ bins in H_1. Let $\widehat{H}_1$ be the $(w, k+1)$-core of H_1. Then the number of bins that begin heavy in the H_1-process and become light in that process but remain heavy in the H_2-process up to step τ is at most the number of bins/vertices not in $\widehat{H}_1$ which receive at least one marked balls after dropping $h\epsilon\bar{n}$ marked balls u.a.r. into the $\bar{n}$ bins. By Lemma 4.3.7, the number of vertices/bins in $\widehat{H}_1$ is a.a.s. $(1-O(e^{-\Theta(k)}))\bar{n}$. Then for each marked ball, the probability that it is dropped into a bin not in $\widehat{H}_1$ is $O(e^{-\Theta(k)})$. By Lemma 4.3.4, the number of marked balls dropped into bins not in $\widehat{H}_1$ is a.a.s. $O(e^{-\Theta(k)}\epsilon\bar{n})$. Hence the number of bins that are not in $\widehat{H}_1$ and receive at least one marked balls is a.a.s. $O(e^{-\Theta(k)}\epsilon\bar{n})$. ∎

Proof of Theorem 4.2.3. Let H_1 be a random uniform multihypergraph with average degree $\bar{\mu}$ and let H_2 be a random uniform multihypergraph obtained from $H_1 \cup \mathcal{E}$, where $\mathcal{E}$ is a set of $\epsilon\bar{n}$ hyperedges, each of which is a multiset of h vertices, each of which is uniformly chosen from $[\bar{n}]$.

Let $m^{(i)}_{h-j}$, $i = 1, 2$, be the number of hyperedges in the $(w, k+1)$-core of H_i. We first show that

$$\sum_{j=0}^{w-1}(w-j)m^{(2)}_{h-j} - \sum_{j=0}^{w-1}(w-j)m^{(1)}_{h-j} \geq w\epsilon\bar{n}/2.$$

Clearly the $(w, k+1)$-core of H_1 is a subgraph of the $(w, k+1)$-core of H_2. Let $[n_1]$ denote the set of vertices in the $(w, k+1)$-core of H_1. By Lemma 4.3.7, a.a.s. $n_1 = (1-O(e^{-\Theta(k)}))\bar{n}$. Then for any hyperedge $x \in \mathcal{E}$, the probability that all vertices in x are contained in $[n_1]$ is $1 - O(e^{-\Theta(k)})$. So the expected number of hyperedges in $\mathcal{E}$ lying completely in $[n_1]$ is $(1 - O(e^{-\Theta(k)}))\epsilon\bar{n}$. By the Chernoff bound, orginally given in [17, Theorem 1], we have a.a.s. the number of hyperedges in $\mathcal{E}$ lying completely in $[n_1]$ is at least $\epsilon\bar{n}/2$ for sufficiently large k. So it follows immediately that

$$\sum_{j=0}^{w-1}(w-j)m^{(2)}_{h-j} - \sum_{j=0}^{w-1}(w-j)m^{(1)}_{h-j} \geq w\epsilon\bar{n}/2.$$

For simplicity, let $S(i)$ denote $\sum_{j=0}^{w-1}(w-j)m^{(i)}_{h-j}$ for $i = 1, 2$. Then a.a.s.

$$\frac{S(2)}{n_2} - \frac{S(1)}{n_1} \geq \frac{(S(1) + w\epsilon\bar{n}/2) - S(1) \cdot n_2/n_1}{n_2}.$$

By Lemma 4.3.8, a.a.s. $n_2 - n_1 = O(e^{-\Theta(k)})\epsilon\bar{n}$, i.e. $n_2/n_1 = 1 + f(k)\epsilon$ for some function $f(k) = O(e^{-\Theta(k)})$. Then a.a.s.

$$\frac{S(2)}{n_2} - \frac{S(1)}{n_1} \geq \frac{w\epsilon\bar{n}/2 - O(f(k)\epsilon S(1))}{n_2} \geq w\epsilon/4 > 0, \tag{4.3.30}$$

for sufficiently large k and for every $\epsilon > 0$, since $S(1) = O(k)\bar{n}$ and $n_2 = (1 - O(e^{-\Theta(k)}))\bar{n}$.

Recall that $\kappa(H_i)$ denotes $\sum_{j=0}^{w-1}(w-j)m_{h-j}^{(i)}/n_i$ for $i = 1, 2$. By Theorem 4.2.1, for given $h > w > 0$ and sufficiently large k, a.a.s. $\kappa(\widehat{H}) = c(\bar{\mu}) + o(1)$, where $c(\bar{\mu})$ is a constant depending only on $\bar{\mu}$. The inequality (4.3.30) implies that $c(\bar{\mu})$ is an increasing function of $\bar{\mu}$. So there exists a unique critical value of $\bar{\mu}$ such that a.a.s. $\kappa(\widehat{H}) = k + o(1)$ and so there exists a threshold function $\bar{m} = f(\bar{n})$ of $\mathcal{M}_{\bar{n},\bar{m},h}$ for the graph property $\mathcal{T}$. Then this holds as well in $\mathcal{G}_{\bar{n},\bar{m},h}$ since the probability of a multihypergraph in $\mathcal{M}_{\bar{n},\bar{m},h}$ being simple is $\Omega(1)$ as explained in the beginning of Section 4.3, and so any event a.a.s. true in $\mathcal{M}_{\bar{n},\bar{m},h}$ is a.a.s. true in $\mathcal{G}_{\bar{n},\bar{m},h}$. ∎

The differential equations in Theorem 4.2.1 are only used in a theoretical way to show properties of the $(w, k+1)$-core, and we do not have an analytic solution. However, they can be numerically solved when the values of h, w, k and μ are given. Thus, we can numerically find the critical value of $\widetilde{\mu}$, such that $\sum_{j=0}^{w-1}(w-j)m_{h-j} \sim kn$. Table 4.3 gives the results of some computations, where h, w and k are given, $\widetilde{\mu}$ is the critical value and $\widehat{\mu}$ is the corresponding average degree of $\widehat{H}$. Note that $\widehat{\mu}$ must be at least hk/w. It follows from the trivial upper bound of the orientability threshold given in the introduction part that $\widetilde{\mu}$ is at most hk/w.

h	w	k	$\widetilde{\mu}$	$\widehat{\mu}$
3	2	4	5.485	6.65086
3	2	10	14.766	15.5872
3	2	40	59.991	60.0773
10	2	4	19.99999	20.0003

Table 4.1: Some numerical computation results

4.4 The (w, k)-orientability of the $(w, k+1)$-core

In this section we prove Corollary 4.2.7 assuming Theorem 4.2.5, and study the basic network flow formulation of the problem that is used in the next section to prove Theorem 4.2.5.

For the rest of the chapter, let $\epsilon > 0$ and $k \geq 2$ be fixed. Without loss of generality, we may assume that $\epsilon < \frac{1}{2}$. By the hypothesis of Theorem 4.2.5, we consider only $\mathbf{m}$ such that $\sum_{j=0}^{w-1}(w-j)m_{h-j} \leq kn - \epsilon n$. We may also assume that $\sum_{j=0}^{w-1}(w-j)m_{h-j} \geq kn - 2\epsilon n$ since otherwise, by Theorem 4.2.3, we can simply add random hyperedges so that the assumption holds. (Note that we can add random hyperedges because the (w, k)-orientability is a decreasing property. Recall that a property is called decreasing if the property holds in all subgraphs of a (hyper)graph G whenever G has the property.) Let

$$D = \sum_{j=0}^{w-1}(h-j)m_{h-j}, \quad m = \sum_{j=0}^{w-1} m_{h-j}, \quad \mu = \frac{D}{n}. \tag{4.4.1}$$

Since

$$D \cdot \frac{1}{h-w+1} \leq \sum_{j=0}^{w-1}(w-j)m_{h-j} \leq D \cdot \frac{w}{h}, \quad m \leq \sum_{j=0}^{w-1}(w-j)m_{h-j} \leq wm,$$

and

$$kn - 2\epsilon n \leq \sum_{j=0}^{w-1}(w-j)m_{h-j} \leq kn - \epsilon n, \tag{4.4.2}$$

we have

$$h(k-1)/w \leq \mu = D/n \leq (h-w+1)k, \quad \frac{(k-1)n}{w} \leq m \leq \Big(k - \frac{1}{2}\Big)n. \tag{4.4.3}$$

Let $f(\bar{n})$ be the threshold function of property $\mathcal{T}$. Let H be a random h-hypergraph on $\bar{n}$ vertices and $m \leq f(\bar{n})(1-\delta)$ hyperedges for some absolute constant $\delta > 0$. Therefore there exists $\epsilon > 0$ such that a.a.s. $\sum_{j=0}^{w-1}(w-j)m_{h-j} \leq kn - \epsilon n$ in the $(w, k+1)$-core $\widehat{H}$. Without loss of generality we may assume that (4.4.3) is satisfied with the same argument as that above (4.4.1). Let $\bar{\mu}$ be the average degree of H. We recall at this point the trivial upper bound

$$\bar{\mu} \leq hk/w \tag{4.4.4}$$

mentioned in the introduction as a requirement for (w, k)-orientability.

In the rest of the chapter, whenever we refer to the probability space $\mathcal{H}(n, \mathbf{m}, k+1)$ or $\mathcal{M}(n, \mathbf{m}, k+1)$, we assume $\mathbf{m}$ satisfies (4.4.3).

Lemma 4.4.1 *Let $c_1 > 1$ be a constant that can be dependent on k. Then there exists a constant $0 < \gamma = \varphi(k, c_1)$ depending only on k and c_1, such that a.a.s. there exists no $S \subset V(H)$, with $|S| < \gamma\bar{n}$ and at least $c_1|S|$ hyperedges partially contained in S. More specifically, when $c_1 \geq 2$ and $c_1 < h^2e^2\bar{\mu}$, we may choose $\gamma = \varphi(k, c_1) = (c_1/h^2e^2\bar{\mu})^2$.*

Proof Let s be any integer such that $0 < s < n$ and let $r = s/n$. Let Y denote the number of S with $|S| = s$ and at least $c_1 s$ hyperedges partially contained in S. The probability for a given hyperedge to be partially contained in S is at most $\binom{h}{2}(s/\bar{n})^2 < h^2 r^2$. Then the probability that there are at least $c_1 s$ such hyperedges is at most

$$\binom{\bar{m}}{c_1 s}(hr)^{2c_1 s}.$$

Since there are $\binom{\bar{n}}{s}$ ways to choose S,

$$\begin{aligned}
\mathbf{E}(Y) &= \sum_{s \le \gamma\bar{n}} \binom{\bar{n}}{s}\binom{\bar{m}}{c_1 s}(hr)^{2c_1 s} \\
&\le \sum_{\ln\bar{n} \le s \le \gamma\bar{n}} \left(\frac{e\bar{n}}{s}\right)^s \left(\frac{e\bar{m}}{c_1 s}\right)^{c_1 s} (hr)^{2c_1 s} + \sum_{1 \le s \le \ln\bar{n}} \bar{n}^s \bar{m}^{c_1 s} \left(\frac{hs}{\bar{n}}\right)^{2c_1 s} \\
&= \sum_{\ln\bar{n} \le s \le \gamma\bar{n}} \left(h^{2c_1} e^{1+c_1} r^{c_1-1} \left(\frac{\bar{\mu}}{c_1}\right)^{c_1}\right)^s + \sum_{1 \le s \le \ln\bar{n}} \left(\frac{(\bar{\mu} h^2 s^2)^{c_1}}{\bar{n}^{c_1-1}}\right)^s \\
&\le \sum_{\ln\bar{n} \le s \le \gamma\bar{n}} \left(\bar{C} r^{c_1-1} \bar{\mu}^{c_1}\right)^s + \ln\bar{n} \cdot \frac{(\bar{\mu} h^2 \ln^2\bar{n})^{c_1}}{\bar{n}^{c_1-1}} \\
&= \sum_{\ln\bar{n} \le s \le \gamma\bar{n}} \left(\bar{C} r^{c_1-1} \bar{\mu}^{c_1}\right)^s + o(1),
\end{aligned}$$

for some constant $0 < \bar{C} = \bar{C}(c_1) \le (h^2/c_1)^{c_1}\, e^{c_1+1}$. Choose

$$\gamma < \left(\frac{c_1}{h^2 e\bar{\mu}}\right)^{\frac{c_1}{c_1-1}} e^{-\frac{1}{c_1-1}}.$$

Then $\bar{C}\gamma^{c_1-1}\bar{\mu}^{c_1} < 1$. So there exist $0 < \beta < 1$, such that $\bar{C}\gamma^{c_1-1}\bar{\mu}^{c_1} < \beta$, for all $r \le \gamma$. When $c_1 \ge 2$ and $c_1/h^2e^2\bar{\mu} < 1$,

$$\left(\frac{c_1}{h^2 e\bar{\mu}}\right)^{\frac{c_1}{c_1-1}} e^{-\frac{1}{c_1-1}} > \left(\frac{c_1}{h^2 e^2\bar{\mu}}\right)^{c_1/(c_1-1)} > \left(\frac{c_1}{h^2 e^2\bar{\mu}}\right)^2.$$

Hence we may simply choose $\gamma = (c_1/h^2e^2\bar{\mu})^2$. Then

$$\sum_{\ln\bar{n} \le s \le \gamma\bar{n}} \left(\bar{C} r^{c_1-1} \bar{\mu}^{c_1}\right)^s < \sum_{\ln\bar{n} \le s \le \gamma\bar{n}} \beta^s = O(\beta^{\ln\bar{n}}) = o(1).$$

Hence we have $\mathbf{E}(Y) = o(1)$. ∎

The following corollary shows that the same property is shared by $\widehat{H}$.

Corollary 4.4.2 *Let c_1 be a constant that can be dependent on k with the constraint that $2 \le c_1 < h^2e^2\bar{\mu}$. Let $0 < \gamma = \varphi(k, c_1) = (c_1/h^2e^2\bar{\mu})^2$. Then a.a.s. for all $S \subset V(\widehat{H})$ with $|S| < \gamma n$, the number of hyperedges partially contained in S is less than $c_1|S|$.*

Proof Let n be the number of vertices in $\widehat{H}$ and D the sum of degrees of vertices in $\widehat{H}$. For any hyperedge $x \in \widehat{H}$, let x^+ denote its corresponding hyperedge in H. Obviously $n \le \bar{n}$. Combining with Lemma 4.4.1 and the fact that for any $S \subset V(\widehat{H})$, a hyperedge x is partially contained in S only if x^+ is partially contained S in H, Corollary 4.4.2 follows. ∎

The following corollary shows that $\widehat{H}$ a.a.s. has property $\mathcal{A}(\gamma)$, defined in Definition 4.2.4, for some certain $0 < \gamma < 1$.

Corollary 4.4.3 *Let $\gamma = e^{-4}h^{-6}/4$. Then a.a.s. $\widehat{H}$ has property $\mathcal{A}(\gamma)$ provided $k \ge 4w$.*

Proof The average degree $\bar{\mu}$ of H satisfies that $\bar{\mu} \le hk/w$ by (4.4.4). Apply Corollary 4.4.2 with $c_1 = k/2w$. Clearly $c_1 < ch^2e^2k$, and $c_1 \ge 2$ provided $k \ge 4w$. Then $\gamma \le \phi(k, c_1)$ by (4.4.4). By Definition 4.2.4, $\widehat{H}$ a.a.s. has property $\mathcal{A}(\gamma)$. ∎

Proof of Corollary 4.2.7 Let $\widehat{H}$ be the $(k+1)$-core of the random hypergraph $H \in \mathcal{G}_{\bar{n},\bar{m},h}$. Let $\epsilon > 0$ be any constant. By Theorem 4.2.3, there exists a constant $\delta > 0$, such that a.a.s. if $\bar{m} \le f(\bar{m}) - \epsilon\bar{n}$, then $\sum_{j=0}^{w-1}(w-j)m_{h-j} \le kn - \delta n$. By Theorem 4.2.5 and Corollary 4.4.3, there exists a constant N depending only on h and w such that provided $k > N$, $\widehat{H}$ a.a.s. has a (w, k)-orientation. On the other hand, if $\bar{m} \ge f(\bar{m}) + \epsilon\bar{n}$, then a.a.s. $\sum_{j=0}^{w-1}(w-j)m_{h-j} \ge kn + \delta n$, and hence clearly $\widehat{H}$ is not (w, k)-orientable. Therefore $f(\bar{n})$ is a sharp threshold function for the (w, k)-orientation of random hypergraphs. ∎

Note that in any multihypergraph $G \in \mathcal{M}(n, \mathbf{m}, k+1)$, the size of hyperedges varies between $h - w + 1$ to h. In the rest of the chapter, we will use the following notations. Let $E_{h-j} := \{x \in E(G) : |x| = h - j\}$. For any given $S \subset [n]$, let $m_{h-j,i}(S) := |\{x \in E_{h-j} : |x \cap S| = i\}|$ for any $0 \le i \le h - j$. When the context is clear of which set S is referred to, we may drop S from the notation. Let $\overline{S}$ denote the set $[n] \setminus S$ and let $d(S)$ denote the sum of degrees of vertices in S.

Recall the definition of induced subgraph with parameter w above the statement of Corollary 4.2.7 in Section 4.2. The proof of the following lemma uses network flow and the max-flow min-cut theorem. The techniques of converting numerous combinatorial problems into a network flow problem have been discussed in [20, 57]. The following Lemma generalises Hakimi's theorem [32, Theorem 4] for graphs.

Lemma 4.4.4 *Any multihypergraph $G \in \mathcal{M}(n, \mathbf{m}, 0)$ has a (w, k)-orientation if and only if*

$$d(H_S) - (h - w)e(H_S) \le k|S|, \quad \textit{for all } \; S \subset V(G).$$

Proof Formulate a network flow problem on a network G^* as follows. Let L be a set of vertices, each of which represents a hyperedge of G, and R be a set of n vertices, each of which represents a vertex in G. For any $u \in L$, and $v \in R$, uv is an edge in G^* if and only if $v \in u$ in G. Add vertices a and b to G^*, such that a is linked to every vertex in L, and b is linked to every vertex in R. Let $c : E(G^*) \to \mathbf{N}^+$ be defined as $c(au) = w - j$ for every $u \in L$ such that the degree of u is $h - j$, $c(vb) = k$ for every $v \in R$, and $c(uv) = 1$ for every $uv \in E(G^*)$. Then G has a (w, k)-orientation if and only if G^* has a flow of size $\sum_{j=0}^{w-1}(w - j)m_{h-j}$ from a to b. By the max-flow min-cut Theorem, G^* has a flow with all edges incident with a saturated if and only if

$$c(\delta(C)) \geq \sum_{j=0}^{w-1}(w - j)m_{h-j}, \quad \text{for all (a,b)-cuts } C. \tag{4.4.5}$$

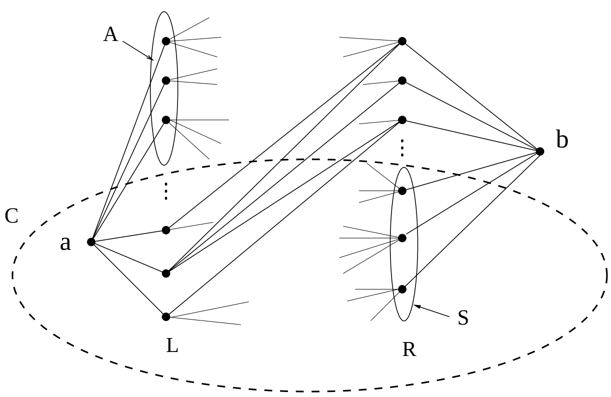

Figure 4.1: *A cut C in the graph G^**

As an example in Figure 4.1, $A \subset L$ is a set of hyperedges in G, and $S \subset R$ is a set of vertices in G. Let $C = \{a\} \cup \overline{A} \cup S$ define a cut of G^*. Then the condition in (4.4.5) is equivalent to

$$\forall C,\ \ c(\delta(C)) = k|S| + \sum_{j=0}^{w-1}\left(\sum_{x \in A \cap E_{h-j}}(w - j) + \sum_{x \in E_{h-j} \setminus A}|x \cap \overline{S}|\right) \geq \sum_{j=0}^{w-1}(w - j)m_{h-j}, \tag{4.4.6}$$

Let $A^* := \{x \in E_{h-j} : |x \cap S| \leq h - w\}$. Clearly A^* minimizes $c(\delta(C))$ for a given S. Therefore we only need to check (4.4.6) when $A = A^*$. The condition in (4.4.6) is then equivalent to

$$\sum_{j=0}^{w-1}\left(\sum_{x\in E_{h-j}\setminus A}(w-j)-\sum_{x\in E_{h-j}\setminus A}|x\cap\overline{S}|\right)\le k|S|.$$

Since

$$\sum_{j=0}^{w-1}\left(\sum_{x\in E_{h-j}\setminus A}(w-j)-\sum_{x\in E_{h-j}\setminus A}|x\cap\overline{S}|\right)=\sum_{j=0}^{w-1}\sum_{x\in E_{h-j}\setminus A}(w-j)-(h-j-|x\cap S|)$$
$$=\sum_{x\notin A}|x\cap S|-\sum_{x\notin A}(h-w)=d(H_S)-(h-w)e(H_s), \tag{4.4.7}$$

Lemma 4.4.4 follows. ∎

The next corollary follows immediately.

Corollary 4.4.5 *Any hypergraph H in $\mathcal{G}_{\bar{n},\bar{m},h}$ has a (w,k)-orientation if and only if for any $S\subset V(H)$,*

$$d(H_S)-(h-w)e(H_S)\le k|S|.$$

Proof of Corollary 4.2.7. This follows directly from Corollary 4.2.7 and Corollary 4.4.5. ∎

For any vertex set S, define

$$\partial^*(S)=d(S)-\sum_{j=0}^{w-1}\sum_{i=w-j+1}^{h-j}(i-(w-j))m_{h-j,i}, \tag{4.4.8}$$

For each hyperedge x of size $h-j$ which intersects S with i vertices, its contribution to $\partial^*(S)$ is $w-j\ge 0$ if $i\ge w-j+1$ and $i\ge 0$ otherwise. Therefore $\partial^*(S)\ge 0$ for any S. Recall that in a graph, the expansion rate of a vertex set is defined as $\partial(S)/|S|$, where $\partial(S)$ is the number of edges with exactly one end in S. It can be easily seen that $\partial^*(S)$ is closely related to $\partial(S)$. For instance, given the sum of degrees of vertices in S, both $\partial^*(S)$ and $\partial(S)$ are maximized when S is an independent set and minimized when S is a component of the graph. The following lemma shows that the (w,k)-orientation of G exists if and only if G satisfies some expansion property.

Lemma 4.4.6 *The following two graph properties of a multihypergraph $G\in\mathcal{M}(n,\mathbf{m},k+1)$ are equivalent.*

(*i*) $d(H_S)-(h-w)e(H_S)\le k|S|,$ *for all* $S\subset V(G)$;

(*ii*) $\partial^*(S)\ge k|S|+\left(\sum_{j=0}^{w-1}(w-j)m_{h-j}\right)-kn,$ *for all* $S\subset V(G)$.

Proof We show that for any $S \subset V(G)$, $d(H_S) - (h-w)e(H_S) \le k|S|$ if and only if $\partial^*(\overline{S}) \ge k|\overline{S}| + \left(\sum_{j=0}^{w-1}(w-j)m_{h-j}\right) - kn$. Then Lemma 4.4.6 follows immediately. Note from the definition of A^*, we have for any $x \in E_{h-j} \setminus A^*$, $|x \cap S| \ge h-w+1$ and hence $|x \cap \overline{S}| \le (h-j)-(h-w+1) = w-j-1$. By (4.4.7), for any $S \subset V(G)$,

$$
\begin{aligned}
& d(H_S) - (h-w)e(H_S) \le k|S| \\
& \Longleftrightarrow \sum_{j=0}^{w-1}\left(\sum_{x \in E_{h-j}\setminus A^*}(w-j) - \sum_{x \in E_{h-j}\setminus A^*}|x \cap \overline{S}|\right) \le kn - k|\overline{S}| \\
& \Longleftrightarrow \sum_{j=0}^{w-1}\sum_{i=0}^{w-j-1}(w-j-i)m_{h-j,i}(\overline{S}) \le kn - k|\overline{S}| \\
& \Longleftrightarrow \sum_{j=0}^{w-1}(w-j)m_{h-j} - \sum_{j=0}^{w-1}\left(\sum_{i=w-j}^{h-j}(w-j)m_{h-j,i}(\overline{S}) + \sum_{i=0}^{w-j-1} i m_{h-j,i}(\overline{S})\right) \le kn - k|\overline{S}| \\
& \Longleftrightarrow \partial^*(\overline{S}) \ge k|\overline{S}| + \left(\sum_{j=0}^{w-1}(w-j)m_{h-j}\right) - kn. \quad \blacksquare
\end{aligned}
$$

It follows from Lemma 4.4.4 and Lemma 4.4.6 that any multihypergraph $G \in \mathcal{M}(n, \mathbf{m}, k+1)$ is (w,k)-orientable if and only if Lemma 4.4.6 (ii) holds.

Without loss of generality, we assume $\sum_{j=0}^{w-1}(w-j)m_{h-j} - kn \le 0$. Otherwise, condition (4.4.5) is violated by taking $C = \{a\} \cup L \cup R$. The following lemma shows that, instead of checking conditions in Lemma 4.4.6 (ii), we can check that certain other events do not occur.

For any $S \subset V(G)$, let

$$
q_{h-j}(S) = \sum_{i=1}^{h-j} i m_{h-j,i}, \quad \eta(S) = \sum_{j=0}^{w-1}\sum_{i=1}^{h-j-1} m_{h-j,i}. \tag{4.4.9}
$$

In other words, $q_{h-j}(S)$ denotes the contribution to $d(S)$ from hyperedges of size $h-j$ and $\eta(S)$ denotes the number of hyperedges which intersect both S and $\overline{S}$. When the context is clear, we may use q_{h-j} and η instead to simplify the notation.

Recall that given a vertex set S, a hyperedge x is partially contained in S if $|x \cap S| \ge 2$. Let $\rho(S)$ denote the number of hyperedges partially contained in S and let $\nu(S)$ denote the number of hyperedges intersecting S.

Lemma 4.4.7 *Suppose that for some $S \subset V(G)$,*

$$\partial^*(S) < k|S| + \left(\sum_{j=0}^{w-1}(w-j)m_{h-j}\right) - kn. \qquad (4.4.10)$$

Then all of the following hold:

(i) $\rho(\overline{S}) > k|\overline{S}|/w$;

(ii) $\nu(S) < k|S|$;

(iii) $(h-w)\rho(S) > d(S) - k|S|$;

(iv) If, in addition, $\displaystyle\sum_{j=0}^{w-1}\frac{w-j}{h-j}q_{h-j}(S) \geq (1-\delta)k|S|$ *for some* $\delta > 0$, *then* $\eta(S) < h^2\delta k|S|$.

Proof Let s and $\bar{s}$ denote $|S|$ and $|\overline{S}|$ respectively. If (4.4.10) is satisfied, then

$$d(S) - \sum_{j=0}^{w-1}\sum_{i=w-j+1}^{h-j}(i-(w-j))m_{h-j,i} < ks - \left(kn - \sum_{j=0}^{w-1}(w-j)m_{h-j}\right) = \sum_{j=0}^{w-1}(w-j)m_{h-j} - k\bar{s}.$$

Hence

$$\begin{aligned}
k\bar{s} &< \sum_{j=0}^{w-1}(w-j)m_{h-j} - \sum_{j=0}^{w-1}\sum_{i=1}^{h-j} i m_{h-j,i} + \sum_{j=0}^{w-1}\sum_{i=w-j+1}^{h-j}(i-(w-j))m_{h-j,i} \\
&= \sum_{j=0}^{w-1}(w-j)m_{h-j,0} + \sum_{j=0}^{w-1}\sum_{i=1}^{w-1-j}(w-j-i)m_{h-j,i} \\
&\leq w\sum_{j=0}^{w-1}\left(m_{h-j,0} + \sum_{i=1}^{w-1-j} m_{h-j,i}\right).
\end{aligned}$$

Since

$$m_{h-j,0} = |\{x \in E_{h-j} : |x \cap \overline{S}| = h-j\}|,$$

and

$$\sum_{i=1}^{w-1-j} m_{h-j,i} \leq w|\{x \in E_{h-j} : 2 \leq |x \cap \overline{S}| \leq h-j-1\}|,$$

(this is because $1 \leq i \leq w-1-j$ and so $h-j-i \leq h-j-1$ and $h-j-i \geq h-(w-1) \geq 2$), we have

$$k\bar{s} < w|\{x \in E(G) : |x \cap \overline{S}| \geq 2\}|.$$

This proves part (i). Again, if (4.4.10) is satisfied, then

$$\sum_{j=0}^{w-1}\sum_{i=w-j+1}^{h-j}(i-(w-j))m_{h-j,i} > d(S) - ks + \left(kn - \sum_{j=0}^{w-1}(w-j)m_{h-j}\right).$$

Since

$$\sum_{j=0}^{w-1}\sum_{i=w-j+1}^{h-j}(i-(w-j))m_{h-j,i} \le \sum_{i=2}^{h}(i-1)|\{x : |x \cap S| = i\}| = d(S) - \nu(S),$$

we have

$$d(S) - \nu(S) > d(S) - ks + \left(kn - \sum_{j=0}^{w-1}(w-j)m_{h-j}\right).$$

Since $kn - \sum_{j=0}^{w-1}(w-j)m_{h-j} > 0$, this directly leads to part (ii). Since

$$\sum_{j=0}^{w-1}\sum_{i=w-j+1}^{h-j}(i-(w-j))m_{h-j,i} \le (h-w)|\{x : |x \cap S| \ge 2\}|,$$

we have

$$|\{x : |x \cap S| \ge 2\}| > d(S) - ks + \left(kn - \sum_{j=0}^{w-1}(w-j)m_{h-j}\right).$$

Since $kn - \sum_{j=0}^{w-1}(w-j)m_{h-j} > 0$, this proves part (iii). Now we prove part (iv). Let $t_{h-j} = 1 - (h-j)m_{h-j,h-j}/q_{h-j}$. Note that $d(S) = \sum_{j=0}^{w-1} q_{h-j}$ and $q_{h-j} = \sum_{i=1}^{h-j} i m_{h-j,i}$. For each hyperedge x of size $h-j$ which intersects S with i vertices, its contribution to q_{h-j} (and thus to $\partial^*(S)$) is

- $i \cdot (w-j)/(h-j)$, if $i = h-j$;
- $i \cdot (w-j)/i \ge i \cdot (w-j)/(h-j-1)$, if $w-j+1 \le i \le h-j-1$;
- $i \ge i \cdot (w-j)/(h-j-1)$, if $1 \le i \le w-j$;

Then

$$\begin{aligned}
\partial^*(S) &\ge \sum_{j=0}^{w-1}\left(\frac{w-j}{h-j}q_{h-j}(1-t_{h-j}) + \frac{w-j}{h-j-1}q_{h-j}t_{h-j}\right) \\
&= \sum_{j=0}^{w-1}\frac{w-j}{h-j}q_{h-j} + \sum_{j=0}^{w-1}\frac{w-j}{(h-j)(h-j-1)}q_{h-j}t_{h-j} \\
&\ge \sum_{j=0}^{w-1}\frac{w-j}{h-j}q_{h-j} + \frac{1}{h^2}\sum_{j=0}^{w-1}(w-j)q_{h-j}t_{h-j}.
\end{aligned}$$

If $\sum_{j=0}^{w-1} \frac{w-j}{h-j} q_{h-j} \geq (1-\delta)ks$ for some $\delta > 0$, then (4.4.10) implies that

$$\frac{1}{h^2} \sum_{j=0}^{w-1} (w-j) q_{h-j} t_{h-j} < \delta ks.$$

Therefore $\eta(S) \leq \sum_{j=0}^{w-1} q_{h-j} t_{h-j} < h^2 \delta ks$. This proves part (iv). ∎

4.5 Proof of Theorem 4.2.5

Recall that $\epsilon > 0$ and $k \geq 2$ are fixed. Let $\mathbf{m}$ be an integer vector with the constraint (4.4.2). Given $\mathbf{m}$, let D, μ be as defined in (4.4.1). Let G be a random multihypergraph from the probability space $\mathcal{M}(n, \mathbf{m}, k+1)$.

Since h and w are given, we consider them as absolute constants. Therefore, whenever we refer to $g = O(f)$, it means that there exists a constant C such that $g \leq Cf$, where C can depend on h and w. We also use notation $g = O_\gamma(f)$, which means that there exists a constant C depending on γ only such that $g \leq Cf$. The same convention applies to $o(f)$, $\Omega(f)$, $\Theta(f)$ and $o_\gamma(f)$, $\Omega_\gamma(f)$ and $\Theta_\gamma(f)$.

In this section we prove Theorem 4.2.5. To start, we explain our use of the probability spaces in this section. By Lemma 4.3.2, $\mathcal{H}(n, \mathbf{m}, k+1)$ is the same as $\mathcal{M}(n, \mathbf{m}, k+1)$ restricted to simple multihypergraphs. The next lemma shows that any event that holds a.a.s. in $\mathcal{M}(n, \mathbf{m}, k+1)$ holds a.a.s. in $\mathcal{H}(n, \mathbf{m}, k+1)$. In view of this, we only need to prove the orientability in $\mathcal{M}(n, \mathbf{m}, k+1)$. Since $\mathcal{M}(n, \mathbf{m}, k+1)$ is equivalent to $\mathcal{P}(n, \mathbf{m}, \mathbf{0}, k+1)$, defined in Section 4.3, the partition-allocation model will often be considered in this section.

Lemma 4.5.1 *For any sequence of events A_n, if $\mathbf{P}_{\mathcal{M}(n,\mathbf{m},k+1)}(A_n) = o(1)$, then $\mathbf{P}_{\mathcal{H}(n,\mathbf{m},k+1)}(A_n) = o(1)$.*

Proof The degree sequence of multihypergraphs resulting from $\mathcal{M}(n, \mathbf{m}, k+1)$ is asymptotically that of truncated Poisson (see [16, Lemma 1] for a short proof). Since $D = \sum_{j=0}^{w-1} (h-j) m_{h-j} = \Theta(k)n$, we have a.a.s. $\sum_{i=1}^{n} d_i(d_i - 1) = \Theta(k^2)n$. The proof of [43, Theorem 5.2] now trivially adapts to show that the probability that a multihypergraph in $\mathcal{M}(n, \mathbf{m}, k+1)$ is simple is $\Theta(1)$. Thus, for any event A_n, if $\mathbf{P}_{\mathcal{M}(n,\mathbf{m},k+1)}(A_n) = o(1)$, then by Lemma 4.3.2, $\mathbf{P}_{\mathcal{H}(n,\mathbf{m},k+1)}(A_n) \leq \mathbf{P}_{\mathcal{M}(n,\mathbf{m},k+1)}(A_n)/\Theta(1) = o(1)$. ∎

We next sketch the proof of Theorem 4.2.5. Let $q_{h-j}(S)$ and $\eta(S)$ be defined as in (4.4.9). Note that the partition-allocation model gives a good foundation for proving that certain events hold concerning the distribution of vertex degrees and intersections of hyperedge sets with vertex sets. Using this and various probabilistic tools, we show that

(a) every hypergraph G with property $\mathcal{A}(\gamma)$ contains no sets S with $|S| > (1-\gamma)n$ for which Lemma 4.4.7 (i) holds;

(b) the probability that G has property $\mathcal{A}(\gamma)$ and contains some set S with $|S| < \gamma n$ for which both Lemma 4.4.7 (ii) and (iii) holds is $o(1)$;

(c) there exists $\delta > 0$, such that when k is large enough, a.a.s. $\sum_{j=0}^{w-1} \frac{w-j}{h-j} q_{h-j} \geq (1-\delta)k|S|$, and the probability of G containing some set S with $\gamma n \leq |S| \leq (1-\gamma)n$ and $\eta(S) < h^2\delta k|S|$ is $o(1)$.

It follows that the probability that G has property $\mathcal{A}(\gamma)$ and contains some set S for which all parts (i)-(iv) of Lemma 4.4.7 hold is $o(1)$. Then by Lemma 4.4.6,

$$\mathbf{P}(G \in \mathcal{A}(\gamma) \wedge G \text{ is not } (w,k)\text{-orientable}) = o(1).$$

We start with a few concentration properties. As discussed in Section 4.3, the degree sequence of $G \in \mathcal{M}(n, \mathbf{m}, k+1)$ obeys the multinomial distribution. The following lemma bounds the probability of rare degree (sub)sequences where the degree distribution is independent truncated Poisson. We will use this result to bound the probability of rare degree sequences in $\mathcal{M}(n, \mathbf{m}, k+1)$.

Lemma 4.5.2 *Let $s \geq w(n)$ for some $w(n) \to \infty$ as $n \to \infty$ and let $Y_1, \ldots, Y_s$ be independent copies of Z defined in (4.3.2) with λ satisfying $\lambda f_k(\lambda) = \mu f_{k+1}(\lambda)$. Let $0 < \delta < 1$ be any constant. Then there exist $N > 0$ and $0 < \alpha < 1$ both depending only on δ, such that provided $k > N$,*

$$\mathbf{P}\left(\Big|\sum_{i=1}^{s} Y_i - \mu s\Big| \geq \delta \mu s\right) \leq \alpha^{\mu s}.$$

Proof Let $G(x)$ be the probability generating function of Y_i. Then

$$G(x) = \sum_{j \geq k+1} \mathbf{P}(Z = j)x^j = \frac{e^{-\lambda}}{f_{k+1}(\lambda)} \left(e^{\lambda x} - \sum_{j=0}^{k} \frac{(\lambda x)^j}{j!}\right) \leq \frac{e^{\lambda x - \lambda}}{f_{k+1}(\lambda)},$$

for all $x \geq 0$. For any nonnegative integer ℓ,

$$\mathbf{P}\left(\sum_{i=0}^{s} Y_i = l\right) \leq \frac{G(x)^s}{x^\ell}, \quad \forall x \geq 0.$$

Putting $x = \ell/s\lambda$ gives

$$\mathbf{P}\left(\sum_{i=0}^{s} Y_i = \ell\right) \leq \frac{e^{\ell - \lambda s}}{(\ell/(\lambda s))^\ell f_{k+1}(\lambda)^s} = \left(\frac{es\lambda}{\ell}\right)^\ell \left(\frac{e^{-\lambda}}{f_{k+1}(\lambda)}\right)^s. \tag{4.5.1}$$

It is easy to check that the right hand side of (4.5.1) is an increasing function of l when $l \leq \lambda s$ and decreasing function of l when $l \geq \lambda s$. By Proposition 4.3.3, there exists a constant N_0 depending only on δ such that provided $k > N_0$, $(1-\delta)\mu < \lambda$. Thus, for any $\ell \leq (1-\delta)\mu s$,

$$\mathbf{P}\left(\sum_{i=0}^{s} Y_i = \ell\right) \leq \left(\frac{es\lambda}{(1-\delta)\mu s}\right)^{(1-\delta)\mu s} \left(\frac{e^{-\lambda}}{f_{k+1}(\lambda)}\right)^s,$$

and so

$$\mathbf{P}\left(\sum_{i=1}^{s} Y_i \leq (1-\delta)\mu s\right) \leq \mu s \left(\frac{e\lambda}{(1-\delta)\mu}\right)^{(1-\delta)\mu s} \left(\frac{e^{-\lambda}}{f_{k+1}(\lambda)}\right)^s.$$

The expectation of Y_1 is $\lambda f_k(\lambda)/f_{k+1}(\lambda) = \mu$. By Proposition 4.3.3, we have $\mu \geq \lambda$ and $\mu - \lambda \to 0$ as $k \to \infty$. Therefore

$$\begin{aligned}\mathbf{P}\left(\sum_{i=1}^{s} Y_i \leq (1-\delta)\mu s\right) &\leq \mu s \left(\frac{\exp(\mu - \lambda - \delta\mu)}{(1-\delta)^{(1-\delta)\mu} f_{k+1}(\lambda)}\right)^s \\ &= \mu s \left(\frac{\exp(\mu-\lambda)}{f_{k+1}(\lambda)} \cdot \left(\frac{\exp(-\delta)}{(1-\delta)^{(1-\delta)}}\right)^{\mu}\right)^s.\end{aligned}$$

Since $0 < \delta < 1$,

$$0 < \frac{\exp(-\delta)}{(1-\delta)^{(1-\delta)}} < 1.$$

Since

$$\exp(\mu - \lambda) \to 1, \quad f_{k+1}(\lambda) \to 1, \quad \text{as } k \to \infty$$

by Proposition 4.3.3, there exists $N_1 > 0$ and $0 < \alpha_1 < 1$, both depending only on δ, such that provided $k > N_1$,

$$\mathbf{P}\left(\sum_{i=1}^{s} Y_i \leq (1-\delta)\mu s\right) \leq \alpha_1^{\mu s}.$$

Now we bound the upper tail of $\sum_{i=1}^{s} Y_i$. Let $j = 1, 2, \ldots$. For any ℓ satisfying $(1+j)\mu s \leq \ell < (2+j)\mu s$, as with the lower tail bound,

$$\mathbf{P}\left(\sum_{i=0}^{s} Y_i = \ell\right) \leq \left(\frac{es\lambda}{(1+j)\mu s}\right)^{(1+j)\mu s} \left(\frac{e^{-\lambda}}{f_{k+1}(\lambda)}\right)^s = \left(\frac{\exp(\mu-\lambda)}{f_{k+1}(\lambda)} \cdot \left(\frac{e^j}{(1+j)^{(1+j)}}\right)^{\mu}\right)^s,$$

and so

$$\mathbf{P}\left((1+j)\mu s \leq \sum_{i=1}^{s} Y_i < (2+j)\mu s\right) \leq \mu s \left(\frac{\exp(\mu-\lambda)}{f_{k+1}(\lambda)} \cdot \left(\frac{e^j}{(1+j)^{(1+j)}}\right)^{\mu}\right)^s. \quad (4.5.2)$$

Similarly we have

$$\mathbf{P}\left((1+\delta)\mu s \leq \sum_{i=1}^{s} Y_i < 2\mu s\right) \leq \mu s \left(\frac{\exp(\mu-\lambda)}{f_{k+1}(\lambda)} \cdot \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}\right)^{s}. \tag{4.5.3}$$

Since $0 < e^{\delta}/(1+\delta)^{(1+\delta)} < 1$ for any $\delta > 0$, we may bound the right side of (4.5.2) and (4.5.3) by $\alpha_2^{\mu s}$ where $0 < \alpha_2 < 1$ is some constant depending only on δ. Also, since $e/(1+j) < 1$ for all $j \geq 2$, the right side of (4.5.2) is at most $\exp(-\Theta(\mu)js)$ for $j \geq 2$ provided k is large enough. Hence there exists $N_2 > 0$ and $0 < \alpha_3 < 1$ depending only on δ, such that provided $k > N_2$,

$$\mathbf{P}\left(\sum_{i=1}^{s} Y_i \geq (1+\delta)\mu s\right) \leq \alpha_3^{\mu s}.$$

The lemma follows by choosing $\alpha = \max\{\alpha_1, \alpha_3\}$ and $N = \max\{N_1, N_2\}$. ∎

Lemma 4.5.3 *Let $k \geq -1$ be an integer. Drop D balls independently at random into n bins. Let $\mu = D/n$ and let λ be defined as $\lambda f_k(\lambda) = \mu f_{k+1}(\lambda)$. Assume $D - (k+1)n \to \infty$ as $n \to \infty$. Then the probability that each bin contains at least $k+1$ balls is $\Omega(f_{k+1}(\lambda)^n)$.*

Proof Let $\mathbf{d}$ denote $(d_1, \ldots, d_n)$. Let $\mathscr{D} = \{\mathbf{d} : d_i \geq k+1 \ \forall i \in [n], \ \sum_{i=1}^{n} d_i = D\}$. Let $\mathbf{P}(B)$ denote the probability that each bin contains at least $k+1$ balls. Then

$$\mathbf{P}(B) = \sum_{\mathbf{d} \in \mathscr{D}} \binom{D}{d_1, \ldots, d_n} \Big/ n^D = \frac{D!}{n^D} \sum_{\mathbf{d} \in \mathscr{D}} \prod_{i=1}^{n} \frac{1}{d_i!}.$$

Let $Y_1, \ldots, Y_n$ be n independent truncated Poisson variables which are copies of $Z_{(\geq k+1)}$ as defined in (4.3.2) with parameter λ satisfying $\lambda f_k(\lambda) = \mu f_{k+1}(\lambda)$. Then

$$\mathbf{P}\left(\sum_{i=1}^{n} Y_i = D\right) = \sum_{\mathbf{d} \in \mathscr{D}} \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{d_i}}{f_{k+1}(\lambda) d_i!} = \frac{e^{-\lambda n}\lambda^D}{f_{k+1}(\lambda)^n} \sum_{\mathbf{d} \in \mathscr{D}} \prod_{i=1}^{n} \frac{1}{d_i!}.$$

Since $D - (k+1)n \to \infty$ as $n \to \infty$, $\mathbf{P}\left(\sum_{i=1}^{n} Y_i = D\right) = \Omega(D^{-1/2})$ (see [47, Theorem 4(a)] for a short proof),

$$\sum_{\mathbf{d} \in \mathscr{D}} \prod_{i=1}^{n} \frac{1}{d_i!} = \Omega\left(\frac{e^{\lambda n} f_{k+1}(\lambda)^n}{\lambda^D D^{1/2}}\right).$$

So, using Stirling's formula,

$$\begin{aligned} \mathbf{P}(B) &= \Omega\left(\frac{D!}{n^D} \cdot \frac{e^{\lambda n} f_{k+1}(\lambda)^n}{\lambda^D D^{1/2}}\right) = \Omega\left(\sqrt{D}\left(\frac{D}{en}\right)^{D} \cdot \frac{e^{\lambda n} f_{k+1}(\lambda)^n}{\lambda^D D^{1/2}}\right) \qquad (4.5.4) \\ &= \Omega\left(\left(\frac{\mu}{\lambda} e^{\lambda/\mu - 1}\right)^{\mu n} f_{k+1}(\lambda)^n\right). \end{aligned}$$

Since $(\mu/\lambda) \cdot e^{\lambda/\mu - 1} \geq 1$, $\mathbf{P}(B) = \Omega(f_{k+1}(\lambda)^n)$. ∎

Corollary 4.5.4 *Let $k \geq -1$ be an integer. Let $\mathscr{D} = \{\mathbf{d} : d_i \geq k+1, \forall i \in [n], \sum_{i=1}^n d_i = D\}$ and let A_n be any subset of $\mathscr{D}$. Let $\mu = D/n$. Let $\mathbf{P}(A_n)$ denote the probability that the degree sequence $\mathbf{d}$ of $G \in \mathcal{M}(n, \mathbf{m}, k+1)$ is in A_n and let $\mathbf{P}_{TP}(A_n)$ be the probability that $(Y_1, \ldots, Y_n) \in A_n$ where Y_i are independent copies of the random variable $Z_{(\geq k+1)}$ as defined in (4.3.2) with the parameter λ satisfying $\lambda f_k(\lambda) = \mu f_{k+1}(\lambda)$. Assume $D - (k+1)n \to \infty$ as $n \to \infty$. Then*

$$\mathbf{P}(A_n) = O\left(\sqrt{D}\right) \mathbf{P}_{TP}(A_n).$$

Proof Let A_n be any subset of $\mathscr{D}$ and let $\mathbf{P}(B)$ denote the probability that each bin contains at least $k+1$ balls by dropping D balls independently and randomly into n bins. Consider the partition-allocation model that generates $\mathcal{M}(n, \mathbf{m}, k+1)$, which allocates the partitioned D balls randomly into n bins with the restriction that each bin contains at least $k+1$ balls. Then

$$\mathbf{P}(A_n) = \sum_{\mathbf{d} \in A_n} \frac{1}{\mathbf{P}(B)} \cdot \binom{D}{d_1, \ldots, d_n} \Big/ n^D = \frac{D!}{n^D \mathbf{P}(B)} \sum_{\mathbf{d} \in A_n} \prod_{i=1}^n \frac{1}{d_i!},$$

and

$$\mathbf{P}_{TP}(A_n) = \sum_{\mathbf{d} \in A_n} \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{d_i}}{f_{k+1}(\lambda) d_i!} = \frac{e^{-\lambda n} \lambda^D}{f_{k+1}(\lambda)^n} \sum_{\mathbf{d} \in A_n} \prod_{i=1}^n \frac{1}{d_i!}.$$

Therefore

$$\mathbf{P}(A_n) = \frac{D! e^{\lambda n} f_{k+1}^n}{n^D \mathbf{P}(B) \lambda^D} \mathbf{P}_{TP}(A_n) = O\left(\sqrt{D}\right) \mathbf{P}_{TP}(A_n),$$

since $\mathbf{P}(B) = \Omega\left(\left(\frac{\mu}{\lambda} e^{\lambda/\mu - 1}\right)^{\mu n} f_{k+1}(\lambda)^n\right)$ by Lemma 4.5.3 (4.5.4). ∎

A significant difficulty in this work is to ensure that various constants do not depend on the choice of ϵ. In particular, we emphasize that the constants such as α and N in the following results do not depend on ϵ.

The next is a corollary of Lemma 4.5.2 and Corollary 4.5.4.

Corollary 4.5.5 *Let $0 < \delta < 1$ be any constant. Then there exist two constants $N > 0$ and $0 < \alpha < 1$ both depending only on δ, such that provided $k > N$, for any vertex set $S \subset V(G)$ with $|S| \geq \log^2 n$,*

$$\mathbf{P}(|d(S) - \mu|S|| \geq \delta \mu |S|) \leq \alpha^{\mu |S|}.$$

Proof Let $Y_1, \ldots, Y_n$ be independent copies of the truncated Poisson random variable Z as defined in (4.3.2). Let $S \subset V(G)$ and let $s = |S|$. Then by Lemma 4.5.2, there exist $N > 0$ and $0 < \hat{\alpha} < 1$, both depending only on δ, such that provided $k > N$,

$$\mathbf{P}\left(\Big|\sum_{i\in S} Y_i - \mu s\Big| \geq \delta\mu s\right) \leq \hat{\alpha}^{\mu s},$$

By Corollary 4.5.4,

$$\mathbf{P}(|d(S) - \mu s| \geq \delta\mu s) \;\; \leq \;\; O(D^{1/2})\hat{\alpha}^{\mu s} = \left(\exp\left(\frac{\ln\Theta(\sqrt{\mu n})}{\mu s}\right)\hat{\alpha}\right)^{\mu s}.$$

Since $s \geq \log^2 n$ and so

$$\frac{\ln\Theta(\sqrt{\mu n})}{\mu s} \to 0, \quad \text{as } n \to \infty.$$

Let $\alpha = 1/2 + \hat{\alpha}/2$. Then $0 < \hat{\alpha} < \alpha < 1$ and α depends only on δ. Then provided $k > N$, $\mathbf{P}(|d(S) - \mu s| \geq \delta\mu s) \leq \alpha^{\mu s}$. ∎

The following corollary shows that $d(S)$ is very concentrated when S is not too small.

Corollary 4.5.6 *Let $\delta > 0$ and $0 < \gamma < 1$ be arbitrary constants. Then there exists a constant N depending only on δ and γ, such that provided $k > N$,*

$$\mathbf{P}(\exists S \subset V(G), s \geq \gamma n, |d(S) - \mu s| \geq \delta\mu s) = o(1).$$

Proof For any $S \subset V(G)$, let $s = |S|$. By Corollary 4.5.5, there exists $N_1 > 0$ and $0 < \alpha < 1$, both depending only on δ, such that provided $k > N_1$, for any $S \subset V(G)$,

$$\mathbf{P}(|d(S) - \mu s| \geq \delta\mu s) \leq \alpha^{\mu s}.$$

Let N_2 be the smallest integer such that $e\alpha^{N_2}/\gamma < 1/2$. Let $N = \max\{N_1, N_2\}$. Then N depends only on δ and γ. For all $\mu > N$,

$$\mathbf{P}(\exists S \subset [n], s \geq \gamma n, |d(S) - \mu s| \geq \delta\mu s) \leq \sum_{\gamma n \leq s \leq n} \binom{n}{s}\alpha^{\mu s} \leq \sum_{\gamma n \leq s \leq n} \left(\frac{en}{s} \cdot \alpha^{\mu}\right)^s$$
$$\leq \sum_{\gamma n \leq s \leq n} \left(\frac{e}{\gamma} \cdot \alpha^{\mu}\right)^s = O\left(2^{-\gamma n}\right) = o(1). \;\; ∎$$

The following lemma will be used later to prove that a.a.s. $\sum_{j=0}^{w-1}(w-j)q_{h-j}/(h-j) \geq (1-\delta)ks$ provided k is large enough.

Lemma 4.5.7 *Let $0 < \delta < 1$ be any given constant. Let $\mathcal{C} = \{c_0, \ldots, c_{w-1}\}$ be a set of colours and let there be D balls coloured with some colour in $\mathcal{C}$. For any $0 \le j \le w-1$, let p_j denote the proportion of balls that are coloured c_j. Randomly choose q balls among all. Let q_j be the number of balls chosen that are coloured with c_j. Then for any $0 \le j \le w-1$,*

$$\mathbf{P}(|q_j - p_j q| \ge \delta p_j q) \le \exp(-\Theta(\delta^2 p_j q)).$$

Proof For any $0 \le j \le w-1$ any $\ell > 0$,

$$\mathbf{P}(q_j = \ell) = \binom{p_j D}{\ell}\binom{D - p_j D}{q - \ell} \Big/ \binom{D}{q}.$$

Let p_ℓ denote $\mathbf{P}(q_j = \ell)$. Put $\ell_0 = p_j q$, $\ell_1 = (1-\delta/2)p_j q$ and $\ell_2 = (1-\delta)p_j q$. Then for any $\ell \le \ell_1$,

$$\frac{p_{\ell-1}}{p_\ell} = \frac{\ell(D(1-p_j) - q + \ell)}{(p_j D - \ell + 1)(q - \ell + 1)} \le \frac{\ell_1(D(1-p_j) - q + \ell_0)}{(p_j D - \ell_0)(q - \ell_0)} = 1 - \frac{\delta}{2}.$$

Then

$$p_{\ell_2} \le (1-\delta/2)^{\delta p_j q/2} p_{\ell_1} \le \exp\left(\frac{\delta p_j q}{2} \ln\left(1 - \frac{\delta}{2}\right)\right) \le \exp(-\delta^2 p_j q/4).$$

So

$$\mathbf{P}(q_j \le (1-\delta)p_j q) = \sum_{\ell \le \ell_2} p_\ell \le \frac{1}{\delta} p_{\ell_2} \le \exp(-\Theta(\delta^2 p_j q)).$$

Similarly we can bound the upper tail and then Lemma 4.5.7 follows. ∎

Lemma 4.5.8 *Let $0 < \delta < 1$ and $0 < \gamma < 1$ be two arbitrary constants. Given $S \subset V(G)$, let $q_{h-j} = q_{h-j}(S)$ be as defined in (4.4.9). Then there exists $N > 0$ depending only on δ and γ such that for all $k > N$,*

$$\mathbf{P}\left(\exists S \subset V(G), |S| \ge \gamma n, \sum_{j=0}^{w-1} \frac{w-j}{h-j} q_{h-j} < (1-\delta)k|S|\right) = o(1).$$

Proof For any $0 \le j \le w-1$, let p_j denoted $(h-j)m_{h-j}/D$. Let $J := \{j : p_j > \delta/8w\}$. We first show that given $S \subset V(G)$ with $|S| \ge \gamma n$, if

$$\sum_{j=0}^{w-1} \frac{w-j}{h-j} q_{h-j} < (1-\delta)k|S|, \tag{4.5.5}$$

then there exists $j \in J$ such that $q_{h-j}(S) \le (1-\delta/8)p_j d(S)$. Assume there is no such j by contradiction. Then

$$\begin{aligned}
\sum_{j=0}^{w-1} \frac{w-j}{h-j} q_{h-j}(S) &\ge \sum_{j\in J} \frac{w-j}{h-j} q_{h-j}(S) > (1-\delta/8)d(S) \sum_{j\in J} \frac{w-j}{h-j} p_j \\
&= (1-\delta/8)d(S) \left(\sum_{j=0}^{w-1} \frac{w-j}{h-j} p_j - \sum_{j\notin J} \frac{w-j}{h-j} p_j \right) \ge (1-\delta/8)d(S) \left(\sum_{j=0}^{w-1} \frac{w-j}{h-j} p_j - \frac{w}{h}\frac{\delta}{8w} \right) \\
&\ge (1-\delta/8)d(S) \sum_{j=0}^{w-1} \frac{w-j}{h-j} p_j (1-\delta/8) \ge (1-\delta/4)d(S) \sum_{j=0}^{w-1} \frac{w-j}{h-j} p_j. \qquad (4.5.6)
\end{aligned}$$

Let $s = |S|$ and let $r = s/n$. Then by Corollary 4.5.6, there exists $N_2 > 0$ depending on δ and γ only, such that a.a.s. $d(S) \ge (1-\delta/4)Dr$ whenever $k > N_2$. Therefore, combining with (4.5.6), we get a.a.s. provided $k > \max\{N_1, N_2\}$,

$$\sum_{j=0}^{w-1} \frac{w-j}{h-j} q_{h-j}(S) > (1-\delta/2)Dr \sum_{j=0}^{w-1} \frac{w-j}{h-j} p_j \ge (1-\delta/2)r(kn - 2\epsilon n) = (1-\delta/2)(k-2\epsilon)s.$$

For any $k > 2/\delta \ge 4\epsilon/\delta$, we have $(1-\delta/2)(k-2\epsilon)s > (1-\delta)ks$. Take $N = \max\{N_1, N_2, 2/\delta\}$. Then for any $k > N$, we have a.a.s.

$$\sum_{j=0}^{w-1} \frac{w-j}{h-j} q_{h-j}(S) > (1-\delta)ks,$$

which contradicts (4.5.5). It follows that there exists $j \in J$ such that $q_{h-j}(S) \le (1-\delta/8)p_j d(S)$.

Consider the partition-allocation model that generates $\mathcal{M}(n, \mathbf{m}, \mathbf{0}, k+1)$. Let $\mathcal{C} = \{c_0, \ldots, c_{w-1}\}$ be a set of colours. For balls partitioned into parts that are of size $h-j$ for some $0 \le j \le w-1$, colour them with c_j. Then the w colours are distributed u.a.r. among the D balls. By Lemma 4.5.7, for any $S \subset V(G)$,

$$\mathbf{P}\big(q_{h-j}(S) \le (1-\delta/8)p_j d(S)\big) \le \exp\big(-\Theta(\delta^2 p_j d(S))\big).$$

Then there exists a constant N_1 depending only on δ and γ such that,

$$\begin{aligned}
&\mathbf{P}\big(\exists S, j \in J, s \ge \gamma n, q_{h-j}(S) \le (1-\delta/8)p_j d(S)\big) \\
&\qquad \le w2^n \exp\big(-\Theta(\delta^3 d(S))\big) \le w\big(2\exp(-\Theta(\delta^3\gamma k))\big)^n = o(1).
\end{aligned}$$

Note that the inequality holds because $|J| \le w$, the number of sets S with $|S| \ge \gamma n$ is at most 2^n, $\delta/8w \le p_j < 1$ for all $j \in J$ and $d(S) \ge (k+1)|S| > k\gamma n$. It follows that a.a.s. there exists no set S with $|S| \geq \gamma n$ for which there exists $j \in J$ such that $q_{h-j}(S) \leq (1 - \delta/8)p_j d(S)$. Lemma 4.5.8 then follows. ∎

Recall that $\rho(S)$ is the number of hyperedges partially contained in S and $\nu(S)$ is the number of hyperedges intersecting S by the definition above Lemma 4.4.7.

Lemma 4.5.9 *Let $\delta > 0$ be any constant and let $\mu = \mu(G) = D/n$ as defined in (4.4.1). Then there exists a constant $N > 0$ depending only on δ such that provided $k > N$, a.a.s. there exists no $S \subset V(G)$, for which* $\log^2 n \leq |S| \leq n$, *$d(S) < (1-\delta)\mu|S|$, and $\nu(S) < k|S|$.*

This lemma will be proved after the proof of Theorem 4.2.5.

Proof of Theorem 4.2.5. By Lemma 4.4.6 and 4.4.7, it is enough to show that the expected number of sets S contained in a hypergraph $G \in \mathcal{M}(n, \mathbf{m}, k+1)$ with property $\mathcal{A}(\gamma)$ for which all of Lemma 4.4.7 (i)–(iv) are satisfied is $o(1)$. We call a set $S \subset V(G)$ is interesting if it lies in a hypergraph G with property $\mathcal{A}(\gamma)$. Let X be the number of interesting sets $S \subset V(G)$ such that (4.4.10) holds. Similarly, let $X_{<a}$ (or $X_{>b}$ or $X_{[a,b]}$) for any $0 < a < b < n$ denote the number of interesting $S \subset [n]$ such that (4.4.10) holds and $|S| < a$ (or $|S| > b$ or $a \leq |S| \leq b$) respectively. For any set S under discussion, let s denote $|S|$ and $\bar{s}$ denote $|\overline{S}|$.

Case 1: $s < \epsilon n/k$. By theorem's hypothesis

$$\left(\sum_{j=0}^{w-1}(w-j)m_{h-j}\right) - kn < -\epsilon n,$$

any S satisfying (4.4.10) must satisfy

$$\partial^*(S) < ks - \epsilon n. \tag{4.5.7}$$

When $s < \epsilon n/k$, $ks - \epsilon n < 0$. However $\partial^*(S) \geq 0$ as observed below (4.4.8). Hence (4.5.7) cannot hold. Thus $X_{<\epsilon n/k} = 0$.

Case 2: $s > (1-\gamma)n$. part (i) of Lemma 4.4.7 says that (4.4.10) holds only if the number of hyperedges partially contained in $\overline{S}$ is at least $k\bar{s}/w$. But X counts only interesting sets, i.e. sets that lie in a hypergraph with property $\mathcal{A}(\gamma)$. By the definition of property $\mathcal{A}(\gamma)$, there are no such interesting sets and so $X_{\geq(1-\gamma)n} = 0$.

Case 3: $\epsilon n/k \leq s < \gamma n$. Let $\delta_1 = (h-w)/2h$. By Lemma 4.5.9, there exists $N_1 > 0$ such that provided $k > N_1$, the expected number of S with $d(S) < (1-\delta_1)\mu s$ for which Lemma 4.4.7 (ii) is satisfied and $\epsilon n/k \leq s \leq n$ is $o(1)$. We now show that there exists no interesting sets $S \subset V(G)$ with $|S| < \gamma n$ for which Lemma 4.4.7 (iii) holds and $d(S) \geq (1-\delta_1)\mu s$. If $d(S) \geq (1-\delta_1)\mu s$, $d(S) \geq \frac{h+w}{2w}ks$ provided $k \geq h+w$ since $\mu \geq h(k-1)/w$ by (4.4.3). Then it follows that

$$\frac{d(S)-ks}{h-w} \geq \frac{ks}{2w}.$$

Lemma 4.4.7 (iii) implies that (4.4.10) holds only if the number of hyperedges partially contained in S is at least $ks/2w$. By the definition of property $\mathcal{A}(\gamma)$, there is no such interesting sets S when $s < \gamma n$. So provided $k > \max\{N_1, h+w\}$, a.a.s. there exists no interesting sets S, with $s < \gamma n$ for which both Lemma 4.4.7 (ii) and (iii) hold. Then $\mathbf{E}(X_{<\gamma n}) = o(1)$.

Note that $k|S|/2w$ in the definition of property $\mathcal{A}(\gamma)$ can be modified to be $Ck|S|$ for any positive constant C, and it can be checked straightforwardly that there exists a constant γ depending on C only, such that Corollary 4.4.3 holds. Therefore, any $0 < \delta_1 < 1 - w/h$ would work here by choosing some appropriate C to modify the definition of property $\mathcal{A}(\gamma)$.

Case 4: $\gamma n \leq s \leq (1-\gamma)n$. Let $0 < \delta_2 < 1$ be chosen later. By Lemma 4.5.8, there exists $N_2 > 0$ depending only on δ_2 such that provided $k \geq N_2$, a.a.s.

$$\sum_{j=0}^{w-1} \frac{w-j}{h-j} q_{h-j} \geq (1-\delta_2)ks \quad \text{for all } S \text{ with } \gamma n \leq |S| \leq (1-\gamma)n.$$

For any $S \subset V(G)$, let $\eta = \eta(S)$ be as defined in (4.4.9). Then by Lemma 4.4.7, to show $\mathbf{E}(X_{[\gamma n,(1-\gamma)n]}) = o(1)$, it is enough to show that the expected number of sets S with $\gamma n \leq s \leq (1-\gamma)n$ for which $\eta(S)$ is at most $h^2\delta_2 ks$, is $o(1)$. Consider the probability space $\mathcal{M}(n, \mathbf{m}, \mathbf{0}, 0)$, which is generated by placing each hyperedge uniformly and randomly on the n vertices. Let B be the event that all bins contain at least $k+1$ balls. Then $\mathcal{M}(n, \mathbf{m}, k+1)$ equals $\mathcal{M}(n, \mathbf{m}, \mathbf{0}, 0)$ conditioned on the event B. By Lemma 4.5.3 $\mathbf{P}(B) = \Omega(f_{k+1}(\lambda)^n)$ where $\lambda f_k(\lambda) = \mu f_{k+1}(\lambda)$. Given any set S, let $r = s/n$. For any hyperedge of size $h-j$, the probability for it to intersect both S and $\overline{S}$ is $p_{j,r} = 1 - r^{h-j} - (1-r)^{h-j}$. Then $p_{j,r} \geq 1 - \gamma^{h-j} - (1-\gamma)^{h-j} \geq 1 - \gamma^{h-w+1} - (1-\gamma)^{h-w+1}$ for any set S and any $0 \leq j \leq w-1$. Recall from (4.4.1) that m is the total number of hyperedges in G. Then $\mathbf{E}\eta(S) = \sum_{j=0}^{w-1} p_{j,r} m_{h-j} \geq m(1 - \gamma^{h-w+1} - (1-\gamma)^{h-w+1})$ for any given S. Since $m \geq (k-1)n/w$ by (4.4.3),

$$\mathbf{E}\eta(S) \geq (1-\gamma^{h-w+1}-(1-\gamma)^{h-w+1})(k-1)n/w = \Theta_\gamma(k)n, \text{ for any } S \text{ with } \gamma n \leq |S| \leq (1-\gamma)n.$$

Choose

$$\delta_2 = \frac{1-\gamma^{h-w+1}-(1-\gamma)^{h-w+1}}{4wh^2(1-\gamma)}.$$

Then δ_2 depends only on γ and so N_2 also depends only on γ. By the Chernoff bound [17],

$$\mathbf{P}(\eta(S) < h^2\delta_2 ks) \le \mathbf{P}(\eta(S) < h^2\delta_2 k(1-\gamma)n) \le \mathbf{P}\left(\eta(S) < \frac{1}{2}\mathbf{E}\eta(S)\right) \le \exp(-\mathbf{E}\eta(S)/16).$$

Note that the second inequality holds because of the choice of δ_2. So there exists some constant $C > 0$ s.t.

$$\mathbf{P}(\eta(S) < h^2\delta_2 ks \mid B) \le C\exp\left(-\mathbf{E}\eta(S)/16\right) f_{k+1}(\lambda)^{-n} = C\left(\exp\left(-\frac{\mathbf{E}\eta(S)}{16n} - \ln f_{k+1}(\lambda)\right)\right)^n.$$

The number of sets S with $\gamma n \le |S| \le (1-\gamma)n$ is at most 2^n. So the expected number of sets S with $\gamma n \le s \le (1-\gamma)n$ and $\eta(S) < h^2\delta_2 ks$ in $\mathcal{M}(n, \mathbf{m}, k+1)$ is at most

$$C\left(2\exp\left(-\frac{\mathbf{E}\eta(S)}{16n} - \ln f_{k+1}(\lambda)\right)\right)^n.$$

Clearly $f_{k+1}(\lambda) \to 1$ as $k \to \infty$ and $\mathbf{E}\eta(S) = \Theta_\gamma(k)n$ as observed before. Then there exists a constant $N_3 > 0$ depending only on γ such that provided $k > N_3$,

$$2\exp\left(-\frac{\mathbf{E}\eta(S)}{16n} - \ln f_{k+1}(\lambda)\right) < 1.$$

Then provided $k \ge \max\{N_2, N_3\}$,

$$\mathbf{E}(X_{[\gamma n,(1-\gamma)n]}) = o(1).$$

Combining all cases, let $N = \max\{N_1, N_2, N_3, h+w\}$. Then N depends only on γ. We have shown that provided $k > N$, $\mathbf{E}X = o(1)$. Theorem 4.2.5 holds for the probability space $\mathcal{M}(n, \mathbf{m}, k+1)$. Then Theorem 4.2.5 follows by Lemma 4.5.1. ∎

Proof of Lemma 4.5.9. The idea of the proof is as follows. When S is big, by Corollary 4.5.6 there are no such sets with $d(S) < (1-\delta)\mu|S|$. We will see later that $\nu(S) < k|S|$ requires a lot of hyperedges partially contained in S, which is unlikely to happen when S is small enough.

Let $G \in \mathcal{M}(n, \mathbf{m}, k+1)$. Let $D = \sum_{j=0}^{w-1}(h-j)m_{h-j}$ and $\mu = D/n$ as defined in (4.4.1). For any S, let $\rho(S,i)$ denotes the number of hyperedges with exactly i vertices contained in S. Then $\nu(S) < ks$ if and only if $\sum_{i=2}^{h}(i-1)\rho(S,i) > d(S) - ks$. By Corollary 4.5.6, there exists $N_1 > 0$ depending only on δ such that provided $k > N_1$, a.a.s. there is no S such that $s > n/h$ and $d(S) < (1-\delta)\mu s$. So we only need to consider sets S with

$|S| \leq n/h$. We call a vertex set $S \in G$ *bad* if $\log^2 n \leq |S| \leq n/h$, $d(S) < (1-\delta)\mu|S|$ and $\sum_{i=2}^{h}(i-1)\rho(S,i) > d(S) - k|S|$. Let s denote $|S|$.

For any given S, let $p(S)$ denote the probability of S being bad. By Corollary 4.5.5, there exists $N_2 > 0$ and $0 < \alpha < 1$, both depending only on δ, such that provided $k > N_2$, the probability that $d(S) < (1-\delta)\mu s$ is at most $\alpha^{\mu s}$. Let $p(q,t)$ be the probability that that $\sum_{i=2}^{h}(i-1)\rho(S,i)$ is at least t conditional on $d(S) = q$. Then

$$p(S) = \sum_{(k+1)s \leq q \leq (1-\delta)\mu s} p(q, q-ks)\mathbf{P}(d(S) = q). \tag{4.5.8}$$

For the small value of q (or s), we need the following claim, to be proved later.

Claim 4.5.10 *If $q < D/h$, then*

$$p(q,t) \leq \left(\exp\left(\frac{h \ln t}{t}\right) \frac{eh(h-1)^2 q^2}{4tD}\right)^t.$$

In particular, if $t \to \infty$ as $n \to \infty$, then

$$p(q,t) \leq \left(\frac{eh^3 q^2}{4tD}\right)^t.$$

Case 1: $s < 2n/eh^3(k+1)$. Since $(k+1)s \leq q \leq (1-\delta)\mu s < D/h$, we have

$$\frac{q}{q-ks} \leq \frac{(k+1)s}{(k+1)s - ks} = k+1, \qquad \frac{q}{D} \leq \frac{(1-\delta)\mu s}{\mu n} < \frac{s}{n}, \qquad q - ks \geq s \geq \log^2 n.$$

So $q - ks \to \infty$ as $n \to \infty$. By (4.5.8) and the particular case of Claim 4.5.10, we have

$$\begin{aligned} p(S) &\leq \sum_{(k+1)s \leq q \leq (1-\delta)\mu s} \left(\frac{eh^3(k+1)s}{4n}\right)^{q-ks} \mathbf{P}(d(S) = q) \\ &\leq \left(\frac{eh^3(k+1)s}{4n}\right)^s \mathbf{P}\Big((k+1)s \leq d(S) \leq (1-\delta)\mu s\Big) \leq \left(\frac{eh^3(k+1)s}{4n}\right)^s \alpha^{\mu s}. \end{aligned}$$

Note that the second inequality above holds because $q - ks \geq s$ and $0 < eh^3(k+1)s/4n < 1$ since $s < 2n/eh^3(k+1)$.

Then the expected number of bad sets S with $|S| = s$, for any fixed $\log^2 n \leq s < 2n/eh^3(k+1)$, is at most

$$\binom{n}{s}\left(\frac{eh^3(k+1)s}{4n}\right)^s \alpha^{\mu s} \leq \left(\frac{en}{s} \cdot \alpha^{\mu} \cdot \frac{eh^3(k+1)s}{4n}\right)^s = \left(e^2 h^3 (k+1)\alpha^{\mu}/4\right)^s.$$

Since $\mu \geq h(k-1)/w$ by (4.4.3), this is at most $\exp(-s)$ provided $k \geq N_3$ for some $N_3 > 0$ depending only on α.

Case 2: $s \geq 2n/eh^3(k+1)$. Take $p(q, q-ks) \leq 1$ since $p(q, q-ks)$ is a probability. So the expected number of bad sets S with $|S| = s$, for any fixed $2n/eh^3(k+1) \leq s \leq (1-\delta)\mu s$, is at most

$$\binom{n}{s} \cdot \alpha^{\mu s} = \left(\frac{en}{s}\alpha^{\mu}\right)^s \leq \left(e^2 h^3(k+1)\alpha^{\mu}/2\right)^s \leq \exp(-s),$$

whenever $k > N_4$ for some $N_4 > 0$ depending only on α. Since α depends only on δ, N_3 and N_4 also depend only on δ. Let $N = \max\{N_1, N_2, N_3, N_4\}$. Then N depends only on δ and provided $k > N$, the expected number of bad S is at most

$$\sum_{\log^2 n \leq s \leq n/h} \exp(-s) = o(1).$$

Lemma 4.5.9 follows thereby. ∎

It only remains to prove Claim 4.5.10.

Proof of Claim 4.5.10. To illustrate the method of computing $p(q, t)$, we show in detail the case $h = 2$ first. Conditional on that $d(S) = q$, we want to estimate the probability that there are at least t edges in S. Consider the alternative algorithm that generates the probability space of the partition-allocation model $\mathcal{P}(n, m_2, 0, k+1)$. Fix any allocation which allocates q balls into bins representing vertices in S with each bin containing at least $k+1$ balls. There are at most

$$\binom{q}{2t}\frac{(2t)!}{2^t t!}$$

partial partitions that contain t parts within S. The probability of every such partial partition to occur is

$$\prod_{i=0}^{t-1} \frac{1}{D-1-2i}.$$

So

$$p(q,t) \leq \binom{q}{2t}\frac{(2t)!}{2^t t!} \cdot \prod_{i=0}^{t-1} \frac{1}{D-1-2i},$$

which is at most

$$\frac{[q]_t}{2^t t!} \prod_{i=0}^{t-1} \frac{q-t-i}{D-1-2i} \leq \left(\frac{eq}{2t}\right)^t \left(\frac{q-t}{D-1}\right)^t \leq \left(\frac{eq}{2t} \cdot \frac{q}{D}\right)^t$$

Note that the second inequality holds since $q < D/2$ and so $q - t < (D-1)/2$.

Now we estimate $p(q,t)$ in the general case $h \geq 2$. Consider the alternative algorithm that generates the probability space of the partition-allocation model $\mathcal{P}(n, \mathbf{m}, \mathbf{0}, k+1)$, defined in Section 4.3. Fix any allocation that allocates exactly q balls into S with each bin containing at least $k+1$ balls. The algorithm uniformly randomly partitions balls into parts such that there are exactly m_{h-j} parts with size $h-j$ for $j = 0, \ldots, w-1$. Let $\mathcal{U} = \{(u_2, \ldots, u_h) \in I\!N^{(h-1)} : \sum_{i=2}^{h}(i-1)u_i = t\}$. Let $\mathbf{u} = (u_2, \ldots, u_h)$ be an arbitrary vector from $\mathcal{U}$. We over estimate the probability that $\rho(S,i)$ is at least u_i for all $i = 2, \ldots, h$, conditional on $d(S) = q$. Let $p(q, \mathbf{u})$ denote this probability. Then clearly $p(q,t) \leq \sum_{\mathbf{u} \in \mathcal{U}} p(q, \mathbf{u})$. The number of partial partitions that contain u_i partial parts of size i within S is

$$\binom{q}{u_1, 2u_2, 3u_3, \ldots, hu_h} \frac{(2u_2)!}{2!^{u_2} u_2!} \cdots \frac{(hu_h)!}{h!^{u_h} u_h!}, \tag{4.5.9}$$

where $u_1 = q - \sum_{i=2}^{h} iu_i$. For any such partial partition we compute the probability that it occurs. The algorithm starts from picking a ball v unpartitioned in S and then it chooses at most $h-1$ balls that are u.a.r. chosen from all the unpartitioned balls to be partitioned into the part containing v.

The probability of the occurrence of a given u_2 partial parts of size 2 within S is at most

$$\prod_{i=0}^{u_2} (h-1) \frac{1}{D-1-hi} = (h-1)^{u_2} \frac{1}{D-1} \cdot \frac{1}{D-h-1} \cdots \frac{1}{D-1-h(u_2-1)}.$$

The probability of the occurrence of a given u_3 partial parts of size 3 within S is at most

$$\prod_{i=0}^{u_3-1} \binom{h-1}{2} \frac{1}{D-hu_2-hi-1} \cdot \frac{1}{D-hu_2-hi-2} \leq (h-1)^{2u_3} \prod_{i=0}^{u_3-1} \frac{1}{(D-hu_2-hi-1)^2}.$$

Note that the above inequality holds because $h\sum_{i=2}^{h} u_i \leq hq/2 < D/2$. Keeping the analysis in this procedure, we obtain that the probability of a particular partial partition with u_i partial parts of size i within S is at most

$$(h-1)^{u_2+2u_3+\cdots+(h-1)u_h} \times \prod_{i=0}^{u_2-1} \frac{1}{D-hi-1} \prod_{i=0}^{u_3-1} \frac{1}{(D-hu_2-hi-1)^2}$$
$$\times \cdots \times \prod_{i=0}^{u_{h-1}-1} \frac{1}{(D-h\sum_{j=2}^{h-2} u_j - hi - 1)^{h-1}} \prod_{i=0}^{u_{h-1}-1} \frac{1}{(D-h\sum_{j=2}^{h-1} u_j - hi - 1)^{h-1}}.$$

The product of this and (4.5.9) gives an upper bound of $p(q, \mathbf{u})$, which is at most

$$\frac{[q]_{\sum_{i=2}^{h} iu_i}(h-1)^t}{u_2!u_3!\cdots u_h! 2!^{u_2} \cdots h!^{u_h}!} \prod_{i=0}^{u_2-1} \frac{1}{D-hi-1} \cdots \prod_{i=0}^{u_h-1} \frac{1}{(D-h\sum_{j=2}^{h-1} u_j - hi - 1)^{h-1}}.$$

Since $2!^{u_2} \cdots h!^{u_h}! \geq 2^t$, this is at most

$$\frac{[q]_t(h-1)^t}{u_2!u_3!\cdots u_h!2^t} \prod_{i=0}^{u_2-1} \frac{q-t-i}{D-hi-1} \times \cdots \times \prod_{i=0}^{u_h-1} \frac{q-t-\sum_{j=2}^{h-1} u_j - i}{(D-h\sum_{j=2}^{h-1} u_j - hi - 1)^{h-1}}.$$

Since $q < D/h$ and so $q-t \leq (D-1)/h$, this is at most

$$\begin{aligned}
&\frac{(eq(h-1)/2)^t}{u_2^{u_2}\cdots u_h^{u_h}} \left(\frac{q-t}{D-1}\right)^{u_2} \left(\frac{q-t-u_2}{(D-hu_2-1)^2}\right)^{u_3} \cdots \left(\frac{q-t-\sum_{j=2}^{h-1} u_j}{(D-h\sum_{j=2}^{h-1} u_j - 1)^{h-1}}\right)^{u_h} \\
&\quad = \frac{(eq(h-1)/2)^t}{u_2^{u_2}\cdots u_h^{u_h}(q-t-u_2)^{u_3}\cdots(q-t-\sum_{j=2}^{h-1} u_j)^{(h-2)u_h}} \\
&\quad \times \left(\frac{q-t}{D-1}\right)^{u_2} \left(\frac{q-t-u_2}{D-hu_2-1}\right)^{2u_3} \cdots \left(\frac{q-t-\sum_{j=2}^{h-1} u_j}{D-h\sum_{j=2}^{h-1} u_j - 1}\right)^{(h-1)u_h} \\
&\quad \leq \frac{(eq(h-1)/2)^t}{u_2^{u_2}(u_3(q-t-u_2))^{u_3}\cdots(u_h(q-t-\sum_{j=2}^{h-1} u_j)^{h-2})^{u_h}} \left(\frac{q-t}{D-1}\right)^t \\
&\quad \leq \frac{(eq(h-1)/2)^t}{u_2^{u_2}(u_3(q-t-u_2))^{u_3}\cdots(u_h(q-t-\sum_{j=2}^{h-1} u_j)^{h-2})^{u_h}} \left(\frac{q}{D}\right)^t.
\end{aligned}$$

Since $q \geq \sum_{i=2}^h iu_i$ and $t = \sum_{i=2}^h (i-1)u_i$, $q-t-\sum_{j=2}^i u_j \geq \sum_{j=i+1}^h u_j \geq u_{i+1}$ for all $2 \leq i \leq h-1$, and so

$$u_2^{u_2}(u_3(q-t-u_2))^{u_3}\cdots\left(u_h\left(q-t-\sum_{j=2}^{h-1} u_j\right)^{h-2}\right)^{u_h} \geq u_2^{u_2}u_3^{2u_3}\cdots u_h^{(h-1)u_h}.$$

We prove the following claim later.

Claim 4.5.11 *Let* $t = \sum_{j=2}^h (j-1)u_j$. *Then*

$$u_2^{u_2}u_3^{2u_3}\cdots u_h^{(h-1)u_h} \geq \left(\frac{2t}{h(h-1)}\right)^t.$$

By Claim 4.5.11, for any $h \geq 2$,

$$p(q,\mathbf{u}) \leq \left(\frac{eqh(h-1)^2}{4t} \cdot \frac{q}{D}\right)^t, \quad \forall \mathbf{u} \in \mathcal{U}.$$

Since $|\mathcal{U}| < t^h$, we have

$$p(q,t) \leq t^h \left(\frac{eh(h-1)^2q^2}{4tD}\right)^t = \left(\exp\left(\frac{h\ln t}{t}\right)\frac{eh(h-1)^2q^2}{4tD}\right)^t.$$

In particular, if $t \to \infty$ as $n \to \infty$, then $h \ln t/t \to 0$ and so $\exp(h \ln t/t) \leq (h/(h-1))^2$ provided n is large enough. So

$$p(q,t) \leq \left(\frac{eh^3q^2}{4tD}\right)^t . \blacksquare$$

Proof of Claim 4.5.11. We solve the following optimization problem

$$\begin{array}{lll} (P_1) & \min & m_2^{m_2} m_3^{2m_3} \cdots m_h^{(h-1)m_h} \\ & s.t. & m_2 + 2m_3 + \cdots + (h-1)m_h = t \\ & & m_2, m_3, \ldots, m_h \geq 0 \end{array}$$

Letting $x_i = (i-1)m_i$ for $2 \leq i \leq h$, and taking the logarithm of the objective function, (P_1) is equivalent to the following optimization problem.

$$\begin{array}{lll} (P_2) & \min & x_2 \ln x_2 + x_3 \ln(x_3/2) + \cdots + x_h \ln(x_h/(h-1)) \\ & s.t. & x_2 + x_3 + \cdots + x_h = t \\ & & x_2, x_3, \ldots, x_h \geq 0 \end{array}$$

For convention, let $x \ln x = 0$ if $x = 0$. Applying the Lagrange multiplier yields $\mathbf{x}^* = (x_2^*, x_3^*, \ldots, x_h^*)$ with $x_i = 2t(i-1)/h(h-1)$, which is a feasible solution of (P_2). In order to show that this is an optimal solution, we need to show that the optimal solution does not appear on the boundary.

Let $\mathbf{x}$ be any solution on the boundary of (P_2). Then there exists $2 \leq i \leq h$ such that $x_i = 0$. There also exists j with $x_j > 0$. Consider $\mathbf{x}'$ with $x'_i = (i-1)x_j/h$, $x'_j = x_j - (i-1)x_j/h$ and $x'_l = x_l$ for any $l \neq i, j$. Then $\mathbf{x}'$ is feasible and it is straightforward to check that

$$x'_i \ln(x'_i/(i-1)) + x'_j \ln(x'_j/(j-1)) < x_i \ln(x_i/(i-1)) + x_j \ln(x_j/(j-1)).$$

Hence $\mathbf{x}'$ cannot be an optimal solution. This proves that $\mathbf{x}^*$ is the minimizer and so the optimal value of (P_1) is $\exp(t \ln(2t/(h(h-1)))) = (2t/(h(h-1)))^t$. $\blacksquare$

4.6 Discussion of a more general setting

In a more general setting, let H be a hypergraph on n vertices and $M(n)$ hyperedges. Each hyperedge x_i is associated with a parameter w_i. Call $\mathbf{w} = (w_i)_{i=1}^M$ the parameter vector. Let h_i be the length of the hyperedge x_i. Then $\mathbf{h} = (h_i)_{i=1}^M$ is called the length vector.

We say that H is $(\mathbf{w}, k)$-orientable if there exists a w_i-orientation for each hyperedge x_i such that the maximum indegree is at most k. Define $\mathcal{G}_{n,M,\mathbf{h}}$ to be the probability space of hypergraphs on n vertices, M hyperedges with the length vector of hyperedges $\mathbf{h}$. We can define the $(\mathbf{w}, k)$-core in a similar way as the (w, k)-core. Let $\widehat{H}$ denote the $(\mathbf{w}, k)$-core of $H \in \mathcal{G}_{n,M,\mathbf{h}}$. Let n', M' denote the number of vertices and the number of hyperedges in $\widehat{H}$. Let $\mathbf{h}' = (h'_i)_{i=1}^{M'}$ and $\mathbf{w}' = (w'_i)_{i=1}^{M'}$ be the length vector and the parameter vector of $\widehat{H}$. Then clearly if $h'_i = h_i - j$ for some j, then $w'_i = w_i - j$. Define the density of $\widehat{H}$ to be $\sum_{i=1}^{M'} w'_i/n'$. If there exists a constant $h > 0$ such that $h_i \leq h$ for any $i \leq M$, then in principle, we can analyse the properties of $\widehat{H}$ using the differential equation method and we believe the density of $\widehat{H}$ is a.a.s. determined by some differential equation system. However, it is hard to determine the threshold when the density of $\widehat{H}$ is at most k. It may be possible to adapt proofs in this chapter to prove the following conjecture, but we haven't checked this.

Conjecture 4.6.1 *A random hypergraph H in $\mathcal{G}_{n,M,\mathbf{h}}$ is a.a.s. $(\mathbf{w}, k)$-orientable if for some constant $\epsilon > 0$, the density of the $(\mathbf{w}, k)$-core of H is at most $k - \epsilon$.*

Chapter 5

Probabilities of induced subgraphs in random regular graphs

5.1 Introduction

A graph G is called a *B-graph with vertex bipartition L and R* if $V(G) = L \cup R$, and L is an independent set of G. A graph G is called a *B-pseudograph with vertex bipartition L and R* if G is a multi-graph, namely, loops and multiple edges are allowed, and $V(G) = L \cup R$ with L being an independent set of G. An edge in a B-graph or B-pseudograph is called a *crossing edge* if its end vertices are in L and R respectively. An edge is called an *embedding edge* if its end vertices are both in R. Given $\mathbf{d}$ as a nonnegative integer vector, let $\mathcal{G}(L, R, \mathbf{d})$ be the set of B-graphs with bipartition L and R and the degree sequence $\mathbf{d}$ and let $g(L, R, \mathbf{d}) = |\mathcal{G}(L, R, \mathbf{d})|$. By convention, let $g(L, R, \mathbf{d}) = 0$ if $\mathbf{d}$ is not nonnegative.

Counting B-graphs directly leads to the study of induced subgraphs. Let S be a vertex subset of $[n]$ and let s denote $|S|$. Without loss of generality, we can assume $S = [s]$ by simply relabelling the vertices in $[n]$. Let H be a graph on the vertex set S with degree sequence $(k_i)_{1\le i\le s}$. Let $\mathbf{d}'$ be the integer vector defined by $d'_i = d - k_i$ for $i \in S$ and $d'_i = d$ for $i \in [n] \setminus S$. Then the number of d-regular graphs with $G_S = H$ is $g(S, [n] \setminus S, \mathbf{d}')$. Therefore, the probability that $G_S = H$ in $\mathcal{G}_{n,d}$ is equal to $g(S, [n] \setminus S, \mathbf{d}')$ divided by the number of d-regular graphs.

The asymptotic formula for $\mathbf{P}(G_S = H)$ in $\mathcal{G}_{n,\mathbf{d}}$ is given in Section 5.2 together with some direct applications of the result. Proofs are given in Section 5.3 together with the introduction of switching operations that are used in our proofs.

5.2 Main results

For any positive integer m, let $[m]$ denote the set $\{1, 2, \dots, m\}$. Given the degree sequence $\mathbf{d} = (d_1, \dots, d_n)$, let $d_{\max} = \max\{d_i, i \in [n]\}$. Let $M(\mathbf{d}) = \sum_{i=1}^{n} d_i$ and $M_2(\mathbf{d}) = \sum_{i=1}^{n} d_i(d_i - 1)$. Define $\mu(\mathbf{d})$ to be $M_2(\mathbf{d})/2M(\mathbf{d})$.

The following theorem gives the asymptotic formula of the number of graphs with given degree sequences when $d_{\max}$ is not too large. Note that the restriction of $d_{\max}$ was relaxed further by McKay and Wormald [43] as introduced in Section 2.3.1, but it requires a few extra terms in the exponential factor of the asymptotic formula. However, the result we cited as follows is enough for the purpose of this chapter.

Theorem 5.2.1 (McKay [42]) *For any degree sequence* $\mathbf{d} = (d_1, \dots, d_n)$ *with* $d_{\max} = o(M(\mathbf{d})^{1/4})$, *the number of graphs with degree sequence* $\mathbf{d}$ *is uniformly*

$$\frac{M(\mathbf{d})!}{2^{M(\mathbf{d})/2}(M(\mathbf{d})/2)! \prod_{i=1}^{n} d_i!} \cdot \exp\left(-\mu(\mathbf{d}) - \mu(\mathbf{d})^2 + O(d_{\max}^4/M(\mathbf{d}))\right).$$

Note that by uniform in the above theorem we mean $O(.)$ is uniform over all $\mathbf{d}$, i.e. $g(\mathbf{d}) = O(f(\mathbf{d}))$ implies there exists an absolute constant $C > 0$ such that $|g(\mathbf{d})| < Cf(\mathbf{d})$ for all $\mathbf{d}$.

A special case of Theorem 5.2.1 gives that the number of d-regular graphs on n vertices is asymptotically

$$\frac{(dn)!}{2^{dn/2}(dn/2)!(d!)^n} \cdot \exp\left(-\frac{d^2 - 1}{4}\right),$$

when $d = o(n^{1/3})$.

Next we estimate $g(L, R, \mathbf{d})$. Let $M(\mathbf{d}) = \sum_{i=1}^{n} d_i$. For any $S \subset L \cup R$, define $M_1(\mathbf{d}, S) = \sum_{i \in S} d_i$ and $M_2(\mathbf{d}, S) = \sum_{i \in S} d_i(d_i - 1)$. Define

$$\mu_0(\mathbf{d}, L, R) = \frac{(M_1(\mathbf{d}, R) - M_1(\mathbf{d}, L))M_2(\mathbf{d}, R)}{2M_1(\mathbf{d}, R)^2}, \tag{5.2.1}$$

$$\mu_1(\mathbf{d}, L, R) = \frac{M_2(\mathbf{d}, R)M_2(\mathbf{d}, L)}{2M_1(\mathbf{d}, R)^2}, \tag{5.2.2}$$

$$\mu_2(\mathbf{d}, L, R) = \mu_0(\mathbf{d}, L, R)^2. \tag{5.2.3}$$

Note that we drop the notations L and R from $\mu_i(\mathbf{d}, L, R)$ for $i = 1, 2$ when the context is clear. Note also that if $M_1(\mathbf{d}, R) < M_1(\mathbf{d}, L)$, then $g(L, R, \mathbf{d})$ is trivially 0. So we may assume that

$$M_1(\mathbf{d}, R) \geq M_1(\mathbf{d}, L). \tag{5.2.4}$$

The following theorem gives the asymptotic formula of $g(L, R, \mathbf{d})$.

Theorem 5.2.2 *For any degree sequence* $\mathbf{d}$ *such that* $d_{\max} = o(M(\mathbf{d})^{1/4})$ *and* $M_1(\mathbf{d}, R) \geq M_1(\mathbf{d}, L)$*, we have uniformly,*

$$g(L, R, \mathbf{d}) = \left(1 + O\left(\frac{d_{\max}^2 + \ln M(\mathbf{d})}{\sqrt{M(\mathbf{d})}}\right)\right) \frac{M_1(\mathbf{d}, R)! \exp(-\mu_0(\mathbf{d}) - \mu_1(\mathbf{d}) - \mu_2(\mathbf{d}))}{2^{(M_1(\mathbf{d},R)-M_1(\mathbf{d},L))/2}((M_1(\mathbf{d}, R) - M_1(\mathbf{d}, L))/2)! \prod_{i=1}^{n} d_i!}.$$

Let $S = [s]$ be a subset of the vertex set $[n]$ and let H be a graph on the vertex set S. Let $\mathbf{P}_{\mathcal{G}_{n,\mathbf{d}}}(S, H)$ denote the probability that $G_S = H$ in $\mathcal{G}_{n,\mathbf{d}}$. Applying Theorem 5.2.2 and Theorem 5.2.1 we directly get the following theorem.

Theorem 5.2.3 *For any degree sequence* $\mathbf{d}$ *with* $d_{\max} = o(M(\mathbf{d})^{1/4})$ *and* $0 < s < n$*, let* $S = [s] \subset [n]$ *and let* H *be a graph on the vertex set* S *with the degree sequence* $\mathbf{k} = (k_1, \ldots, k_s)$*. Let* $h = \sum_{i=1}^{h} k_i$ *and let* $\mathbf{d}' = (d'_1, \ldots, d'_n)$ *with* $d'_i = d_i - k_i$ *for* $i \in S$ *and* $d'_i = d_i$ *for* $i \notin S$*. If* $d'_i < 0$ *for some* $i \in [n]$ *or* $M_1(\mathbf{d}', [n] - S) < M_1(\mathbf{d}', S)$*, then* $\mathbf{P}_{\mathcal{G}_{n,\mathbf{d}}}(S, H) = 0$*. Otherwise, if* $d'_{\max} = o(M(\mathbf{d}')^{\mathbf{1/4}})$*, then uniformly,*

$$\mathbf{P}_{\mathcal{G}_{n,\mathbf{d}}}(S, H) = \exp\left(-\mu_0(\mathbf{d}') - \mu_1(\mathbf{d}') - \mu_2(\mathbf{d}') + \mu(\mathbf{d}) + \mu(\mathbf{d})^2\right) \prod_{i=1}^{s} [d]_{k_i}$$
$$\frac{M_1(\mathbf{d}', [n] - S)! 2^{M_1(\mathbf{d}', S) + h/2} (M(\mathbf{d})/2)!}{((M_1(\mathbf{d}', [n] - S) - M_1(\mathbf{d}', S))/2)! M(\mathbf{d})!} \left(1 + O\left(\frac{{d'_{\max}}^2 + \ln M(\mathbf{d}')}{\sqrt{M(\mathbf{d}')}} + \frac{d_{\max}^4}{M(\mathbf{d})}\right)\right).$$

Proof

$$\mathbf{P}_{\mathcal{G}_{n,\mathbf{d}}}(S, H) = \frac{g(S, [n] - S, \mathbf{d}')}{\#\text{graphs with degree sequence } \mathbf{d}}.$$

The value of $g(S, [n] - S, \mathbf{d}')$ is given by Theorem 5.2.2 and the number of d-regular graphs is given by Theorem 5.2.1. Theorem 5.2.3 then follows. ∎

Let $\mathbf{P}_{\mathcal{G}_{n,d}}(S, H)$ denote the probability that $G_S = H$ in the probability space of random d-regular graphs. Then the following corollary follows from Theorem 5.2.3 and by applying the Stirling's formula.

Corollary 5.2.4 *Given* $0 < s < n$*, let* $S = [s] \subset [n]$ *and let* H *be a graph on the vertex set* S *with the degree sequence* $\mathbf{k} = (k_1, \ldots, k_s)$*. Let* $h = \sum_{i=1}^{h} k_i$*. Assume* $d = o((n-s)^{1/3})$ *and* $dn - 2ds + h \to \infty$*,*

$$\mathbf{P}_{\mathcal{G}_{n,d}}(S, H) = \left(1 + O\left(\frac{d^2 + \ln(dn - h)}{\sqrt{dn - h}}\right)\right) \exp\left(-\mu_0(\mathbf{d}') - \mu_1(\mathbf{d}') - \mu_2(\mathbf{d}') + \frac{d^2 - 1}{4}\right)$$
$$e^{h/2} \prod_{i=1}^{s} [d]_{k_i} \sqrt{\frac{dn - ds}{dn - 2ds + h}} \cdot \frac{(dn - ds)^{dn - ds}}{(dn - 2ds + h)^{(dn - 2ds + h)/2} (dn)^{dn/2}},$$

where $d'_i = d - k_i$ *for* $i \in S$ *and* $d'_i = d$ *for* $i \notin S$.

Proof By the definition of $\mu(\mathbf{d})$, we immediately get that $\mu(\mathbf{d}) + \mu(\mathbf{d})^2 = (d^2 - 1)/4$ when $\mathbf{d}$ is a constant sequence with each term d. We also have that $M(\mathbf{d}')$ in Theorem 5.2.3 equals to $dn - h$ in this special case and $d'_{\max} \le d$. By Theorem 5.2.3,

$$\mathbf{P}_{\mathcal{G}_{n,d}}(S,H) = \left(1 + O\left(\frac{d^2 + \ln(dn-h)}{\sqrt{dn-h}}\right)\right) \exp\left(-\mu_0(\mathbf{d}') - \mu_1(\mathbf{d}') - \mu_2(\mathbf{d}') + \frac{d^2-1}{4}\right)$$
$$\prod_{i=1}^{s} [d]_{k_i} \frac{M_1(\mathbf{d}', [n]-S)! 2^{M_1(\mathbf{d}',S)+h/2} (M(\mathbf{d})/2)!}{((M_1(\mathbf{d}',[n]-S) - M_1(\mathbf{d}',S))/2)! M(\mathbf{d})!}$$

Since $M(\mathbf{d}) = dn$, $M_1(\mathbf{d}', [n]-S) = dn - ds$, $M_1(\mathbf{d}', S) = ds - h$, and $n! \sim \sqrt{2\pi n}(n/e)^n$ by the Stirling's formula, we have

$$\frac{M_1(\mathbf{d}', [n]-S)! 2^{M_1(\mathbf{d}',S)+h/2} (M(\mathbf{d})/2)!}{((M_1(\mathbf{d}',[n]-S) - M_1(\mathbf{d}',S))/2)! M(\mathbf{d})!}$$
$$\sim e^{h/2} \sqrt{\frac{dn - ds}{dn - 2ds + h}} \cdot \frac{(dn-ds)^{dn-ds}}{(dn - 2ds + h)^{(dn-2ds+h)/2} (dn)^{dn/2}},$$

and the Corollary 5.2.4 follows. ∎

The formula in Corollary 5.2.4 can be simplified as follows if the size of H is not too large.

Corollary 5.2.5 *Let* S, H, $\mathbf{k}$ *and* h *be defined as in Corollary 5.2.4. If* $d = o(n^{1/3})$, $s^2 d = o(n)$ *and* $d^2 s = o(n)$, *then*

$$\mathbf{P}_{\mathcal{G}_{n,d}}(S,H) = \left(1 + O((d^{3/2} + \ln n)/\sqrt{n} + s^2 d/n + d^2 s/n)\right)(dn)^{-h/2} \prod_{i=1}^{s} [d]_{k_i}.$$

The proofs of Theorem 5.2.3 and Corollary 5.2.5 are provided in Section 5.3.

The following corollary is a special case of Corollary 5.2.4 where H is an empty graph.

Corollary 5.2.6 *Assume* $d = o(n^{1/3})$. *Then for any* $S \subset [n]$, *if* $0 \le |S|/n < 1/2$,

$$\mathbf{P}(S \text{ is an independent set}) =$$
$$\left(1 + O((d^{3/2} + \ln n)/\sqrt{n})\right) \sqrt{\frac{1-\delta}{1-2\delta}} \left(\frac{(1-\delta)^{1-\delta}}{(1-2\delta)^{(1-2\delta)/2}}\right)^{dn} \exp\left(f(d,\delta)\right),$$

where $\delta = \delta(n) = |S|/n$, *and*

$$f(d,\delta) = -\frac{\delta^2}{4(1-\delta)^2} d^2 + \frac{\delta}{2(1-\delta)^2} d + \frac{\delta^2 - 2\delta}{4(1-\delta)^2}.$$

Proof Apply Corollary 5.2.4 with $h = 0$. ∎

5.3 Proofs

We can use the pairing model to generate B-graphs with the vertex partition $L \cup R$ and the degree sequence $\mathbf{d} = \{d_1, \ldots, d_n\}$. Consider n buckets representing the n vertices. Let each bucket i contain d_i points. Take a random pairing over these points. We say a pairing is *restricted* if no pair has both ends in the buckets representing vertices in L. Let $\mathcal{M}(L, R, \mathbf{d})$ be the class of all restricted pairings. A pair in a pairing is called a *mixed (pure) pair* if it corresponds to a mixed (pure) edge in the corresponding B-pseudograph.

Every restricted pairing corresponds to a B-pseudograph by contracting all points in each bucket to form a vertex. On the other hand, any simple B-graph corresponds to $\prod_{i=1^n} d_i!$ restricted pairing in $\mathcal{M}(L, R, \mathbf{d})$. Then all simple B-graphs occur with the same probability in the pairing model.

The main goal of this section is to compute the probability that a B-pseudograph generated by the pairing model is simple. Recall that $\{\{u_1, u_1'\}, \{u_2, u_2'\}, \{u_3, u_3'\}\}$ is a triple pair if u_1, u_2, u_3 are in one bucket and u_1', u_2', u_3' are in another bucket. We call the two buckets the end vertices of the triple pair. If the buckets represent vertices in L and R respectively, then the triple pair is called a *mixed triple pair.* If both buckets represent vertices in R, then the triple pair is called a *pure triple pair.* Given a random restricted pairing, let T_1 and T_2 be the number of mixed and pure triple pairs respectively. In this section, there is only one degree sequence $\mathbf{d}$ that is referred to. So we drop the notation $\mathbf{d}$ from $M(\mathbf{d})$ and $M_i(\mathbf{d}, L)$, $M_i(\mathbf{d}, R)$, $\mu_i(\mathbf{d})$ for simplicity. Since $M_1(R) \geq M_1(L)$ by the assumption (5.2.4), we have $M_1(R) \geq M/2$.

Lemma 5.3.1 *Assume $d_{\max} = o(M^{1/4})$. Then $\mathbf{E}(T_1) = o(1)$ and $\mathbf{E}(T_2) = o(1)$.*

Proof For any two vertices $i \in L$ and $j \in R$, we compute the probability that there is a triple pair with end vertices i and j. There are $\binom{d_i}{3}$ ways to choose three points from the bucket i and $\binom{d_j}{3}$ ways to choose three points from the bucket j. There are 6 ways to match the six chosen points to form a triple pair. Let $U(m)$ for any positive even integer m denote the number of perfect matchings over m points. Then

$$U(m) = \prod_{i=0}^{m/2-1} (m - 2i - 1) = \frac{m!}{2^{m/2}(m/2)!}.$$

The probability for the three particular pairs to occur is

$$\frac{[M_1(R) - 3]_{M_1(L)-3} U(M_1(R) - M_1(L))}{[M_1(R)]_{M_1(L)} U(M_1(R) - M_1(L))} \sim M_1(R)^{-3}.$$

This is because the number of ways to match the remaining $M_1(R)-3$ points in L to points in R, except for the three chosen points in the vertex j, is $[M_1(R)-3]_{M_1(L)-3}$, and the number of matchings over the remaining $M_1(R)-M_1(L)$ points in R is $U(M_1(R)-M_1(L))$, whilst the total number of matchings over all M points with all points in L matched to ones in R is $[M_1(R)]_{M_1(L)}U(M_1(R)-M_1(L))$. Hence we have

$$\begin{aligned}\mathbf{E}(T_1) &\sim \sum_{i\in L}\sum_{j\in R} 6\binom{d_i}{3}\binom{d_j}{3} M_1(R)^{-3} = O\left(\left(\sum_{i\in L} d_i^3\right)\left(\sum_{j\in R} d_j^3\right)\right) M^{-3}\\ &= O\left(\frac{d_{\max}^4 M_1(L)M_1(R)}{M^3}\right) = O\left(\frac{d_{\max}^4}{M}\right) = o(1),\end{aligned}$$

where the second equality uses $M/2 \le M_1(R) \le M$.

Similarly we can compute $\mathbf{E}(T_2)$. We first compute the probability that a triple pair with ends $i, j \in R$ occurs. There are $\binom{d_i}{3}\binom{d_j}{3}$ ways to choose three points from each bucket and six ways to match them up. The probability for any such particular three pairs to occur is

$$\frac{[M_1(R)-6]_{M_1(L)}U(M_1(R)-M_1(L)-6)}{[M_1(R)]_{M_1(L)}U(M_1(R)-M_1(L))},$$

because the number of ways to match all points in L to points in R, except the six chosen points, three from bucket i and j each, is $[M_1(R)-6]_{M_1(L)}$, and the number of matchings over all remaining points in R is $U(M_1(R)-M_1(L)-6)$. Since

$$\frac{[M_1(R)-6]_{M_1(L)}}{[M_1(R)]_{M_1(L)}} \sim \frac{(M_1(R)-M_1(L))^6}{M_1(R)^6}, \text{ and } \frac{U(M_1(R)-M_1(L)-6)}{U(M_1(R)-M_1(L))} \sim (M_1(R)-M_1(L))^{-3},$$

this probability is asymptotically

$$\frac{(M_1(R)-M_1(L))^3}{M_1(R)^6} = O(M_1(R)^{-3}).$$

Therefore,

$$\begin{aligned}\mathbf{E}(T_2) &\sim \sum_{i\in R}\sum_{j\in R} 6\binom{d_i}{3}\binom{d_j}{3} M_1(R)^{-3} = O\left(\left(\sum_{i\in R} d_i^3\right)\left(\sum_{j\in R} d_j^3\right)\right) M^{-3}\\ &= O\left(\frac{d_{\max}^4 M_1(R)^2}{M^3}\right) = O\left(\frac{d_{\max}^4}{M}\right) = o(1). \quad \blacksquare\end{aligned}$$

Let two loops that start from a common vertex be called a double loop and let I be the number of double loops.

Lemma 5.3.2 *Assume* $d_{\max}^4 = o(M)$. *Then* $\mathbf{E}(I) = o(1)$.

Proof We first compute the probability that a double loop starting at i occurs. There are $\binom{d_i}{4}$ ways to choose 4 distinct points from the i-th bucket and there are 3 ways to match up the four points to form a double loop. The probability that any such particular two pairs occur is

$$\frac{[M_1(R)-4]_{M_1(L)}U(M_1(R)-M_1(L)-4)}{[M_1(R)]_{M_1(L)}U(M_1(R)-M_1(L))} \sim \frac{(M_1(R)-M_1(L))^2}{M_1(R)^4} = O(M_1(R)^{-2}),$$

because the number of ways to match all points in L to points in R, except the four chosen points in i, is $[M_1(R)-4]_{M_1(L)}$, and the number of matchings over all remaining points in R is $U(M_1(R)-M_1(L)-4)$. Therefore,

$$\mathbf{E}(I) \sim \sum_{i\in R} 3\binom{d_i}{4} M_1(R)^{-2} = O\left(d_{\max}^3/M_1(R)\right) = O\left(\frac{d_{\max}^3}{M}\right) = o(1). \quad \blacksquare$$

Lemmas 5.3.1 and 5.3.2 show that a.a.s. there are no triple pairs or double loops in a random restricted pairing. So we only need to consider the loops and the double pairs. Recall that a pair $\{u, u'\}$ in a pairing is called a loop if u and u' are contained in the same bucket and two pairs $\{u_1, u_1'\}, \{u_2, u_2'\}$ are called a double pair if u_1, u_2 are in one bucket and u_1', u_2' are in another bucket. In a restricted pairing, there are two types of double pairs. One is that u_1, u_2 are contained in a bucket in L and u_1', u_2' are contained in a bucket in R. The other is that all of u_1, u_2, u_1' and u_2' are contained in buckets in R. We call the former *type 1* and the latter *type 2*.

Let B_0, B_1 and B_2 be the numbers of loops and double pairs of types 1 and 2 respectively. We first compute the expected value of B_i for $i = 0, 1, 2$. Recall from (5.2.1)–(5.2.3) that

$$\mu_0 = \frac{(M_1(R)-M_1(L))M_2(R)}{2M_1(R)^2}, \quad \mu_1 = \frac{M_2(R)M_2(L)}{2M_1(R)^2}, \quad \mu_2 = \mu_0^2.$$

Lemma 5.3.3 *For* $i = 0, 1$, $\mathbf{E}B_i \sim \mu_i$ *and* $\mathbf{E}B_2 = O(\mu_2)$. *More specifically, if* $d_{\max} = o(M^{1/3})$, *then* $\mathbf{E}B_2 = (1+o(1))\mu_2 + o(1)$.

Proof Using the same patten of proofs as for Lemma 5.3.1 and Lemma 5.3.2, we get immediately

$$
\begin{aligned}
\mathbf{E}B_0 &= \sum_{i\in R}\binom{d_i}{2}\frac{[M_1(R)-2]_{M_1(L)}U(M_1(R)-M_1(L)-2)}{[M_1(R)]_{M_1(L)}U(M_1(R)-M_1(L))} \\
&\sim \sum_{i\in R}\frac{[d_i]_2}{2}\frac{M_1(R)-M_1(L)}{M_1(R)^2}=\mu_0; \\
\mathbf{E}B_1 &= \sum_{i\in L}\sum_{j\in R}2\binom{d_i}{2}\binom{d_j}{2}\frac{[M_1(R)-2]_{M_1(L)-2}U(M_1(R)-M_1(L))}{[M_1(R)]_{M_1(L)}U(M_1(R)-M_1(L))} \\
&\sim \frac{M_2(L)M_2(R)}{2}M_1(R)^{-2}=\mu_1; \\
\mathbf{E}B_2 &= \sum_{i,j\in R,i<j}2\binom{d_i}{2}\binom{d_j}{2}\frac{[M_1(R)-4]_{M_1(L)}U(M_1(R)-M_1(L)-4)}{[M_1(R)]_{M_1(L)}U(M_1(R)-M_1(L))} \\
&= \frac{1}{2}\sum_{i\in R}\sum_{j\in R}2\binom{d_i}{2}\binom{d_j}{2}\frac{[M_1(R)-4]_{M_1(L)}U(M_1(R)-M_1(L)-4)}{[M_1(R)]_{M_1(L)}U(M_1(R)-M_1(L))} \\
&\quad -\frac{1}{2}\sum_{i\in R}2\binom{d_i}{2}\binom{d_i}{2}\frac{[M_1(R)-4]_{M_1(L)}U(M_1(R)-M_1(L)-4)}{[M_1(R)]_{M_1(L)}U(M_1(R)-M_1(L))} \qquad (5.3.1) \\
&= (1+o(1))\frac{M_2(R)^2}{4}\frac{(M_1(R)-M_1(L))^2}{M_1(R)^4}-O(d_{\max}^3/M)=(1+o(1))\mu_2-O(d_{\max}^3/M).
\end{aligned}
$$

Note that the fraction 1/2 appears in (5.3.1) because of the double counting of each pair $i<j$ and the term subtracted accounts for the case $i=j$, which, following the argument in the proof of Lemma 5.3.2, is bounded by $O(d_{\max}^3/M)$. Hence the lemma follows. ∎

The following two corollaries follow directly from Lemma 5.3.3 by the first moment method.

Corollary 5.3.4 *Let* $w(n)=\sqrt{M}/d_{\max}^2$. *Then* $\mathbf{P}\big(B_i>w(n)\mu_i\big)=O(w(n)^{-1})=O(d_{\max}^2/\sqrt{M})$ *for* $i=0,1,2$.

Corollary 5.3.5 *Assume* $d_{\max}^4=o(M)$. *Let* $w(n)=\sqrt{M}/d_{\max}^2$. *If* $M_2(R)\le d_{\max}^3 w(n)$, *then the probability that there exists a loop or a double pair is* $O(d_{\max}/\sqrt{M}+d_{\max}^4/M)$.

Proof If $d_{\max}^4=o(M)$ and $M_2(R)\le d_{\max}^3w(n)$, then $\mathbf{E}B_0=O(M_2(R)/M_1(R))=O(d_{\max}^3w(n)/M)=O(d_{\max}/\sqrt{M})$; $\mathbf{E}B_1=O(M_2(L)d_{\max}^3/M^2)=O(d_{\max}^4/M)$; $\mathbf{E}B_2=O(d_{\max}^6w(n)^2/M^2)=O(d_{\max}^2/M)$. ∎

Given $\mathcal{P}$ as a restricted pairing, we say the ordered pair of pairs $((u_1,u_1'),(u_2,u_2'))$ form a directed 2-path in $\mathcal{P}$ if u_1' and u_2 lie in the same bucket and the three buckets where u_1, u_1' and u_2' lie in respectively are all distinct. We then say that the two pairs (u_1, u_1') and (u_2, u_2') are adjacent. For instance, the ordered pair of pairs $((1,2),(3,4))$ form a directed 2-path in the four examples in Figure 5.1. Note that a directed 2-path in a pairing corresponds to a directed 2-path in the corresponding B-pseudograph. Let v denote the bucket where u_1' and u_2 lie in. We say the directed 2-path $((u_1, u_1'), (u_2, u_2'))$ in $\mathcal{P}$ is *simple* if neither of $\{u_1, u_1'\}$ and $\{u_2, u_2'\}$ is contained in a double pair and there is no loop at v.

We study the four types of directed 2-paths as illustrated in Figure 5.1. The first type refers to those with vertices lying in R; the second type refers to directed 2-paths $((1,2),(3,4))$ with 1 lying in a bucket in L, 2 and 3 both lie in one bucket in R and 4 lies in another bucket in R; the third type refers to directed 2-paths $((1,2),(3,4))$ with 1 and 4 each lying in a bucket in L and 2 and 3 both lying in one bucket in R; the fourth type refers to directed 2-paths $((1,2),(3,4))$ with 1 and 4 each lying in a bucket in R and 2 and 3 both lying in one bucket in L. Let these four types of directed 2-paths be called type 1, 2, 3 and 4 respectively.

Given $\mathcal{P}$, let t be the number of pure pairs in $\mathcal{P}$. Then $t = (M_1(R) - M_1(L))/2$. Let $A_i(\mathcal{P})$ denote the number of simple directed 2-paths of type i for $i = 1,2,3,4$ and let $a_i(l_0, l_1, l_2) = \mathbf{E}(A_i(\mathcal{P}) \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})$. Clearly $A_4(\mathcal{P}) = \sum_{i\in L} d(i)(d(i)-1) - O(l_1 d_{\max}) = M_2(L) - O(l_1 d_{\max})$ for any $\mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}$ since the number of non-simple directed 2-path of type 4 is bounded by $O(l_1 d_{\max})$.

Let $\mathcal{C}_{l_0,l_1,l_2}$ be the class of restricted pairings in $\mathcal{M}(L, R, \mathbf{d})$ that contains l_0 loops, l_1 double pairs of types 1, l_2 double pairs of type 2 and no double loop or triple pairs. We use some switching operations to estimate the ratios of $|\mathcal{C}_{l_0,l_1,l_2}|/|\mathcal{C}_{l_0-1,l_1,l_2}|$, etc. The switching operations we are going to use are ideologically similar to, although look different from, the switching operations used by McKay and Wormald [43]. The switching operations used in [43] do not apply here because the resulting pairing does not remain restricted after they are applied to a given pairing. However, they can be easily adjusted and adapted to our case.

The following two switching operations are used to prove Lemma 5.3.6.

L_1-switching: take a loop $\{2,3\}$ and two pure pairs $\{1,5\}$, $\{4,6\}$ such that the six points are located in the five distinct buckets as drawn in Figure 5.2. Replace the three pairs $\{2,3\},\{1,5\},\{4,6\}$ by $\{1,2\},\{3,4\},\{5,6\}$.

L_2-switching: take a loop $\{2,3\}$ and two mixed pairs $\{1,5\}$, $\{4,6\}$ such that the six points are located in the five distinct buckets as drawn in Figure 5.3. Replace the three pairs $\{2,3\},\{1,5\},\{4,6\}$ by $\{1,2\},\{3,4\},\{5,6\}$.

For any switching operation that converts a pairing P_1 to another pairing P_2, we define the operation that converts P_2 to P_1 the inverse of the switching operation under discussion.

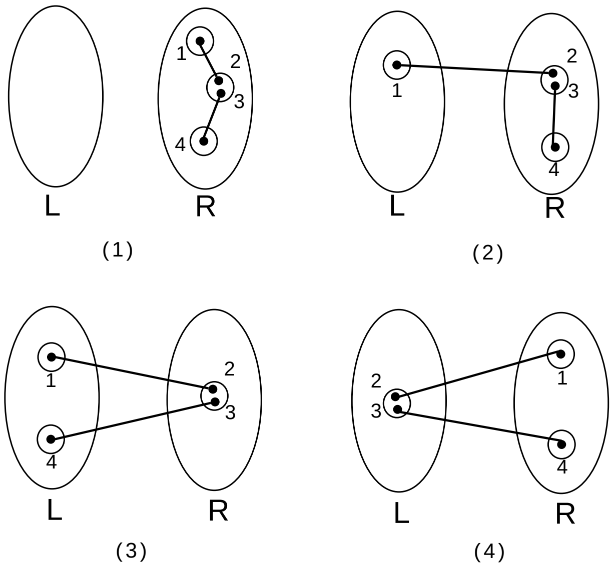

Figure 5.1: *four different types of* 2*-paths*

Thus, the *inverse* L_1*-switching* operation can be defined as follows. Take a 2-directed path (not necessarily simple) $((1, 2), (3, 4))$ of type 1 and a pure pair $\{5, 6\}$ such that the points 1, 2, 4, 5 and 6 lie in five distinct buckets. Replace $\{1, 2\}$, $\{3, 4\}$ and $\{5, 6\}$ by $\{2, 3\}$, $\{1, 5\}$ and $\{4, 6\}$. The inverse L_2-switching can be defined in the same way.

Recall that $a_i(l_0, l_1, l_2) = \mathbf{E}(A_i(\mathcal{P}) \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})$ for $1 \le i \le 4$. The following lemma gives the estimation of the ratio $|\mathcal{C}_{l_0,l_1,l_2}|/|\mathcal{C}_{l_0-1,l_1,l_2}|$, which is approached by counting ways to perform certain L_1-switchings and the inverse L_1-switchings. Counting these switching operations involves the estimation of $a_i(l_0, l_1, l_2)$. Since the computation of $a_i(l_0, l_1, l_2)$ for $i = 1, 2, 3$ is not easy, we postpone their calculation and express the result in Lemma 5.3.6 (and also in Lemmas 5.3.7 and 5.3.8) in terms of $a_i(l_0, l_1, l_2)$.

Lemma 5.3.6 *Let* $a_1 = a_1(l_0 - 1, l_1, l_2)$ *and* $a_3 = a_3(l_0 - 1, l_1, l_2)$*. Assume* $l_0 \ge 1$*. Then*

(i) : *If* $t \ge 1$,

$$\frac{|\mathcal{C}_{l_0,l_1,l_2}|}{|\mathcal{C}_{l_0-1,l_1,l_2}|} = \frac{a_1}{4l_0 t}(1 + O(d_{\max}^2/t + (l_0 + l_2)/t)),$$

(ii) : *If* $M_1(L) \ge 1$ *and* $a_3 \ge 1$,

$$\frac{|\mathcal{C}_{l_0,l_1,l_2}|}{|\mathcal{C}_{l_0-1,l_1,l_2}|} = \frac{ta_3}{l_0 M_1(L)^2}(1 + O(d_{\max}^2/M_1(L) + d_{\max}^3/a_3 + l_1/M_1(L) + (l_0 + l_2)/t)).$$

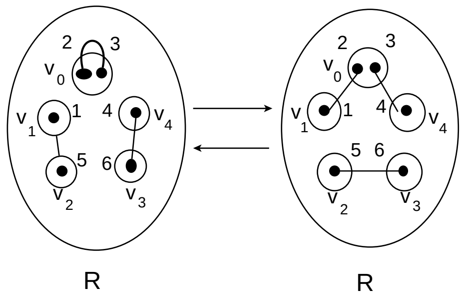

Figure 5.2: L_1-*switching*

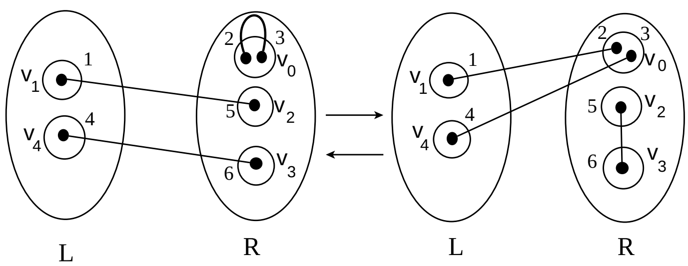

Figure 5.3: L_2-*switching*

Proof Let $\mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}$ and we consider the number of L_1-switching operations that convert $\mathcal{P}$ to some $\mathcal{P}' \in \mathcal{C}_{l_0-1,l_1,l_2}$. For the purpose of counting, we label the points in the pairs that are under consideration as shown in Figure 5.2. So for any pair under consideration, for instance, a given loop or a given double pair, we count how many ways we can label the points in the pair. Let N denote the number of ways to choose the pairs and label the points in them so that an L_1-switching can be applied to these pairs, which converts $\mathcal{P}$ to some $\mathcal{P}' \in \mathcal{C}_{l_0-1,l_1,l_2}$ without any simultaneously created loops or double pairs. This implies that the switching operations counted by N destroy only one loop and there is no simultaneous creation or destruction of other loops or double pairs.

We first give a rough count of N, that includes some forbidden cases (due to creating double pairs, etc) and then estimate the error. There are l_0 ways to choose a loop e_0 and $t(t-1)$ ways to choose (e_1, e_2), an ordered pair of two distinct pure pairs. For any chosen loop e_0, there are two ways to distinguish the two end points to label the points 2 and

3 as shown in Figure 5.2. For each of the other pairs, there are two ways to label its two endpoints, as 1 and 5, or 4 and 6, as the case may be. Hence a rough estimation of N is $8l_0t(t-1)$, including the count of some forbidden choices of e_0, e_1 and e_2, which we estimate next. Let the vertices that contain points $2, 1, 5, 6, 4$ be denoted by v_0, v_1, v_2, v_3, v_4 respectively as shown in Figure 5.2. The only possible exclusions caused by invalid choices in the above are the following:

(a) the loop e_0 is adjacent to e_1 or e_2, or e_1 is adjacent to e_2, in which case, the L_1-switching is not applicable since the definition of the L_1-switching excludes cases where the edges are adjacent because it requires the end vertices to be distinct;

(b) there exists a pair between $\{v_0, v_1\}$, or $\{v_0, v_4\}$, or $\{v_2, v_3\}$ in $\mathcal{P}$, in which case there will be more double pairs created after the L_1-switching is applied;

(c) the pair e_1 or e_2 is a loop or is contained in a double pair, in which case there is a simultaneously destroyed loop or double pair.

First we show that the number of exclusions from case (a) is $O(l_0td_{\max})$. The number of pairs of (e_0, e_1) is at most l_0t. For any given e_0 and e_1, the number of ways to choose a pair e_2 such that e_2 is adjacent to e_0 or e_1 is at most $2d_{\max}$ since both e_0 and e_2 are adjacent to at most $d_{\max}$ pairs. Hence the number of triples of (e_0, e_1, e_2) such that e_2 is adjacent to either e_0 or e_1 is at most $2l_0td_{\max}$. By symmetry, the number of triples of (e_0, e_1, e_2) such that e_1 is adjacent to either e_0 or e_2 is also at most $2l_0td_{\max}$. Hence the number of exclusions from case (a) is $O(l_0td_{\max})$.

Next we show that the number of exclusions from case (b) is $O(l_0td_{\max}^2)$. As just explained, the number of pairs of (e_0, e_1) is at most l_0t. For any given e_0 and e_1, the number of ways to choose a pair e_2 such that v_3 is adjacent to v_2 or v_4 is adjacent to v_0 is at most $2d_{\max}^2$, since both e_0 and e_1 have at most $d_{\max}^2$ edges that are of distance 2 away. Hence the number of triples (e_0, e_1, e_2) such that v_3 is adjacent to v_2 or v_4 is adjacent to v_0 is $O(l_0td_{\max}^2)$. Again symmetry, the number of triples (e_0, e_1, e_2) such that v_3 is adjacent to v_2 or v_0 is adjacent to v_1 is $O(l_0td_{\max}^2)$. Hence the number of exclusions from case (b) is $O(l_0td_{\max}^2)$.

Now we show that the number of exclusions from case (c) is $O(l_0^2t + l_0tl_2)$. The number of ways to choose e_0, e_1, e_2 such that e_1 or e_2 is a loop is at most $2l_0^2t$ and the number of ways to choose these three pairs such that e_1 or e_2 is contained in a double pair at most $2 \cdot l_0t \cdot 2l_2 = O(l_0tl_2)$. Hence the number of exclusions from case (c) is $O(l_0^2t + l_0tl_2)$.

Thus, the number of exclusions in the calculation of N is $O(l_0td_{\max}^2 + l_0^2t + l_0tl_2)$. Hence $N = 8l_0t^2(1 + O(d_{\max}^2/t + (l_0 + l_2)/t))$.

Now choose an arbitrary pairing $\mathcal{P}' \in \mathcal{C}_{l_0-1,l_1,l_2}$. Let N' be the number of ways to choose the pairs and label points in them so that an inverse L_1-switching operation can be applied to these pairs such that $\mathcal{P}'$ is converted to some $\mathcal{P} \in \mathcal{C}_{l_0,t_1,t_2}$ without any simultaneously destroyed loops or double pairs. To apply this operation we need to choose e'_0, e'_1, e'_2, such that (e'_0, e'_1) is a simple directed 2-path of type 1 and e'_2 is a pure pair. We consider the directed 2-path (e'_0, e'_1) because it automatically gives a unique way of distinguishing vertices v_1, v_0 and v_4 and labelling points as 1, 2, 3 and 4 in Figure 5.2. There are $A_1(\mathcal{P}')$ simple directed 2-paths of type 1, and hence $A_1(\mathcal{P}')$ ways to choose the points as 1, 2, 3 and 4. The number of ways to choose a pure pair e'_2 is t and so there are $2t$ ways to fix the vertices v_2, v_3 and the points $\{5, 6\}$. The only possible exclusions to the above choices are listed the following cases.

(a) There exists a pair between $\{v_1, v_2\}$ or $\{v_3, v_4\}$ in $\mathcal{P}'$, since then more double pairs will be created if the inverse L_1-switching is applied.

(b) The pair e'_2 is a loop, in which case the inverse L_1-switching is not applicable, or e'_2 is contained in a double pair, in which case a double pair is destroyed after the application of the inverse L_1-switching.

(c) The pair e'_2 is adjacent to the 2-path or is contained in the 2-path, in which case the inverse L_1-switching operation is not applicable.

The number of exclusions from case (a) is $O(A_1(\mathcal{P}')d_{\max}^2)$ and the numbers of exclusions from case (b) and (c) are $O(A_1(\mathcal{P}')l_0 + A_1(\mathcal{P}')l_2)$ and $O(A_1(\mathcal{P}')d_{\max})$ respectively.

Thus, the number of exclusions from case (a)–(d) is $O(A_1(\mathcal{P}')d_{\max}^2 + A_1(\mathcal{P}')l_0 + A_1(\mathcal{P}')l_2)$. So

$$\mathbf{E}(N') = \mathbf{E}\big(2A_1 t(1+O(d_{\max}^2/t+(l_0+l_2)/t)) \mid \mathcal{P}' \in \mathcal{C}_{l_0-1,l_1,l_2}\big) = 2a_1 t(1+O(d_{\max}^2/t+(l_0+l_2)/t)).$$

We count the pairs of $(\mathcal{P}, \mathcal{P}')$ such that $\mathcal{P} \in \mathcal{C}_{l_0,t_1,t_2}$, $\mathcal{P}' \in \mathcal{C}_{l_0-1,l_1,l_2}$, and $\mathcal{P}'$ is obtained by applying an L_1-switching to $\mathcal{P}$, which destroys only one loop without any simultaneously created loops or double pairs. Then the number of such pairs of pairings equals to both $|\mathcal{C}_{l_0,l_1,l_2}|\mathbf{E}(N)$ and $|\mathcal{C}_{l_0-1,l_1,l_2}|\mathbf{E}(N')$. Thus,

$$\frac{|\mathcal{C}_{l_0,l_1,l_2}|}{|\mathcal{C}_{l_0-1,l_1,l_2}|} = \frac{a_1}{4l_0 t}(1 + O(d_{\max}^2/t + (l_0 + l_2)/t)).$$

This proves part (i) of Lemma 5.3.6. Analogously we can deduce the following by analysing the L_2-switching and its inverse.

$$\begin{aligned}\frac{|\mathcal{C}_{l_0,l_1,l_2}|}{|\mathcal{C}_{l_0-1,l_1,l_2}|} &= \frac{2ta_3 + O(d_{\max}^3 t) + O(l_0 a_3 + l_2 a_3)}{2l_0 M_1(L)^2 + O(d_{\max}^2 M_1(L) l_0 + l_0 M_1(L) l_1)} \\ &= \frac{ta_3}{l_0 M_1(L)^2}(1 + O(d_{\max}^2/M_1(L) + d_{\max}^3/c + (l_0 + l_2)/t + l_1/M_1(L))).\end{aligned}$$

Then we obtain part (ii) of Lemma 5.3.6. ∎

We use the following two switching operations to prove Lemma 5.3.7.

D_1-switching: take a double pair $\{\{3,4\},\{5,6\}\}$ that are of type 1 and also two pure pairs $\{1,2\}$ and $\{7,8\}$ such that the eight points are located in the six distinct buckets as shown in Figure 5.4. Replace the four pairs by $\{1,3\},\{5,7\},\{2,4\},\{6,8\}$.

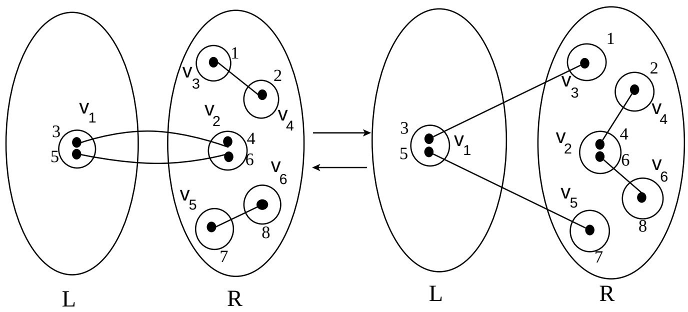

Figure 5.4: D_1*-switching*

D_2-switching: take a double pair $\{\{3,4\},\{5,6\}\}$ that are of type 1 and also two mixed pairs $\{1,2\}$ and $\{7,8\}$ such that the eight points are located in the six distinct buckets as shown in Figure 5.5. Replace the four pairs by $\{1,4\},\{6,7\},\{2,3\},\{5,8\}$.

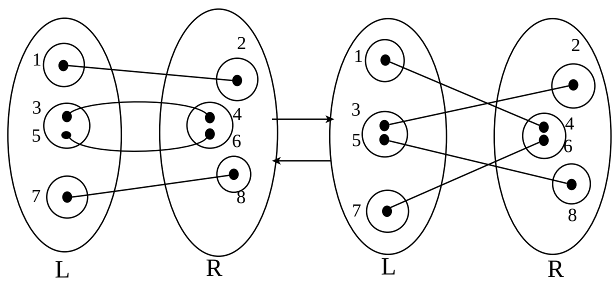

Figure 5.5: D_2*-switching*

The inverse D_i-switching for $i = 1,2$ can be defined in the same way as the inverse L_1-switching. Take $i = 1$ as an example. The inverse D_1-switching is defined as follows.

Take a directed 2-path $((1,3),(5,7))$ of type 4 and a directed 2-path $((2,4),(6,8))$ of type 1 such that the eight points are located in six distinct buckets as shown in Figure 5.4. Replace these four pairs by $\{1,2\}$, $\{3,4\}$, $\{5,6\}$ and $\{7,8\}$.

Lemma 5.3.7 *Let $a_1 = a_1(0, l_1-1, l_2)$ and $a_3 = a_3(0, l_1-1, l_2)$. Assume $l_1 \geq 1$. Then*

(i) : *If $t \geq 1$ and $a_1 \geq 1$,*
$$\frac{|\mathcal{C}_{0,l_1,l_2}|}{|\mathcal{C}_{0,l_1-1,l_2}|} = \frac{M_2(L)a_1}{8l_1t^2}(1 + O(d_{\max}^3/a_1 + d_{\max}^2/t + l_2/t + l_1 d_{\max}/M_2(L)));$$

(ii) : *If $M_1(L) \geq 1$ and $a_3 \geq 1$,*
$$\frac{|\mathcal{C}_{0,l_1,l_2}|}{|\mathcal{C}_{0,l_1-1,l_2}|} = \frac{a_3 M_2(L)}{2l_1 M_1(L)^2}(1 + O(d_{\max}^2/M_1(L) + d_{\max}^3/a_3 + l_1/M_1(L) + l_1 d_{\max}/M_2(L))).$$

Proof For a given pairing $\mathcal{P} \in \mathcal{C}_{0,l_1,l_2}$, let N be the number of ways to choose the pairs and label the points in them so that a D_1-switching can be applied to these pairs such that $\mathcal{P}$ is converted to some $\mathcal{P}' \in \mathcal{C}_{0,l_1-1,l_2}$ without simultaneously creating any loops and double pairs. In order to apply a D_1-switching operation, we need to choose a double pair $\{e_1, e_2\}$ of type 1 and an ordered pair of distinct pure pairs (e_3, e_4). The number of ways to choose $\{e_1, e_2\}$ of type 1 is l_1 in $\mathcal{C}_{0,l_1,l_2}$ and hence the number of ways to label the points as $3,4,5,6$ is $2l_1$. The number of ways to choose the ordered pair of pure pairs (e_3, e_4) is $t(t-1)$. For any chosen (e_3, e_4), there are 4 ways to label points as $1,2,7,8$. Let the vertices that contain points $3,4,1,2,7,8$ be $v_1, v_2, v_3, v_4, v_5, v_6$ as shown in Figure 5.4. Hence a rough count of N is $8l_1t(t-1)$ including the count of a few forbidden choices of e_1, e_2, e_3, e_4, which are listed as follows.

(a) The pair e_1 is adjacent to e_3 or e_4, or e_3 is adjacent to e_4, in which case the D_1-switching is not applicable.

(b) There exists a pair between $\{v_1, v_3\}$, or $\{v_2, v_4\}$, or $\{v_2, v_6\}$, or $\{v_1, v_5\}$ in $\mathcal{P}$, since another double pair will be created after the D_1-switching is applied.

(c) The pair e_3 or e_4 is contained in a double pair, since another double pair is destroyed after the D_1-switching is applied.

The numbers of forbidden choices of e_1, e_2, e_3, e_4 coming from case (a), (b) and (c) are $O(l_1 t d_{\max})$, $O(l_1 t d_{\max}^2)$ and $O(l_1 t l_2)$ respectively. So $N = 8l_1t^2(1 + O(d_{\max}^2/t + l_2/t))$.

For a given pairing $\mathcal{P}' \in \mathcal{C}_{0,l_1-1,l_2}$, let N' be the number of ways to choose the pairs and label the points in them so that an inverse D_1-switching operation can be applied to these pairs which converts $\mathcal{P}'$ to some $\mathcal{P} \in \mathcal{C}_{0,l_1,l_2}$ without destroying any loops or double pairs simultaneously. In order to apply such an operation, we need to choose two simple directed

2-paths, one of type 1 and the other of type 4. There are $A_1(\mathcal{P}')$ simple directed 2-paths of type 1, each of which gives a way of labelling points as $2, 4, 6, 8$, and there are $A_4(\mathcal{P}')$ simple directed 2-paths of type 4, each of which gives a way of labelling points as $1, 3, 5, 7$. Hence a rough count of N' is $A_1(\mathcal{P}')A_4(\mathcal{P}')$ including the counts of a few forbidden choices of such two 2-paths which are listed in the following two cases.

(a) If we have $v_i = v_j$, for $i \in \{3, 5\}$ and $j \in \{2, 4, 6\}$, then the operation is not applicable.

(b) If there already exists a pair between $\{v_1, v_2\}$, or $\{v_3, v_4\}$, or $\{v_5, v_6\}$ in $\mathcal{P}'$, then more than one double pair will be created in this case if the inverse D_1-switching is applied.

The numbers of forbidden choices of the two directed 2-paths from case (a) and (b) are respectively $O(A_4(\mathcal{P}')d_{\max}^2) = O(M_2(L)d_{\max}^2)$ and $O(A_4(\mathcal{P}')d_{\max}^3) = O(M_2(L)d_{\max}^3)$. So $\mathbf{E}(N') = \mathbf{E}\big(A_1(\mathcal{P}')A_4(\mathcal{P}') \mid \mathcal{P}' \in \mathcal{C}_{0,l_1-1,l_2}\big) + O(M_2(L)d_{\max}^3) = a_1(M_2(L) - O(l_1 d_{\max}))(1 + O(d_{\max}^3/a_1))$. Since $l_1 \geq 1$, we have $M_2(L) \geq 1$. Hence

$$\begin{aligned} \frac{|\mathcal{C}_{0,l_1,l_2}|}{|\mathcal{C}_{0,l_1-1,l_2}|} &= \frac{a_1 M_2(L)(1 + O(d_{\max}^3/a_1) + O(l_1 d_{\max}/M_2(L)))}{8 l_1 t^2 (1 + O(d_{\max}^2/t) + O(l_2/t))} \\ &= \frac{a_1 M_2(L)}{8 l_1 t^2}(1 + O(d_{\max}^3/a_1 + d_{\max}^2/t + l_2/t + l_1 d_{\max}/M_2(L))), \end{aligned}$$

and this shows part (i) of Lemma 5.3.7. Similarly we can obtain part (ii) by analysing the D_2-switching and its inverse. ∎

The following two switching operations are used to prove Lemma 5.3.8.

D_3-switching: take a double pair $\{\{1, 2\}, \{3, 4\}\}$ that are of type 2 and also two pure pairs $\{5, 6\}$ and $\{7, 8\}$ such that the eight points are located in the six distinct buckets as shown in Figure 5.6. Replace the four pairs by $\{1, 5\}, \{2, 6\}, \{3, 7\}, \{4, 8\}$.

D_4-switching: take a double pair $\{\{1, 2\}, \{3, 4\}\}$ that are of type 2 and also four mixed pairs $\{5, 6\}$, $\{7, 8\}$, $\{9, 10\}$, $\{11, 12\}$ such that the twelve points are located in the ten distinct buckets as shown in Figure 5.7. Replace the six pairs by $\{6, 10\}$, $\{8, 12\}$, $\{1, 5\}$, $\{3, 9\}$,$\{2, 11\}$, $\{4, 7\}$.

The inverse D_3-switching is be defined as follows whereas the inverse D_4-switching can be defined in a similar way. Take two directed paths of type 1 $((5, 1), (3, 7))$ and $((6, 2), (4, 8))$ such that the eight points are located in six distinct buckets as shown in Figure 5.6. Replace these four pairs by $\{5, 6\}$, $\{1, 2\}$, $\{3, 4\}$, $\{7, 8\}$.

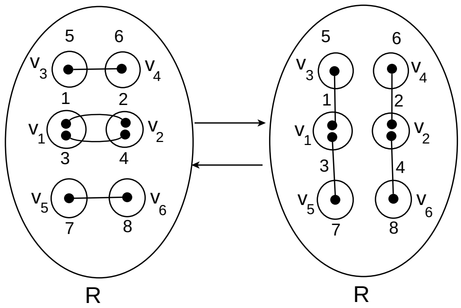

Figure 5.6: *D_3-switching*

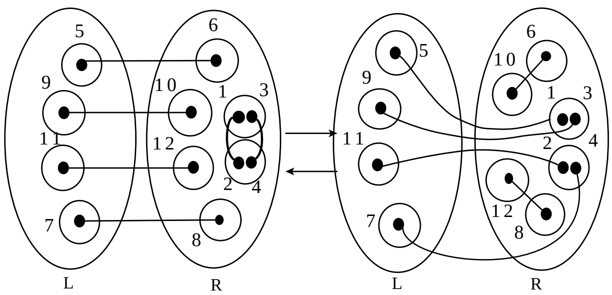

Figure 5.7: *D_4-switching*

Lemma 5.3.8 *Let* $b_1 = \mathbf{E}(A_1(\mathcal{P})^2 \mid \mathcal{P} \in \mathcal{C}_{0,0,l_2-1})$ *and* $b_3 = \mathbf{E}(A_3(\mathcal{P})^2 \mid \mathcal{P} \in \mathcal{C}_{0,0,l_2-1})$. *Assume* $l_2 \geq 1$. *Then*

$$\begin{aligned}(i): \quad & \textit{If } t \geq 1 \textit{ and } b_1 \geq 1, \\ & \frac{|\mathcal{C}_{0,0,l_2}|}{|\mathcal{C}_{0,0,l_2-1}|} = \frac{b_1}{16 l_2 t^2}(1 + O(d_{\max}^2/t + d_{\max}^3 a_1/b_1 + l_2/t)). \\ (ii): \quad & \textit{If } M_1(L) \geq 1 \textit{ and } b_3 \geq 1, \\ & \frac{|\mathcal{C}_{0,0,l_2}|}{|\mathcal{C}_{0,0,l_2-1}|} = \frac{t^2 b_3}{l_2 M_1(L)^4}(1 + O(d_{\max}^3 a_3/b_3 + d_{\max}^2/M_1(L) + l_2/t)).\end{aligned}$$

Proof For a given pairing $\mathcal{P} \in \mathcal{C}_{0,0,l_2}$, let N be the number of ways to choose the pairs and label the points in them so that a D_3-switching operation can be applied, which converts

$\mathcal{P}$ to some $\mathcal{P}' \in \mathcal{C}_{0,0,l_2-1}$ without creating any loops and double pairs simultaneously. In order to apply a D_3-switching operation, we need to choose a double pair $\{e_1, e_2\}$ of type 2 and an ordered pair of distinct pure pairs (e_3, e_4). The number of ways to choose $\{e_1, e_2\}$ is l_2 in $\mathcal{C}_{0,0,l_2}$ and there are four ways to label the points as $1, 2, 3, 4$ for any chosen double pair. The number of ways to choose an ordered pair of two pure pairs (e_3, e_4) is $t(t-1)$ and hence the number of ways to label the points as $5, 6, 7, 8$ is $4t(t-1)$. Hence a rough count of N is $16l_2t(t-1)$ including the counts of forbidden choices of pairs $e_1, \ldots, e_4$ which we estimate next. Let the vertices that contain points $1, 2, 5, 6, 7, 8$ be $v_1, v_2, v_3, v_4, v_5, v_6$ as shown in Figure 5.6. The forbidden choices of the pairs $e_1, \ldots, e_4$ are listed in the following three cases.

(a) When e_1 is adjacent to e_3 or e_4 or when e_3 is adjacent to e_4, then the D_3-switching is not applicable.

(b) If there exists a pair between $\{v_1, v_3\}$, or $\{v_2, v_4\}$, or $\{v_1, v_5\}$, or $\{v_2, v_6\}$ in $\mathcal{P}$, then more double pairs will be created after the application of the switching operation.

(c) If e_3 or e_4 is contained in a double pair, then another double pair would be destroyed after the application of the switching operation.

The numbers of forbidden choices of $e_1, \ldots, e_4$ coming from (a),(b) and (c) are $O(l_2td_{\max})$, $O(l_2td_{\max}^2)$ and $O(l_2^2t)$. So $N = 16l_2t^2(1 + O(d_{\max}^2/t + l_2/t))$.

For any pairing $\mathcal{P}' \in \mathcal{C}_{0,0,l_2-1}$, let N' be the number of ways to choose the pairs and label the points in them so that an inverse D_3-switching can be applied to these pairs, which converts $\mathcal{P}'$ to some $\mathcal{P} \in \mathcal{C}_{0,0,l_2}$ without simultaneously destroying any loops or double pairs. In order to apply such an operation, we need to choose an ordered pair of distinct simple directed 2-paths of type 1. The number of ways to do that is $A_1(\mathcal{P}')(A_1(\mathcal{P}') - 1)$. So the number of ways to label the points $1, 2, \ldots, 8$ is $A_1(\mathcal{P}')(A_1(\mathcal{P}') - 1)$, which gives a rough count of N'. The forbidden choices of the two paths whose counts should be excluded from N' are listed in the following cases.

(a) The two paths shares some common vertex or common pair. In this case the inverse D_3-switching is not applicable.

(b) There exists a pair between $\{v_1, v_2\}$ or $\{v_3, v_4\}$ or $\{v_5, v_6\}$ in $\mathcal{P}'$. In this case, more double pairs will be created after the inverse D_3-switching operation is applied.

The numbers of ways to choose the ordered pair of 2-paths in case (a) and (b) are $O(A_1(\mathcal{P}')d_{\max}^2)$ and $O(A_1(\mathcal{P}')d_{\max}^3)$ respectively. Thus, $\mathbf{E}(N') = b_1(1 + O(d_{\max}^3 a_1/b_1))$.

Hence

$$\frac{|\mathcal{C}_{0,0,l_2}|}{|\mathcal{C}_{0,0,l_2-1}|} = \frac{b_1}{16l_2t^2}(1 + O(d_{\max}^2/t + d_{\max}^3 a_1/b_1 + (l_0 + l_2)/t)).$$

Similarly by analysing the D_4-switching and its inverse, we obtain Lemma 5.3.8 (ii). ∎

Next we estimate $a_i(l_0, l_1, l_2)$ and $b_i(l_0, l_1, l_2)$ for $i = 1, 3$, which appeared in the results of Lemmas 5.3.6–5.3.8. The following two switchings are used to estimates $a_i(l_0, l_1, l_2)$ for $i = 1, 2, 3$.

S_1-switching. Take a mixed pair and label the points in it by $\{1, 2\}$ as shown in Figure 5.8. Take a simple directed 2-path that is vertex disjoint from the chosen mixed pair. Label the points by $3, 4, 5, 6$. Replace these three pairs by $\{2, 3\}$, $\{1, 4\}$ and $\{5, 6\}$. The inverse S_1-switching is defined as follows. Take a pure pair $\{2, 3\}$ and a simple directed 2-path $((1, 4), (5, 6))$ such that the six points are located in five distinct buckets shown as in Figure 5.8. Replace these three pairs by $\{1, 2\}$, $\{3, 4\}$ and $\{5, 6\}$.

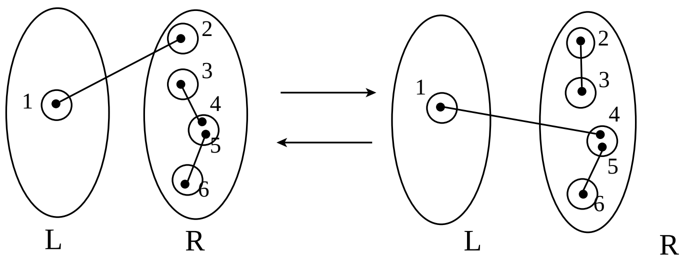

Figure 5.8: *S_1-switching*

S_2-switching. Take a pure pair $\{5, 6\}$ and a simple directed 2-path $((1, 2), (3, 4))$ such that the six points are located in five distinct buckets shown as in Figure 5.9. Replace these three pairs by $\{1, 2\}$, $\{3, 5\}$ and $\{4, 6\}$. The inverse S_2-switching is defined as follows. Take a mixed pair $\{4, 6\}$ and a simple directed 2-path $((1, 2), (3, 5))$ such that the six points are located in five distinct buckets shown as in Figure 5.8. Replace these three pairs by $\{1, 2\}$, $\{3, 4\}$ and $\{5, 6\}$.

Lemma 5.3.9 *Given l_0, l_1 and l_2, let $\ell = l_0 + l_1 + l_2$. Then*

$$\begin{aligned}
&(i): \ \textit{if } M_1(L) \le M/4, \\
&\quad a_1(l_0, l_1, l_2) = \frac{(M_1(R) - M_1(L))^2 M_2(R)}{M_1(R)^2}(1 + O(d_{\max}^2/t + \ell/t + (\ell d_{\max} + l_0 d_{\max}^2)/M_2(R))); \\
&(ii): \ \textit{if } M_1(L) > M/4, \\
&\quad a_3(l_0, l_1, l_2) = \frac{M_1(L)^2 M_2(R)}{M_1(R)^2}(1 + O(d_{\max}^2/M_1(L) + \ell/M_1(L) + (\ell d_{\max} + l_0 d_{\max}^2)/M_2(R))).
\end{aligned}$$

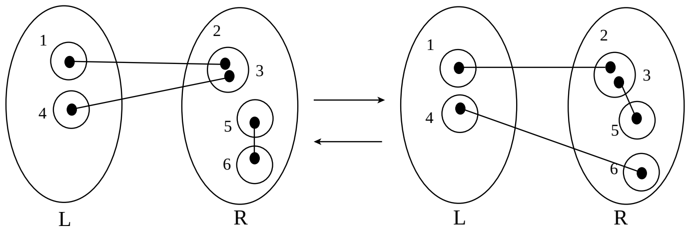

Figure 5.9: S_2-*switching*

Proof Let $a_i = a_i(l_0, l_1, l_2)$ for $i = 1, 2, 3$. We use the S_1-switching to compute the ratio a_1/a_2 and the S_2-switching to compute the ratio a_3/a_2. We count the ordered pairs of pairings $(\mathcal{P}, \mathcal{P}')$ such that both $\mathcal{P}$ and $\mathcal{P}'$ are from $\mathcal{C}_{l_0,l_1,l_2}$, and $\mathcal{P}'$ is obtained from $\mathcal{P}$ by applying an S_1-switching to $\mathcal{P}$ without any creation or destruction of loops or double pairs. Let N_1 denote the number of such ordered pairs of pairings.

We first prove part (i). Assume $M_1(L) \le M/4$. For any directed 2-path of type 1 in $\mathcal{C}_{l_0,l_1,l_2}$, the number of S_1-switching operations that can be applied to it is

$$A_1 M_1(L) + O(A_1 d_{\max}^2 + A_1 l_1) = A_1 M_1(L)\Big(1 + O(d_{\max}^2/M_1(L) + l_1/M_1(L))\Big), \quad (5.3.2)$$

For any directed 2-path of type 2 in $\mathcal{C}_{l_0,l_1,l_2}$, the number of inverse S_1-switching operations that can be applied to it is

$$A_2 \cdot 2t + O(A_2 d_{\max}^2 + A_2(l_0 + l_2)) = A_2 \cdot 2t\Big(1 + O(d_{\max}^2/t + (l_0 + l_2)/t)\Big). \quad (5.3.3)$$

The total number of S_1-switching operations that can be applied to pairings in $\mathcal{C}_{l_0,l_1,l_2}$ is

$$\sum_{\mathcal{P}\in\mathcal{C}_{l_0,l_1,l_2}} A_1(\mathcal{P}) M_1(L)\Big(1 + O((d_{\max}^2 + l_1)/M_1(L))\Big) = a_1 M_1(L)\Big(1 + O((d_{\max}^2 + \ell)/M_1(L))\Big)|\mathcal{C}_{l_0,l_1,l_2}|,$$

and the total number of inverse S_1-switching operations that can be applied to pairings in $\mathcal{C}_{l_0,l_1,l_2}$ is

$$\sum_{\mathcal{P}\in\mathcal{C}_{l_0,l_1,l_2}} A_2(\mathcal{P}) \cdot 2t\Big(1 + O(d_{\max}^2/t + (l_0 + l_2)/t)\Big) = a_2 \cdot 2t\Big(1 + O(d_{\max}^2/t + \ell/t)\Big)|\mathcal{C}_{l_0,l_1,l_2}|.$$

These two numbers are both equal to N_1. Hence

$$\frac{a_2}{a_1} = \frac{M_1(L)}{2t}(1 + O(d_{\max}^2/t + d_{\max}^2/M_1(L) + \ell/M_1(L) + \ell/t)). \quad (5.3.4)$$

Similarly, by the S_2-switching and its inverse we get

$$\frac{a_3}{a_2} = \frac{M_1(L)}{2t}(1 + O(d_{\max}^2/t + d_{\max}^2/M_1(L) + \ell/M_1(L) + \ell/t)). \quad (5.3.5)$$

Then (5.3.4) gives

$$\frac{a_2}{a_1} = \frac{M_1(L)}{2t}\Big(1 + O((d_{\max}^2 + \ell)/t)\Big) + O((d_{\max}^2 + \ell)/t),$$

and (5.3.5) gives

$$\frac{a_3}{a_2} = \frac{M_1(L)}{2t}\Big(1 + O((d_{\max}^2 + \ell)/t)\Big) + O((d_{\max}^2 + \ell)/t).$$

Hence

$$\begin{aligned} a_2 &= a_1\left(\frac{M_1(L)}{2t}\Big(1 + O((d_{\max}^2 + \ell)/t)\Big) + O((d_{\max}^2 + \ell)/t)\right) \\ a_3 &= a_1\left(\frac{M_1(L)}{2t}\Big(1 + O((d_{\max}^2 + \ell)/t)\Big) + O((d_{\max}^2 + \ell)/t)\right)^2. \end{aligned}$$

Since $M_1(L) \le M/4$, we have $M_1(L)/t \le 1$ and so

$$a_3 = a_1\left(\left(\frac{M_1(L)}{2t}\right)^2\Big(1 + O((d_{\max}^2 + \ell)/t)\Big) + O((d_{\max}^2 + \ell)/t)\right).$$

Hence

$$\begin{aligned} a_1 + 2a_2 + a_3 &= a_1\left(1 + \left(2\frac{M_1(L)}{2t} + \left(\frac{M_1(L)}{2t}\right)^2\right)\Big(1 + O((d_{\max}^2 + \ell)/t)\Big) + O((d_{\max}^2 + \ell)/t)\right) \\ &= a_1\left(\left(1 + \frac{M_1(L)}{2t}\right)^2\Big(1 + O((d_{\max}^2 + \ell)/t)\Big) + O((d_{\max}^2 + \ell)/t)\right) \\ &= a_1\left(1 + \frac{M_1(L)}{2t}\right)^2(1 + O((d_{\max}^2 + \ell)/t)). \quad (5.3.6) \end{aligned}$$

For any pairing $\mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}$, the number of simple directed 2-paths in $\mathcal{P}$ is $\sum_{v\in L\cup R} d(v)(d(v)-1) - O(\ell d_{\max} + l_0 d_{\max}^2)$, since the number of non-simple directed 2-path is bounded by

$O(l_0 d_{\max}^2 + l_1 d_{\max} + l_2 d_{\max}) = O(\ell d_{\max} + l_0 d_{\max}^2)$. On the other hand, the number of simple directed 2-paths in $\mathcal{P}$ is $A_1 + 2A_2 + A_3 + A_4$, since $2A_2$ counts the number of directed 2-paths of type 2 and the opposite direction. Then

$$a_1 + 2a_2 + a_3 + M_2(L) - O(l_1 d_{\max}) = \sum_{v \in L \cup R} d(v)(d(v)-1) - O(\ell d_{\max} + l_0 d_{\max}^2).$$

Thus, $a_1 + 2a_2 + a_3 = M_2(R) + O(\ell d_{\max} + l_0 d_{\max}^2) = M_2(R)(1 + O((\ell d_{\max} + l_0 d_{\max}^2)/M_2(R)))$. Combining this with the equation (5.3.6), we have

$$a_1 = \frac{(M_1(R) - M_1(L))^2 M_2(R)}{M_1(R)^2}(1 + O(d_{\max}^2/t + \ell/t + (\ell d_{\max} + l_0 d_{\max}^2)/M_2(R))),$$

which proves part (i).

Next we show part (ii). Assume $M_1(L) > M/4$. We observe that (5.3.4) also gives

$$\frac{a_1}{a_2} = \frac{2t}{M_1(L)}\Big(1 + O((d_{\max}^2 + \ell)/M_1(L))\Big) + O((d_{\max}^2 + \ell)/M_1(L)),$$

and (5.3.5) gives

$$\frac{a_2}{a_3} = \frac{2t}{M_1(L)}\Big(1 + O((d_{\max}^2 + \ell)/M_1(L))\Big) + O((d_{\max}^2 + \ell)/M_1(L)).$$

Thus,

$$\begin{aligned} a_2 &= a_3\left(\frac{2t}{M_1(L)}\Big(1 + O((d_{\max}^2 + \ell)/M_1(L))\Big) + O((d_{\max}^2 + \ell)/M_1(L))\right) \\ a_1 &= a_3\left(\frac{2t}{M_1(L)}\Big(1 + O((d_{\max}^2 + \ell)/M_1(L))\Big) + O((d_{\max}^2 + \ell)/M_1(L))\right). \end{aligned}$$

Since $M_1(L) \geq 1/4$, we have $t/M_1(L) < 1$ and so

$$\begin{aligned} a_1 + 2a_2 + a_3 &= a_3\left(1 + \frac{2t}{M_1(L)}\right)^2 \Big(1 + O((d_{\max}^2 + \ell)/M_1(L))\Big) \\ &= M_2(R)(1 + O((\ell d_{\max} + l_0 d_{\max}^2)/M_2(R))). \end{aligned}$$

Hence

$$a_3 = \frac{M_1(L)^2 M_2(R)}{M_1(R)^2}(1 + O(d_{\max}^2/M_1(L) + \ell/M_1(L) + (\ell d_{\max} + l_0 d_{\max}^2)/M_2(R))).$$

This proves part (ii) of the lemma. ∎

In the next lemma we estimate $b_i(l_0, l_1, l_2)$ using some switching operations that are simple modifications of the S_i-switchings for $i = 1, 3$.

Lemma 5.3.10 *Let $b_1 = \mathbf{E}(A_1(\mathcal{P})^2 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})$, $b_3 = \mathbf{E}(A_3(\mathcal{P})^2 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})$, $a_1 = a_1(l_0,l_1,l_2)$ and $a_3 = a_3(l_0,l_1,l_2)$. Let $\ell = l_0 + l_1 + l_2$. Then*

$$
\begin{aligned}
(i):&\quad \text{if } M_1(L) \le M/4,\quad b_1 = a_1^2(1 + O(d_{\max}^2/t + \ell/t + (\ell d_{\max} + l_0 d_{\max}^2 + d_{\max}^3)/M_2(R)));\\
(ii):&\quad \text{if } M_1(L) > M/4,\\
&\quad b_3 = a_3^2(1 + O(d_{\max}^2/M_1(L) + \ell/M_1(L) + (\ell d_{\max} + l_0 d_{\max}^2 + d_{\max}^3)/M_2(R))).
\end{aligned}
$$

Proof Let $X_1(\mathcal{P})$ denote the number of ordered pairs of vertex disjoint simple 2-paths of type 1 in $\mathcal{P}$, $X_2(\mathcal{P})$ the number of ordered pairs of vertex disjoint simple 2-paths of type 3 in $\mathcal{P}$, $X_3(\mathcal{P})$ the number of pairs of vertex disjoint simple 2-paths of type 1 and type 2 respectively, $X_4(\mathcal{P})$ the number of pairs of vertex disjoint simple 2-paths of type 1 and type 3 respectively, and $X_5(\mathcal{P})$ the number of pairs of vertex disjoint simple 2-paths of type 2 and type 3 respectively.

The S_3-switching and its inverse, as illustrated in Figure 5.10, is a slight modification of the S_1-switching and its inverse. To apply the s_3-switching operation, we need to choose a mixed pair and two simple 2-paths of type 1 such that they are pairwise disjoint. To apply its inverse, we need to choose a pure pair and two simple 2-paths of type 2 and 1 respectively such that they are pairwise disjoint. Compared with the S_1-switching, it is required to take an extra simple directed 2-path of type 1 in order to apply an S_3-switching. However, the pairs in the extra 2-path remain after the S_3-switching is applied. Actually, the only extra restriction to the S_3-switching compared with the S_1-switching is that the mixed pair and the other simple directed 2-path under consideration are vertex disjoint from the extra directed 2-path so that the extra 2-path is not "affected" after the application of the switching. Similarly, the S_4-switching and its inverse, as illustrated in Figure 5.11, is a modification of the S_2-switching and its inverse.

We will first estimate $\mathbf{E}(X_i(\mathcal{P}) \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})$ for $i \in [5]$ and then using this to estimate b_1 and b_3. Following the analogous argument as in Lemma 5.3.9, we can estimate the ratio $\mathbf{E}(X_3(\mathcal{P}) \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})/\mathbf{E}(X_1(\mathcal{P}) \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})$ by counting the ordered pairs of pairings $(\mathcal{P}, \mathcal{P}')$ such that $\mathcal{P}, \mathcal{P}' \in \mathcal{C}_{l_0,l_1,l_2}$ and $\mathcal{P}'$ is obtained by applying an S_3-operation to $\mathcal{P}$ without any creation or destruction of loops or double pairs. The the number of such S_3-switching operations that can be applied to $\mathcal{P}$ is $X_1 M_1(L) + O(X_1 d_{\max}^2 + X_1 l_1)$. The number of such inverse S_3-operations that can be applied to $\mathcal{P}$ is $2tX_3 + O(X_3 d_{\max}^2 + X_3(l_0 + l_2))$. So the ratio $\mathbf{E}(X_3(\mathcal{P}) \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})/\mathbf{E}(X_1(\mathcal{P}) \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})$ equals exactly the right hand side of (5.3.4) and the ratio $\mathbf{E}(X_4(\mathcal{P}) \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})/\mathbf{E}(X_3(\mathcal{P}) \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})$ equals exactly the right hand side of (5.3.5).

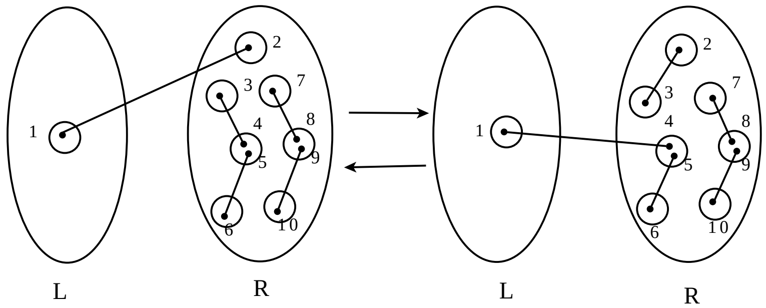

Figure 5.10: S_3-switching

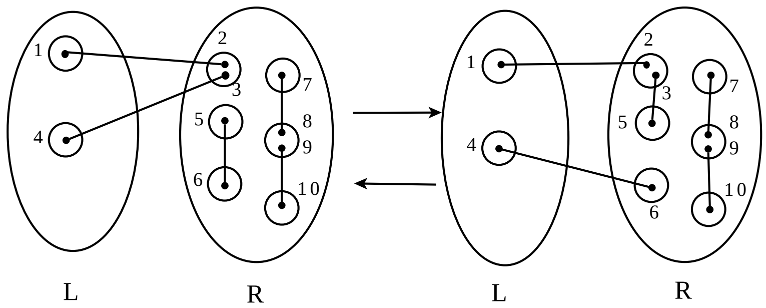

Figure 5.11: S_4-switching

On the other hand,

$$
\begin{aligned}
&\mathbf{E}(A_1^2 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) + 2\mathbf{E}(A_1A_2 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) + \mathbf{E}(A_1A_3 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) \\
&= \mathbf{E}(A_1(A_1 + 2A_2 + A_3) \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) = \mathbf{E}(A \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})M_2(R)(1 + O(\ell d_{\max}^2/M_2(R))) \\
&= a_1(l_0,l_1,l_2)M_2(R)(1 + O((\ell d_{\max} + l_0 d_{\max}^2)/M_2(R))),
\end{aligned}
$$

and

$$X_1 = A_1^2 + O(A_1 d_{\max}^3), \quad X_3 = A_1A_2 + O(A_1 d_{\max}^3), \quad X_4 = A_1A_3 + O(A_1 d_{\max}^3). \qquad (5.3.7)$$

Let $a_1 = a_1(l_0, l_1, l_2)$. Then

$$\begin{aligned}
\mathbf{E}(X_1 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) &= \mathbf{E}(A_1^2 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) + O(a_1 d_{\max}^3),\\
\mathbf{E}(X_3 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) &= \mathbf{E}(A_1A_2 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) + O(a_1 d_{\max}^3),\\
\mathbf{E}(X_4 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) &= \mathbf{E}(A_1A_3 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) + O(a_1 d_{\max}^3).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\mathbf{E}(X_1 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) + 2\mathbf{E}(X_3 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) + \mathbf{E}(X_4 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})\\
&= a_1 M_2(R)(1 + O((\ell d_{\max} + l_0 d_{\max}^2 + d_{\max}^3)/M_2(R))).
\end{aligned}$$

Then part (i) follows from the similar argument of Lemma 5.3.9 and (5.3.7). Similarly, by analysing two switching operations similar to those of S_3-switching and S_4-switching, except that the extra 2-path is of type 3, we can estimate the ratio $\mathbf{E}(X_5 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})/\mathbf{E}(X_4 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})$ and $\mathbf{E}(X_2 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})/\mathbf{E}(X_4 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})$. By the fact that

$$X_5 = A_2A_3 + O(A_3 d_{\max}^3), \quad X_4 = A_1A_3 + O(A_3 d_{\max}^3), \quad X_2 = A_3^2 + O(A_3 d_{\max}^3),$$

and

$$\begin{aligned}
&\mathbf{E}(X_2 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) + 2\mathbf{E}(X_5 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2}) + \mathbf{E}(X_4 \mid \mathcal{P} \in \mathcal{C}_{l_0,l_1,l_2})\\
&= a_3(l_0, l_1, l_2) M_2(R)(1 + O((\ell d_{\max} + l_0 d_{\max}^2 + d_{\max}^3)/M_2(R))),
\end{aligned}$$

together with Lemma 5.3.9 (ii), part (ii) follows from the similar argument as in part (i) and the proof of Lemma 5.3.9 (ii). ∎

The following lemma deals with some error terms that appeared in the previous lemmas.

Lemma 5.3.11 *Assume $d_{\max}^3 = o(M_2(R))$, $d_{\max}^4 = o(M)$. Let $w(n) = \sqrt{M}/d_{\max}^2$. Then for any $l_i \le w(n)\mu_i$, we have*

$$\begin{aligned}
&\ell/M = O(d_{\max}^2 w(n)/M) = o(1), \quad (\ell d_{\max} + l_0 d_{\max}^2)/M_2(R) = O(d_{\max}^2 w(n)/M) = o(1),\\
&(l_0 + l_2)/t = O(d_{\max}^2 w(n)/M) = o(1), \quad l_1 d_{\max}/M_2(L) = O(d_{\max}^2 w(n)/M) = o(1).
\end{aligned}$$

Proof Since $M_2(R) = O(d_{\max} M)$, $M_2(L) = O(d_{\max} M)$ and $M_1(R) = \Theta(M)$, by the definition of μ_i, we have

$$\mu_i = O(d_{\max} M_2(R)/M) = O(d_{\max}^2), \quad \text{for } 1 \le i \le 3. \tag{5.3.8}$$

It follows from (5.3.8) that $\ell/M = O(d_{\max}^2)w(n)/M = O(d_{\max}^2 \ln n/M)$. By (5.2.1)–(5.2.3), we have

$$\begin{aligned}
\frac{\ell d_{\max} + l_0 d_{\max}^2}{M_2(R)} &= O\left(\frac{d_{\max}^2 t w(n)}{M^2} + \frac{d_{\max} M_2(L) w(n)}{M^2} + \frac{d_{\max} t^2 M_2(R) w(n)}{M^4}\right) = O(d_{\max}^2 w(n)/M), \\
\frac{l_0 + l_2}{t} &= O\left(\frac{M_2(R) w(n)}{M^2} + \frac{t M_2(R)^2 w(n)}{M^4}\right) = O(d_{\max}^2 w(n)/M), \\
\frac{l_1 d_{\max}}{M_2(L)} &= O\left(\frac{d_{\max} M_2(R) w(n)}{M^2}\right) = O(d_{\max}^2 w(n)/M). \quad \blacksquare
\end{aligned}$$

The following corollary follows from Lemma 5.3.8 and Lemma 5.3.10.

Corollary 5.3.12 *Assume $d_{\max}^3 = o(M_2(R))$, $d_{\max}^4 = o(M)$. Let $a_1 = a_1(0, 0, l_2 - 1)$ and $a_3 = a_3(0, 0, l_2 - 1)$. Let $w(n) = \sqrt{M}/d_{\max}^2$. Then for any $l_i \leq w(n)\mu_i$, we have*

$$\begin{aligned}
&(i): \text{ if } M_1(L) \leq M/4, \\
&\qquad \frac{|\mathcal{C}_{0,0,l_2}|}{|\mathcal{C}_{0,0,l_2-1}|} = \frac{a_1^2}{16 l_2 t^2}(1 + O((d_{\max}^2 + l_2)/t + d_{\max}^3/M_2(R) + l_2 d_{\max}/M_2(R))), \\
&(ii): \text{ if } M_1(L) > M/4, \\
&\qquad \frac{|\mathcal{C}_{0,0,l_2}|}{|\mathcal{C}_{0,0,l_2-1}|} = \frac{t^2 a_3^2}{l_2 M_1(L)^4}(1 + O(d_{\max}^3/M_2(R) + (d_{\max}^2 + l_2)/M_1(L) + l_2 d_{\max}/M_2(R))).
\end{aligned}$$

Proof If $M_1(L) < M/4$, then $b_1 = a_1^2(1 + O(d_{\max}^2/t + l_2/t + (l_2 d_{\max} + d_{\max}^3)/M_2(R)))$ by Lemma 5.3.10 (i). By Lemma 5.3.9 (i) and Lemma 5.3.11, $a_1 = \Theta(M_2(R))(1 + O(d_{\max}^2/t + l_2/t + l_2 d_{\max}/M_2(R))) = \Theta(M_2(R))$. Since $d_{\max}^3 = o(M_2(R))$, $b_1 = \Omega(1)$. Then the error term $d_{\max}^3 a_1/b_1$ in Lemma 5.3.8 (i) can be bounded by

$$\begin{aligned}
&\frac{d_{\max}^3}{a}(1 + O(d_{\max}^2/t + \ell/t + (l_2 d_{\max} + d_{\max}^3)/M_2(R))) \\
&= O(d_{\max}^3/M_2(R)) + O(d_{\max}^3/M_2(R)) \cdot O(d_{\max}^2/t + l_2/t + (l_2 d_{\max} + d_{\max}^3)/M_2(R)) \\
&= O(d_{\max}^3/M_2(R)) + o(d_{\max}^2/t + l_2/t + (l_2 d_{\max} + d_{\max}^3)/M_2(R)).
\end{aligned}$$

Then part (i) follows by combining this with Lemma 5.3.8 (i) and Lemma 5.3.10 (i). Similarly we can show part (ii) from Lemma 5.3.8 (ii), Lemma 5.3.10 (ii) and Lemma 5.3.11, following the same routine. $\blacksquare$

Lemma 5.3.13 *Assume $d_{\max}^3 = o(M_2(R))$, $d_{\max}^4 = o(M)$. Let $w(n) = \sqrt{M}/d_{\max}^2$. Then for any $l_i \leq w(n)\mu_i$, we have*

$$\begin{aligned}
\frac{|\mathcal{C}_{l_0,l_1,l_2}|}{|\mathcal{C}_{l_0-1,l_1,l_2}|} &= \frac{\mu_0}{l_0}(1+O(d_{\max}^3/M_2(R)+d_{\max}^2 w(n)/M)), \quad l_0 \ge 1;\\
\frac{|\mathcal{C}_{0,l_1,l_2}|}{|\mathcal{C}_{0,l_1-1,l_2}|} &= \frac{\mu_1}{l_1}(1+O(d_{\max}^3/M_2(R)+d_{\max}^2 w(n)/M)), \quad l_1 \ge 1;\\
\frac{|\mathcal{C}_{0,0,l_2}|}{|\mathcal{C}_{0,0,l_2-1}|} &= \frac{\mu_2}{l_2}(1+O(d_{\max}^3/M_2(R)+d_{\max}^2 w(n)/M)), \quad l_2 \ge 1.
\end{aligned}$$

Proof Let $\delta = M_1(L)/M$. Then $0 \le \delta \le 1/2$ by the assumption that $M_1(R) \ge M_1(L)$. If $0 \le \delta \le 1/4$, then $t = \Theta(M)$. By Lemma 5.3.9 (i) and Lemma 5.3.11, $a_1 = \Theta(M_2(R))$. So the error term $d_{\max}^3/a_1$ is bounded by $O(d_{\max}^3/M_2(R))$. By Lemma 5.3.6 (i), Lemma 5.3.7 (i), Corollary 5.3.12 (i) and Lemma 5.3.11,

$$\begin{aligned}
\frac{|\mathcal{C}_{l_0,l_1,l_2}|}{|\mathcal{C}_{l_0-1,l_1,l_2}|} &= \frac{a_1}{4l_0t}(1+O(d_{\max}^2/t+(l_0+l_2)/t))\\
&= \frac{\mu_0}{l_0}(1+O(d_{\max}^3/M_2(R)+d_{\max}^2 w(n)/M_2(R))),\\
\frac{|\mathcal{C}_{0,l_1,l_2}|}{|\mathcal{C}_{0,l_1-1,l_2}|} &= \frac{\mu_1}{l_1}(1+O(d_{\max}^3/M_2(R)+d_{\max}^2 w(n)/M_2(R))),\\
\frac{|\mathcal{C}_{0,0,l_2}|}{|\mathcal{C}_{0,0,l_2-1}|} &= \frac{\mu_2}{l_2}(1+O(d_{\max}^3/M_2(R)+d_{\max}^2 w(n)/M_2(R))).
\end{aligned}$$

If $1/4 < \delta \le 1/2$, which implies that $M_1(L) = \Theta(M)$, then by Lemma 5.3.9 (ii), $a_3 = \Theta(M_2(R))$. So $d_{\max}^3/a_3 = O(d_{\max}^3/M_2(R))$. By Lemma 5.3.6 (ii), Lemma 5.3.7 (ii), Corollary 5.3.12 (ii) and Lemma 5.3.11,

$$\begin{aligned}
\frac{|\mathcal{C}_{l_0,l_1,l_2}|}{|\mathcal{C}_{l_0-1,l_1,l_2}|} &= \frac{ta_3}{l_0M_1(L)^2}(1+O(d_{\max}^2/M_1(L)+d_{\max}^3/c+l_1/M_1(L)+(l_0+l_2)/t))\\
&= \frac{\mu_0}{l_0}(1+O(d_{\max}^3/M_2(R))+d_{\max}^2 w(n)/M),\\
\frac{|\mathcal{C}_{0,l_1,l_2}|}{|\mathcal{C}_{0,l_1-1,l_2}|} &= \frac{\mu_1}{l_1}(1+O(d_{\max}^3/M_2(R))+d_{\max}^2 w(n)/M),\\
\frac{|\mathcal{C}_{0,0,l_2}|}{|\mathcal{C}_{0,0,l_2-1}|} &= \frac{\mu_2}{l_2}(1+O(d_{\max}^3/M_2(R))+d_{\max}^2 w(n)/M). \quad \blacksquare
\end{aligned}$$

Let $\mathbf{P}(\mathbf{d})$ be the probability that $\mathcal{P} \in \mathcal{M}(L, R, \mathbf{d})$ does not contain loops or double pairs. For any $i = 0, 1, 2$, define

$$k_i = \lfloor \min\{w(n)\mu_i, \max\{8\mu_i, \ln M\}\}\rfloor. \tag{5.3.9}$$

Lemma 5.3.14 *Assume $d_{\max}^3 = o(M_2(R))$ and $d_{\max}^4 = o(M)$. Let $w(n) = \sqrt{M}/d_{\max}^2$. Then*

$$\frac{1}{\mathbf{P}(\mathbf{d})} = \left(1 + O(d^2/\sqrt{M})\right) \sum_{l_0=0}^{k_0} \sum_{l_1=0}^{k_1} \sum_{l_2=0}^{k_2} \frac{|\mathcal{C}_{l_0,l_1,l_2}|}{|\mathcal{C}_{0,0,0}|},$$

and $\mu_i^{k_i+1}/(k_i+1)! = O(d_{\max}^2/\sqrt{M})$ for all $i = 0, 1, 2$.

Proof of Lemma 5.3.14. It follows immediately from Lemma 5.3.1, Lemma 5.3.2 and Corollary 5.3.4 that

$$\frac{1}{\mathbf{P}(\mathbf{d})} = (1 + O(d_{\max}^2/\sqrt{M})) \sum_{l_0=0}^{\lfloor w(n)\mu_0 \rfloor} \sum_{l_1=0}^{\lfloor w(n)\mu_1 \rfloor} \sum_{l_2=0}^{\lfloor w(n)\mu_2 \rfloor} \frac{|\mathcal{C}_{l_0,l_1,l_2}|}{|\mathcal{C}_{0,0,0}|}. \tag{5.3.10}$$

We first show that given any $l_1 \le w(n)\mu_1$ and $l_2 \le w(n)\mu_2$,

$$\sum_{l_0=0}^{w(n)\mu_0} |\mathcal{C}_{l_0,l_1,l_2}| = (1 + O(M^{-1})) \sum_{l_0=0}^{k_0} |\mathcal{C}_{l_0,l_1,l_2}|, \tag{5.3.11}$$

and $\mu_0^{k_0+1}/(k_0+1)!$. Note that the k_i are integers. However, in the following analysis we omit the floor functions if the result of the analysis is not affected by changing the values of the k_i by at most 1. The equality (5.3.11) is trivially true if $k_0 = w(n)\mu_0$ and it this case

$$\frac{\mu_0^{k_0+1}}{(k_0+1)!} \le \left(\frac{e\mu_0}{k_0+1}\right)^{k_0+1} \le (e/w(n))^{k_0+1} \le e/w(n) = O(d_{\max}^2/\sqrt{M}).$$

So we assume $k_0 < w(n)\mu_0$. Then $k_0 = \max\{8\mu_0, \ln M\}$. Since $k_0 < w(n)\mu_0$ and $k_0 \ge 8\mu_0$, by Lemma 5.3.13,

$$\frac{|\mathcal{C}_{l_0,l_1,l_2}|}{|\mathcal{C}_{l_0-1,l_1,l_2}|} \sim \frac{\mu_0}{l_0} < \frac{1}{2}, \quad \text{for all } k_0 \le l_0 \le w(n)\mu_0.$$

Thus,

$$\sum_{l_0=0}^{\infty} |\mathcal{C}_{l_0,l_1,l_2}| = \sum_{l_0=0}^{k_0} |\mathcal{C}_{l_0,l_1,l_2}| + \sum_{l_0=k_0+1}^{\infty} |\mathcal{C}_{l_0,l_1,l_2}| = \sum_{l_0=0}^{k_0} |\mathcal{C}_{l_0,l_1,l_2}| + O\left(|\mathcal{C}_{k_0+1,l_1,l_2}|\right) \tag{5.3.12}$$

By Lemma 5.3.13, we have

$$\frac{|\mathcal{C}_{k_0+1,l_1,l_2}|}{|\mathcal{C}_{0,l_1,l_2}|} = \frac{\mu_0^{k_0+1}}{(k_0+1)!}\Big(1 + O(k_0 d_{\max}^3/M_2(R) + k_0 d_{\max}^2 w(n)/M)\Big).$$

Clearly

$$\frac{\mu_0^{k_0+1}}{(k_0+1)!} \le \left(\frac{e\mu_0}{k_0+1}\right)^{k_0+1} \le \left(\frac{e}{8}\right)^{k_0+1} \le \left(\frac{e}{8}\right)^{\ln M} \le M^{-1} = O(d_{\max}^2/\sqrt{M}). \qquad (5.3.13)$$

If $k_0 = 8\mu_0$, then by (5.3.8),

$$k_0 d_{\max}^3/M_2(R) = O(d_{\max}M_2(R)/M) \cdot d_{\max}^3/M_2(R) = O(d_{\max}^4/M) = o(1),$$
$$k_0 d_{\max}^2 w(n)/M = O(d_{\max}^2) \cdot d_{\max}^2 w(n)/M = O(d_{\max}^4/M \cdot \sqrt{M/d^4}) = O(d_{\max}^2/\sqrt{M}) = o(1).$$

Otherwise, $k_0 = \ln M$. Since $\ln M < \mu_0 w(n)$, we have $M_2(R) = \Omega(M^2 \ln M/tw(n))$. Then

$$k_0 d_{\max}^3/M_2(R) = (\ln M)d_{\max}^3 \cdot O(tw(n)/M^2 \ln M) = O(d_{\max}^3 w(n)/M) = o(1),$$
$$k_0 d_{\max}^2 w(n)/M = (\ln M)d_{\max}^2 w(n)/M = o(1).$$

Therefore, $|\mathcal{C}_{k_0+1,l_1,l_2}| = O(M^{-1})|\mathcal{C}_{0,l_1,l_2}|$. Together with (5.3.12) we have shown that $\sum_{l_0=0}^{\lfloor w(n)\mu_0 \rfloor} |\mathcal{C}_{l_0,l_1,l_2}| = (1+O(M^{-1}))\sum_{l_0=0}^{k_0} |\mathcal{C}_{l_0,l_1,l_2}|$. Similarly we can show that $\mu_i^{k_i+1}/(k_i+1)! = O(d_{\max}^2/\sqrt{M})$ for $i = 1,2$ and

$$\sum_{l_1=0}^{\lfloor w(n)\mu_1 \rfloor} |\mathcal{C}_{0,l_1,l_2}| = (1+O(M^{-1}))\sum_{l_1=0}^{k_1} |\mathcal{C}_{0,l_1,l_2}|, \quad \text{for any } 0 \le l_2 \le w(n)\mu_2,$$
$$\sum_{l_2=0}^{\lfloor w(n)\mu_2 \rfloor} |\mathcal{C}_{0,0,l_2}| = (1+O(M^{-1}))\sum_{l_2=0}^{k_2} |\mathcal{C}_{0,0,l_2}|,$$

and the lemma follows thereby. ∎

Proof of Theorem 5.2.2. Recall that $U(m)$ is the number of perfect matchings on m points. Then the total number of pairings in $\mathcal{M}(L,R,\mathbf{d})$ is equal to

$$[M_1(R)]_{M_1(L)}U(M_1(R)-M_1(L)) = \frac{M_1(R)!}{2^{(M_1(R)-M_1(L))/2}((M_1(R)-M_1(L))/2)!}.$$

Hence the number of restricted pairings corresponds to simple B-graphs is asymptotically

$$\frac{M_1(R)!\mathbf{P}(\mathbf{d})}{2^{(M_1(R)-M_1(L))/2}((M_1(R)-M_1(L))/2)!}.$$

Since each simple B-graph corresponds to $\prod_{i=1}^n d_i$ pairings in $\mathcal{M}(L,R,\mathbf{d})$, we have

$$g(L,R,\mathbf{d}) = \frac{M_1(R)!\mathbf{P}(\mathbf{d})}{2^{(M_1(R)-M_1(L))/2}((M_1(R)-M_1(L))/2)!\prod_{i=1}^n d_i!}.$$

It only remains to show that

$$\mathbf{P}(\mathbf{d}) = (1 + O((d_{\max}^2 + \ln M)/\sqrt{M}))e^{-\mu_0-\mu_1-\mu_2}. \tag{5.3.14}$$

If $M_2(R) \le d_{\max}^3 w(n)$, by Corollary 5.3.5, $\mathbf{P}(\mathbf{d}) = 1 - O(d_{\max}^4/M + d_{\max}/\sqrt{M}) = 1 - O(d_{\max}^2/\sqrt{M})$. So (5.3.14) is true in this case. Now assume $M_2(R) > d_{\max}^3 w(n)$, which implies $d_{\max}^3 = o(M_2(R))$. Then by Lemma 5.3.14, it is enough to evaluate

$$\sum_{l_2=0}^{k_2}\sum_{l_1=0}^{k_1}\sum_{l_0=0}^{k_0} \frac{|\mathcal{C}_{l_0,l_1,l_2}|}{|\mathcal{C}_{0,0,0}|}.$$

Let $\alpha = d_{\max}^3/M_2(R) + d_{\max}^2 w(n)/M$. By Lemma 5.3.13,

$$\sum_{l_2=0}^{k_2}\sum_{l_1=0}^{k_1}\sum_{l_0=0}^{k_0} |\mathcal{C}_{l_0,l_1,l_2}| = \sum_{l_2=0}^{k_2}\sum_{l_1=0}^{k_1} |\mathcal{C}_{0,l_1,l_2}| \left(\sum_{l_0=0}^{k_0} \mu_0^{l_0}/l_0!\right)\Big(1 + O(k_0\alpha)\Big)$$

Since $\mu_0/k \le 1/8$ for all $k \ge k_0+1$, and $\mu_0^{k_0+1}/(k_0+1)! = O(d_{\max}^2/\sqrt{M})$ by Lemma 5.3.14, we have

$$\sum_{l_0=0}^{k_0} \mu_0^{l_0}/l_0! = \sum_{l_0=0}^{\infty} \mu_0^{l_0}/l_0! + O(\mu_0^{k_0+1}/(k_0+1)!) = e^{\mu_0} + O(d_{\max}^2/\sqrt{M}).$$

So

$$\sum_{l_2=0}^{k_2}\sum_{l_1=0}^{k_1}\sum_{l_0=0}^{k_0} |\mathcal{C}_{l_0,l_1,l_2}| = \sum_{l_2=0}^{k_2}\sum_{l_1=0}^{k_1} |\mathcal{C}_{0,l_1,l_2}|\Big(e^{\mu_0} + O(d_{\max}^2/\sqrt{M})\Big)\Big(1 + O(k_0\alpha)\Big)$$

Similarly we can show that

$$\begin{aligned}\sum_{l_1=0}^{k_1} |\mathcal{C}_{0,l_1,l_2}| &= |\mathcal{C}_{0,0,l_2}|(e^{\mu_1} + O(d_{\max}^2/\sqrt{M}))(1 + O(k_1\alpha)), \quad \text{for all } 0 \le l_2 \le k_2,\\ \sum_{l_2=0}^{k_2} |\mathcal{C}_{0,0,l_2}| &= |\mathcal{C}_{0,0,0}|(e^{\mu_2} + O(d_{\max}^2/\sqrt{M}))(1 + O(k_2\alpha)).\end{aligned}$$

Thus,

$$\sum_{l_2=0}^{k_2}\sum_{l_1=0}^{k_1}\sum_{l_0=0}^{k_0} |\mathcal{C}_{l_0,l_1,l_2}| = |\mathcal{C}_{0,0,0}|e^{\mu_0+\mu_1+\mu_2}\Big(1 + O(d_{\max}^2/\sqrt{M} + (k_0+k_1+k_2)\alpha)\Big).$$

Then we obtain, in this case, that

$$\mathbf{P}(\mathbf{d}) = e^{-\mu_1-\mu_2-\mu_3}\Big(1 + O(d_{\max}^2/\sqrt{M} + (k_0+k_1+k_2)\alpha)\Big).$$

The argument below (5.3.13) shows that for any $i = 0, 1, 2$, $k_i\alpha = O(d_{\max}^2/\sqrt{M})$ if $k_i = \lfloor \max\{8\mu_i, \ln M\}\rfloor$. Next we show that for any $i = 0, 1, 2$, $k_i\alpha = O(d_{\max}^2/\sqrt{M})$ if $k_i = \lfloor w(n)\mu_i\rfloor$. Then $k_i \leq \ln M$ by the definition of k_i in (5.3.9). By (5.3.8), $\mu_i = O(d_{\max}M_2(R)/M)$. Hence

$$k_i d_{\max}^3/M_2(R) = O(w(n)d_{\max}M_2(R)/M)d_{\max}^3/M_2(R) = O(d_{\max}^4 w(n)/M) = O(d_{\max}^2/\sqrt{M}),$$
$$k_i d_{\max}^2 w(n)/M = O(\ln M \cdot d_{\max}^2 w(n)/M) = O(\ln M/\sqrt{M}).$$

Hence $\mathbf{P}(\mathbf{d}) = (1 + O((d_{\max}^2 + \ln M)/\sqrt{M}))e^{-\mu_1-\mu_2-\mu_3}$. Theorem 5.2.2 follows. ∎

Proof of Corollary 5.2.4. First we observe that $\ln x/\sqrt{x}$ is a decreasing function for sufficiently large x. Since $d^2 s = o(n)$, we have $s = o(n)$. Hence we have that for sufficiently large n,

$$\frac{d^2 + \ln(dn-h)}{\sqrt{dn-h}} = \frac{d^2}{\sqrt{dn-h}} + \frac{\ln(dn-h)}{\sqrt{dn-h}} \leq \frac{d^2}{\sqrt{dn-ds}} + O\left(\frac{\ln n}{\sqrt{n}}\right).$$

The last inequality holds because $h \leq ds$ and $dn - h \geq dn - ds \geq n - s = n(1 + o(1))$ and so

$$\frac{\ln(dn-h)}{\sqrt{dn-h}} = O\left(\frac{\ln(n(1+o(1)))}{\sqrt{n(1+o(1))}}\right) = O\left(\frac{\ln n + o(1)}{\sqrt{n}}(1+o(1))\right) = O(\ln n/\sqrt{n}).$$

We also have that

$$\frac{d^2}{\sqrt{dn-ds}} = d^{3/2}/\sqrt{n-s} = d^{3/2}/\sqrt{n}(1+o(1)) = O(d^{3/2}/\sqrt{n}).$$

Hence

$$\frac{d^2 + \ln(dn-h)}{\sqrt{dn-h}} = O\left(\frac{d^{3/2} + \ln n}{\sqrt{n}}\right). \tag{5.3.15}$$

Next we estimate $\mu_i(\mathbf{d}')$ for $i = 0, 1, 2$. By the definition of μ_i in (5.2.1)–(5.2.3), we have

$$\begin{aligned}
\mu_0(\mathbf{d}') &= \frac{(d(n-s)-(ds-h))d(d-1)(n-s)}{2d^2(n-s)^2} = \frac{(dn-2ds+h)(d-1)}{2d(n-s)} \\
&= \frac{d-1}{2}\left(1+O(s/n+h/dn)\right) = \frac{d-1}{2}\left(1+O(s/n)\right) = \frac{d-1}{2} + O(ds/n) \\
\mu_1(\mathbf{d}') &= \frac{d(d-1)(n-s)\sum_{i=1}^{s}[d-k_i]_2}{2d^2(n-s)^2} \le \frac{(d-1)d^2 s}{2d(n-s)} \\
&= \frac{(d-1)d^2}{2d}\frac{s}{n-s} = O\left(d^2 s/n\right) \\
\mu_2(\mathbf{d}') &= \mu_0(\mathbf{d}')^2 = \frac{(d-1)^2}{4}(1+O(s/n)) = \frac{(d-1)^2}{4} + O(d^2 s/n).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\exp\left(-\mu_0(\mathbf{d}') - \mu_1(\mathbf{d}') - \mu_2(\mathbf{d}') + \frac{d^2-1}{4}\right) \\
&\quad = \exp\left(-\frac{d-1}{2} - \frac{(d-1)^2}{4} + \frac{d^2-1}{4} + O(d^2 s/n)\right) = 1 + O(d^2 s/n). \quad (5.3.16)
\end{aligned}$$

Next we estimate

$$\sqrt{\frac{dn-ds}{dn-2ds+h}} \quad \text{and} \quad \frac{(dn-ds)^{dn-ds}}{(dn-2ds+h)^{(dn-2ds+h)/2}(dn)^{dn/2}}.$$

It follows easily that

$$\sqrt{\frac{dn-ds}{dn-2ds+h}} = \sqrt{\frac{dn(1+O(s/n))}{dn(1+O(s/n+h/dn))}} = 1 + O(s/n), \qquad (5.3.17)$$

and

$$\begin{aligned}
&\frac{(dn-ds)^{dn-ds}}{(dn-2ds+h)^{(dn-2ds+h)/2}(dn)^{dn/2}} \\
&= \exp\left((dn-ds)\ln(dn-ds) - \frac{dn-2ds+h}{2}\ln(dn-2ds+h) - \frac{dn}{2}\ln(dn)\right) \\
&= \exp\left((dn-ds)\Big(\ln(dn) + \ln\left(1-\frac{s}{n}\right)\Big) - \frac{dn-2ds+h}{2}\Big(\ln(dn) + \ln\left(1-\frac{2ds-h}{dn}\right)\Big) - \frac{dn}{2}\ln(dn)\right).
\end{aligned}$$

Since

$$\ln\left(1-\frac{s}{n}\right) = -\frac{s}{n} + O\left(\frac{s^2}{n^2}\right),$$

and

$$\ln\left(1-\frac{2ds-h}{dn}\right)=-\frac{2ds-h}{dn}+O\left(\frac{(2ds-h)^2}{(dn)^2}\right)=-\frac{2ds-h}{dn}+O\left(\frac{d^2s^2}{(dn)^2}\right)=-\frac{2ds-h}{dn}+O\left(\frac{s^2}{n^2}\right),$$

the above expression equals to

$$\exp\left(-\frac{h}{2}\ln(dn)-\frac{h}{2}+O(s^2d/n)\right). \tag{5.3.18}$$

Therefore, by (5.3.15)–(5.3.18), the probability $\mathbf{P}_{\mathcal{G}_{n,d}}(S,H)$ in Corollary 5.2.4 can be simplified to

$$\begin{aligned}&\left(1+O\left(\frac{d^{3/2}+\ln n}{\sqrt{n}}\right)\right)(1+O(d^2s/n))e^{h/2}\prod_{i=1}^{s}[d]_{k_i}\exp\left(-\frac{h}{2}\ln(dn)-\frac{h}{2}+O(s^2d/n)\right)\\&=\left(1+O\left(\frac{d^{3/2}+\ln n}{\sqrt{n}}\right)+O(d^2s/n+s^2d/n)\right)(dn)^{-h/2}\prod_{i=1}^{s}[d]_{k_i}\end{aligned}$$

as required. ∎

Chapter 6

Conclusion

6.1 Pegging algorithm for odd $d \geq 3$

The pegging operation for d-regular graphs was only defined for d even in Chapter 3. We extend it to the case of d odd and discuss the properties of the random regular graphs generated by the pegging model in the general case $d \geq 3$. The pegging operation for odd $d \geq 3$ is as follows, illustrated in Figure 6.1 for $d = 3$.

Pegging Operation for Odd d
Input: a d-regular graph G, where d is odd.

1. Let $c := \lfloor d/2 \rfloor$ and choose a set $E_1 = \{u_1u_2, u_3u_4, \ldots, u_{2c-1}u_{2c}\}$ of c pariwise non-adjacent edges in $E(G)$ u.a.r., and another set $E_2 = \{u_{2c+1}u_{2c+2}, \ldots, u_{4c-1}u_{4c}\}$ of c pairwise non-adjacent edges in $E(G) \setminus E_1$ u.a.r.
2. $G := (G \setminus (E_1 \cup E_2)) \cup \{u, v\} \cup E_3 \cup \{uv\}$, where u and v are new vertices added to $V(G)$, and $E_3 = \{uu_1, \ldots, uu_{2c}, vu_{2c+1}, \ldots, vu_{4c}\}$.
3. Output: G.

The following theorem shows that all results in Section 3.2 and the upper bound in Theorem 3.3.1 can be extended to the case of general $d \geq 3$. Moreover, the lower bound in Theorem 3.3.2 was extended to the general case of all even $d \geq 4$ in [26], though the authors believe that it can be extended to the general case of all $d \geq 3$ by analogous arguments. Interested readers can refer to [27, 26] for sketches of the proofs.

Theorem 6.1.1 ([27], Theorem 2.3) *For fixed $k \geq 3$, any integer $d \geq 3$, and fixed initial d-regular graph G_0,*

$$\mathbf{E}Y_{t,d,k} = \frac{(d-1)^k - (d-1)^2}{2k} + O\left(n_t^{-1}\right).$$

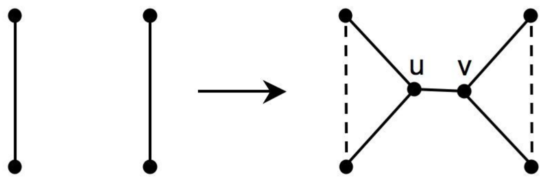

Figure 6.1: *Pegging operation when* $d = 3$

Moreover,

(i) $Y_{t,d,3}, Y_{t,d,4}, \ldots, Y_{t,d,k}$ *have a limiting joint distribution equal to that of independent Poisson variables with means* $\mu_{d,3}$*,* $\mu_{d,4}, \ldots, \mu_{d,k}$*, where* $\mu_{d,i} = ((d-1)^i - (d-1)^2)/(2i)$ *for any fixed integer* $i \geq 3$*,*

(ii) the ϵ*-mixing time satisfies* $\tau^*_\epsilon\big((\sigma_{t,d,k})_{t\geq 0}\big) = O\left(\epsilon^{-1}\right)$*.*

(iii) if d *is even, then the* ϵ*-mixing time satisfies* $\tau^*_\epsilon\big((\sigma_{t,d,k})_{t\geq 0}\big) \neq o\left(\epsilon^{-1}\right)$*.*

6.2 More discussion of contiguity

Theorem 3.2.1 implies that the random d-regular graphs generated by the pegging algorithm are not uniformly distributed, since, in the uniform distribution, the expected number of k-cycles is asymptotic to $(d-1)^k/2k$. Nonetheless, the theorem indicates the possibility that the pegging model and the uniform model might be close in the sense of contiguity. Let $\mathcal{PG}(G_0, t, d)$ be the probability space of all random d-regular graphs generated by the pegging algorithm at time t, starting with graph G_0. Recall that $\mathcal{G}_{n,d}$ denotes the probability space of all random d-regular graphs with uniform distribution. Let $n = n_0 + t$, and $\widehat{G}$ be an arbitrary d-regular graph from $\mathcal{G}_{n,d}$. Let $Y_t(\widehat{G})$ be the number of ways that $\widehat{G}$ could be obtained in the pegging algorithm, i.e. $Y_t(\widehat{G})$ is the total number of all sequences $(G_0, G_1, \ldots, G_t) \in \mathcal{P}(G_0, d)$ such that $G_t = \widehat{G}$. If we can show that $Y_t/\mathbf{E}Y_t$ converges in distribution to some random variable W as $t \to \infty$, and $W > 0$ a.s., it follows that $\mathcal{PG}(G_0, t, d)$ is contiguous with $\mathcal{G}_{n,d}$, where $n = n_0 + t$. If true, this means that properties a.a.s. true in one model are also true in the other model.

The small subgraph conditioning method introduced in Section 2.1.5 gives a way of proving the convergence of $Y_t/\mathbf{E}Y_t$. It is well known that for any integer $m \geq 3$, $X_3, X_4, \ldots, X_m$ are independent Poisson random variables with means $\lambda_i = (d-1)^i/(2i)$, where X_i is the number of i-cycles in a graph $G \in \mathcal{G}_{n,d}$. Theorem 6.1.1 shows that for any integer $d \geq 3$, the numbers of short cycles in random d-regular graphs generated by pegging are asymptotically independent Poisson random variables with means $\mu_i = ((d-1)^i - (d-1)^2)/(2i)$. To use the small subgraph conditioning method, one computes

$$\delta_i = \frac{\mu_i}{\lambda_i} - 1 = -\frac{(d-1)^2}{(d-1)^i},$$

and then it is easy to check that

$$\sum_{i=3}^{k} \lambda_i \delta_i^2 < \infty.$$

It then becomes conceivable that the variation in probabilities in the pegging model is strongly associated with the varying numbers of short cycles (see [59] and [36]). According to the method, this would be proved if we could show that $\mathbf{E}Y_t^2/(\mathbf{E}Y_t)^2 \leq \exp\left(\sum_{i=3}^{k} \lambda_i \delta_i^2\right) + o(1)$. However, the estimation of $\mathbf{E}Y_t^2$ seems to be out of reach at present. Thus, we make the following conjecture.

Conjecture 6.2.1 *For all fixed $d \geq 3$ and every fixed d-regular graph G_0, $\mathcal{PG}(G_0, t, d) \approx \mathcal{G}_{n,d}$, where $t = n - n_0$.*

For any fixed $d \geq 3$, it is well known that the random d-regular graphs with the uniform distribution are d-connected and have diameter at most $1 + \lceil \log_{d-1}((2+\epsilon)dn \log n) \rceil$ a.a.s. (See [62] for facts not referenced here.) These properties are of central interest where the graphs are used as communication networks. We have determined the connectivity of random regular graphs in $\mathcal{P}(G_0, d)$ in the previous section, which supports Conjecture 6.2.1. If the conjecture holds, it implies that the random regular graphs in $\mathcal{PG}(G_0, t, d)$ are a.a.s. d-connected with diameter $O(\log n)$, where $n = n_0 + t$. In any case, the logarithmic diameter is common among random networks with average degree above 1. In the Erdős-Rényi model of random graphs, the components of the random graph a.a.s. all have diameter $O(\log n)$ if the edge probability p is at least c/n for some $c > 1$. Ferholz and Ramachandran [24] showed that the diameter of random sparse graphs with given degree sequences is a.a.s. $c(1 + o(1)) \log n$, when the degree sequences satisfy some natural convergence conditions, and they determined the value of c. Bollobás and Riordan [12] proved that the random graphs generated by the preferential attachment model a.a.s. have diameter asymptotically $\log n / \log \log n$. Gerke et al. [30] recently showed that the diameter of random regular graphs in $\mathcal{PG}(G_0, t, d)$ is a.a.s. $O(\log n)$. The proof uses the result that $G \in \mathcal{PG}(G_0, t, d)$ is a.a.s. d-connected.

6.3 More about future work

In Section 4.6, we discussed a more general setting of the hypergraph orientation problem. However, more interesting future work is to extend our result to smaller values of k or to find linear time algorithms that a.a.s. find an optimal solution. Extending the results in Section 4.2 to any $k \geq 1$ could be quite difficult since studies of the problem for all values of k require more knowledge of the solution to the differential equations in Section 4.3. For instance, we need to know the existence of thresholds of the appearance of the (w, k)-core, and if a threshold exists, where it occurs. Currently we are not able to cope with this without an analytical solution to the differential equations. However, it might be possible to improve our results for the case $w = 1$ and general $h \geq 2$, for which the k-core of a random uniform hypergraph has been well studied.

As discussed in Section 2.2.1, there are a few natural ways of generalising the algorithm used in [15], but the analysis would become even more complicated than the random graph case. It is desirable to find some "clever" algorithm which is easier to analyse.

Lastly we discuss some application of the results in Chapter 5. Consider the following problem.

Let G be a random graph from $\mathcal{G}(n, p)$ and let G_O be the subgraph of G induced by vertices of G whose degrees are odd. What is the probability space generated by G_O?

This problem is interesting because it helps to study the famous Chinese postman problem. Given any connected graph G, the Chinese postman problem asks for its shortest closed trail which travels each edge at least once. It is well known that if the subgraph of G induced by the odd degree vertices has a perfect matching M, then the set of edges of the shortest closed trail that travels each edge at least once is $E(G) \cup M$, and a linear time algorithm outputs the required trail. Therefore, studies of the properties of the subgraph induced by the odd degree vertices are of particular interest. We plan to apply the results in Section 5.2 to study G_O in $\mathcal{G}(n, p)$. When $\log n/n \leq p \ll n^{-2/3}$, intuitively, the probability space G_O should be equivalent or close to $\mathcal{G}(n/2, p)$, because around half of the vertices would be of odd degree and the odd degree vertices are distributed almost randomly since each vertex has approximately the same chance to be of odd degree. Then the probability of an edge to occur between any pair of vertices in G_O should be around p and almost independently of other edges. Studying G_O for $p = O(1/n)$ can also be interesting.

References

[1] David Aldous. Mixing time for a markov chain on cladograms. (english summary). *Combin. Probab. Comput.*, 9(3):191–204, 2000.

[2] David Aldous, László Lovász, and Peter Winkler. Mixing times for uniformly ergodic markov chains. (english summary). *Stochastic Process. Appl.*, 71(2):165–185, 1997.

[3] N. Alon and J. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ. xviii+352 pp., third edition, 2008.

[4] Y. Azar, A. Z. Broder, A. R. Karlin, and E. Upfal. Balanced allocations. *SIAM J. Comput.*, 29(1):180–200, 1999.

[5] A. Barabási and R. Albert. Emergence of scaling in random networks. (english summary). *Science*, 286(5439):509–512, 1999.

[6] E. Bender. The asymptotic number of non-negative integer matrices with given row and column sums. *Discrete Math*, 10:217–223, 1974.

[7] E. Bender and E. Canfield. The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory Ser. A*, 24(3):296–307, 1978.

[8] P. Berenbrink, A. Czumaj, A. Steger, and B. Vőcking. Balanced allocations: The heavily loaded case. *32th Annual ACM Symposium on Theory of Computing*, pages 745–754, 2000.

[9] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Europ. J. Combinatorics*, 1:311–316, 1980.

[10] B. Bollobás. *Random graphs, Combinatorics, (Swansea, 1981), London Math. Soc. Lecture Note Ser., 52*. Cambridge Univ. Press, Cambridge-New York, 1981.

[11] B. Bollobás and B. McKay. The number of matchings in random regular graphs and bipartite graphs. *J Combin Theory Ser B*, 41:80–91, 1986.

[12] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24(1):5–34, 2004.

[13] Béla Bollobás and Alan M. Frieze. On matchings and hamiltonian cycles in random graphs. In *Random graphs '83 (Poznan', 1983)*, pages 23–46. North-Holland Math. Stud., 118, North-Holland, Amsterdam, 1985.

[14] V. Bourassa and F. Holt. Swan: Small-world wide area networks. In *Proceeding of International Conference on Advances in Infrastructures (SSGRR 2003w)*. North-Holland Math. Stud., 118, North-Holland, Amsterdam, L'Aquila, Italy, 2003.

[15] J. Cain, P. Sanders, and N. Wormald. The random graph threshold for k-orientiability and a fast algorithm for optimal multiple-choice allocation. January 2007.

[16] J. Cain and N. Wormald. Encores on cores. *Electronic Journal of Combinatorics*, 13:RP 81, 2006.

[17] Herman Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statistics*, 23:493–507, 1952.

[18] V. Chvátal. Almost all graphs with $1.44n$ edges are 3-colorable. *Random Structures Algorithms*, 2(1):11–28, 1991.

[19] A. Bruce Clarke and Ralph L. Disney. *Probability and Random Processes: A First Course with Applications.* John Wiley & Sons, second edition, 1985.

[20] William J. Cook, William H. Cunningham, William R. Pulleyblank, and Alexander Schrijver. *Combinatorial optimization. (English summary).* Wiley-Interscience Series in Discrete Mathematics and Optimization. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, x+355 pp., 1998.

[21] C. Cooper, M. Dyer, and C. Greenhill. Sampling regular graphs and a peer-to-peer network. In *Proceedings of the sixteenth annual ACM-SIAM Symposium on Discrete Algorithms*, pages 980–988, 2005.

[22] M. Dietzfelbinger and F. Meyer auf der Heide. Simple, efficient shared memory simulations. In *5th ACM Symposium on Parallel Algorithms and Architectures, Velen, Germany, SIGACT and SIGARCH*, pages 110–119, June 30July 2, 1993.

[23] P. Erdős and A. Rényi. *On random graphs. I.*, volume 6. Publ. Math. Debrecen, 1959.

[24] D. Fernholz and V. Ramachandran. The k-orientability thresholds for $G_{n,p}$. In *Proc. ACM-SIAM Symposium on Discrete Algorithms (SODA)*, January 2007.

[25] P. Gao. Models of generating random regular graphs and their short cycle distribution. `http://www.math.uwaterloo.ca/~p3gao/writings/P.GAO.pdf`, December 2006.

[26] P. Gao and N. Wormald. Rate of convergence of the short cycle distribution in random regular graphs generated by pegging. *Elec. J. Combinatorics*, 16(1):R44, 17 pp., 2009.

[27] P. Gao and N. Wormald. Short cycle distributions in random regular graphs recursively generated by pegging. *Random Struct. Algorithms*, 34(1):54–86, 2009.

[28] P. Gao and N. Wormald. Load balancing and orientability thresholds for random hypergraphs. submitted, 2010.

[29] Z. Gao and N. Wormald. Distribution of subgraphs of random regular graphs. *Random Struct. Algorithms*, 32(1):38–48, 2008.

[30] S. Gerke, A. Steger, and N. Wormald. Pegging graphs yields a small diameter (manuscript).

[31] C. Greenhill and B. McKay. Random dense bipartite graphs and directed graphs with specified degrees. to appear.

[32] S. L. Hakimi. On the degrees of the vertices of a directed graph. *J. Franklin Inst.*, 279:290–308, 1965.

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

[34] F.B. Holt, V. Bourassa, A.M. Bosnjakovic, and J. Popovic. Swan - highly reliable and efficient networks of true peers. In *Handbook on Theoretical and Algorithmic Aspects of Sensor, Ad Hoc Wireless, and Peer-to-Peer Networks*, pages 799–824. CRC Press, Boca Raton, Florida, 2005.

[35] S. Janson. Random regular graphs: asymptotic distributions and contiguity. *Combin. Probab. Comput.*, (4):369–405, 1995.

[36] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs.* Wiley, New York, 2000.

[37] Mark Jerrum. A very simple algorithm for estimating the number of k-colorings of a low-degree graph. (english summary). *Random Structures Algorithms*, (2):157–165, 1995.

[38] R. M. Karp, M. Luby, and F. Meyer auf der Heide. Efficient pram simulation on a distributed memory machine. In *24th ACM Symp. on Theory of Computing*, pages 318–326, May 1992.

[39] M. Krivelevich, B. Sudakov, and N. Wormald. Regular induced subgraphs of a random graph. *preprint*, 2009.

[40] T. Lindvall. *Lectures on the Coupling Method*. Wiley, New York, 1992.

[41] B. McKay. Subgraphs of random graphs with specified degrees. In *Proceedings of the Twelfth Southeastern Conference on Combinatorics, Graph Theory and Computing*, volume II, pages 213–223, 1981.

[42] B. McKay. Asymptotics for symmetric 0-1 matrices with prescribed row sums. *Ars Combin.*, 19(A):15–25, 1985.

[43] B.D. McKay and N.C. Wormald. Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$. *Combinatorica*, 11(4):369–382, 1991.

[44] F. Meyer auf der Heide, C. Scheideler, and V. Stemann. Exploiting storage redundancy to speed up randomized shared memory simulations. *Theoret. Comput. Sci.*, 162(2):245–281, August 1996.

[45] M. Mitzenmacher, A. Richa, and R. Sitaraman. The power of two random choices: a survey of techniques and results. In *Handbook of randomized computing*, volume I, II, pages 255–312. Comb. Optim., 9, Kluwer Acad. Publ., Dordrecht, 2001.

[46] G. Peano. Démonstration de l'intégrabilité des équations différentielles ordinaires. *(French) Math. Ann.*, 37(2):182–228, 1890.

[47] B. Pittel and N. Wormald. Asymptotic enumeration of sparse graphs with a minimum degree constraint. *J. Combin. Theory Ser. A*, 101(2):249–263, 2003.

[48] M. M. Rao and R. J. Swift. *Probability theory with applications*. Mathematics and Its Applications (Springer) 582. Springer, New York, 2006. xviii+527 pp., second edition, 2006.

[49] H. Robalewska. Random graphs generated by the star 2-process. *Random Structures and Algorithms*, 11:125–149, 1997.

[50] H. Robalewska and N.C. Wormald. Random star d-processes. *Combinatorics, Probability and Computing*, 9:33–43, 2000.

[51] R.W. Robinson and N.C. Wormald. Almost all cubic graphs are hamiltonian. *Random Structures Algorithms*, 3:117–125, 1992.

[52] A. Ruciński. Subgraphs of random graphs: a general approach. In Random graphs *'83 (Poznan, 1983)*, pages 221–229. North-Holland Math. Stud., 118, North-Holland, Amsterdam, 1985.

[53] A. Ruciński. Induced subgraphs in a random graph. In Random graphs *'85 (Poznan, 1985)*, pages 275–296. North-Holland Math. Stud., 144, North-Holland, Amsterdam, 1987.

[54] A. Ruciński. When are small subgraphs of a random graph normally distributed? *Probab. Theory Related Fields*, 78(1):1–10, 1988.

[55] A. Ruciński and N.C. Wormald. Random graphs processes with degree restrictions. *Combinatorics, Probability and Computing*, 1:169–180, 1992.

[56] P. Sanders, S. Egner, and J. Korst. Fast concurrent access to parallel disks. *Algorithmica*, 35(1):21–55, 2003.

[57] Alexander Schrijver. *Combinatorial optimization. Polyhedra and efficiency*, volume B. Algorithms and Combinatorics, 24, Springer-Verlag, Berlin, 2003. pp. i–xxxiv.

[58] Herbert Simon. On a class of skew distribution functions. *Biometrika*, 42:425–440, 1955.

[59] N.C. Wormald. The asymptotic distribution of short cycles in random regular graphs. *J. Combinatorial Theory. Ser. B*, 31:168–182, 1981.

[60] N.C. Wormald. Differential equations for random processes and random graphs. *The Annals of Applied Probability*, 5(4):1217–1235, November 1995.

[61] N.C. Wormald. The differential equation method for random graph processes and greedy algorithms. In M. Karonski and H.J. Proemel, editors, *Lectures on Approximation and Randomized Algorithms*, pages 73–155. PWN, Warsaw, 1999.

[62] N.C. Wormald. Models of random regular graphs. In J.D. Lamb and D.A. Preece, editors, *Surveys in Combinatorics, 1999*, pages 239–298. London Mathematical Society Lecture Note Series **267** Cambridge University Press, Cambridge, 1999.

[63] G. Udny Yule. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Philosophical Transactions of the Royal Society of London. Series B*, pages 21–87, 1925.