Asymptotic enumeration of sparse uniform hypergraphs, with applications

Author:
Aldosari, Haya

Publication Date:
2020

DOI:
https://doi.org/10.26190/unsworks/22187

ASYMPTOTIC ENUMERATION OF SPARSE UNIFORM HYPERGRAPHS, WITH APPLICATIONS

Haya Saeed Aldosari

Supervisor: Professor Catherine Greenhill

A thesis in fulfilment of the requirements for the degree of

Doctor of Philosophy

School of Mathematics and Statistics

Faculty of Science

UNSW Sydney

June 2020

Thesis/Dissertation Sheet

Surname/Family Name	:	**Aldosari**
Given Name/s	:	**Haya Saeed**
Abbreviation for degree as give in the University calendar	:	**PhD**
Faculty	:	**Faculty of Science**
School	:	**School of Mathematics and Statistics**
Thesis Title	:	**Asymptotic enumeration of sparse uniform hypergraphs, with applications**

Abstract 350 words maximum: (PLEASE TYPE)

A hypergraph is a generalisation of a graph where the edges may contain more than two vertices. This thesis concentrates on asymptotic enumeration of some families of sparse uniform hypergraphs, when the number of vertices is sufficiently large and each vertex degree is given. Our first result gives an asymptotic formula for the number of simple uniform hypergraphs with a given degree sequence which contain no edges of another specified hypergraph. This formula holds under some restrictions on the number of forbidden edges, and the maximum degree and edge size of the hypergraph. We apply a combinatorial argument known as the switching method to obtain our estimates.

This formula allows us to calculate the probability that a random uniform hypergraph with given degrees contains a specified set of edges. As a result, we find an asymptotic formula for the expected number of perfect matchings and the expected number of loose Hamilton cycles in random regular uniform hypergraphs. As another application of our first result, we study the average number of uniform spanning hypertrees.

Finally, we give a new proof of Blinovsky and Greenhill's asymptotic formula for the number of sparse linear uniform hypergraphs with given degrees. Our proof uses a more complicated switching and extends the range of degrees and edge size for which the formula holds.

ORIGINALITY STATEMENT

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed ……………………………………………………............

Date ……26/06/2020……………………………………............

COPYRIGHT STATEMENT

Signed

Date 26/06/2020

AUTHENTICITY STATEMENT

‘I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis.’

Signed ……………………………………………..........................

Date ……………………………………………..............................

INCLUSION OF PUBLICATIONS STATEMENT

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in their thesis in lieu of a Chapter if:

- The candidate contributed greater than 50% of the content in the publication and is the "primary author", ie. the candidate was responsible primarily for the planning, execution and preparation of the work for publication
- The candidate has approval to include the publication in their thesis in lieu of a Chapter from their supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis

Please indicate whether this thesis contains published material or not:

☐ This thesis contains no publications, either published or submitted for publication
(if this box is checked, you may delete all the material on page 2)

☒ Some of the work described in this thesis has been published and it has been documented in the relevant Chapters with acknowledgement
(if this box is checked, you may delete all the material on page 2)

☐ This thesis has publications (either published or submitted for publication) incorporated into it in lieu of a chapter and the details are presented below

CANDIDATE'S DECLARATION

I declare that:

- I have complied with the UNSW Thesis Examination Procedure
- where I have used a publication in lieu of a Chapter, the listed publication(s) below meet(s) the requirements to be included in the thesis.

Candidate's Name	Signature	Date (dd/mm/yy)
Haya Saeed Aldosari		26/06/20

Contents

Dedication

I dedicate this thesis to my mother (Fahidah Aldosari) for her priceless support.

Acknowledgements

The decision to choose which country to do my PhD overseas was a hard one, luckily, I decided on Australia and found myself in Sydney. It was a turning point in my life not only for completing my research in Australia but also for having valuable friendship and infinite experiences along the way.

So, I would like to thank the government of Saudi Arabia and Prince Sattam bin Abdulaziz University for the scholarship and financial support they gave me during my PhD journey. I am so indebted and grateful to my Supervisor, Prof. Catherine Greenhill for the best supervision in doing my PhD research. It would have been impossible to complete this research without my supervisor. So thanks Catherine for encouraging me and supporting me financially to present my talks at some conferences. I cannot thank you enough in few words, but I owe all the knowledge I acquired about combinatorics. Moving forward, these knowledge will positively influence my academic life. I am grateful to Prof. Brendan McKay (ANU) for his helpful suggestions during my research. I would also like to thank the School of Mathematics and Statistics at UNSW, and special thanks to the past and current members of the combinatorics group at UNSW.

I would like to express my sincere gratitude to my mum for being the kindest person in my life and for her emotional supports during my stay in Australia. Unfortunately, I lost her before finishing my PhD. I would also thank my father and siblings for their emotional support during my study. Special thanks to my cousin (Jozza) for encouraging me and being my close friend despite the thousands of miles between us.

I would also like to thank all my doctoral colleagues and friends for sharing their experiences with me and showing their support during my study. Last but not least, I am also thankful to my friends Sarah Rashed and Shaymaa Al-Shakarchi for exchanging our academic experiences and sharing the emotional support and sharing the details of our life in the last few years.

Abstract

A hypergraph is a generalisation of a graph where the edges may contain more than two vertices. This thesis concentrates on asymptotic enumeration of some families of sparse uniform hypergraphs, when the number of vertices is sufficiently large and each vertex degree is given. Our first result gives an asymptotic formula for the number of simple uniform hypergraphs with a given degree sequence which contain no edges of another specified hypergraph. This formula holds under some restrictions on the number of forbidden edges, and the maximum degree and edge size of the hypergraph. We apply a combinatorial argument known as the switching method to obtain our estimates.

This formula allows us to calculate the probability that a random uniform hypergraph with given degrees contains a specified set of edges. As a result, we find an asymptotic formula for the expected number of perfect matchings and the expected number of loose Hamilton cycles in random regular uniform hypergraphs. As another application of our first result, we study the average number of uniform spanning hypertrees.

Finally, we give a new proof of Blinovsky and Greenhill's asymptotic formula for the number of sparse linear uniform hypergraphs with given degrees. Our proof uses a more complicated switching and extends the range of degrees and edge size for which the formula holds.

List of Publications

- H. S. Aldosari and C. Greenhill, Enumerating sparse uniform hypergraphs with given degree sequence and forbidden edges, *European Journal of Combinatorics* **77** (2019), 68-77.
- H. S. Aldosari and C. Greenhill, The average number of spanning hypertrees in sparse uniform hypergraphs, *Discrete Mathematics, (to appear)* arXiv:1907.04993.

List of Figures

Chapter 1

Introduction

A hypergraph is a generalisation of a graph, which can be used to model the relationships between objects in a discrete (finite) system. Graphs capture relationships between pairs of objects, whereas hypergraphs can have edges of any size. Hypergraphs are used to model complex discrete systems in many areas, including ecology [56], quantum computing [72], social networks [4], computer science [22], medicine [88] and chemistry [51].

Graphs and random graphs have been extremely well-studied over many years, partly due to their usefulness in modelling in various areas. However, there are a limited number of analogous results in hypergraphs. In particular, there are relatively few asymptotic enumeration results for hypergraphs. This lack of extension motivates us in conducting this research, where we focus on asymptotic enumeration of hypergraphs.

In asymptotic enumeration, we seek an approximation for the number of combinatorial objects from a certain family, such that the relative error in the formula tends to zero as the size of the objects increases. Having an approximate formula is useful in situations where exact enumeration is impossible, and allows us to understand the asymptotic behaviour of the function being approximated. For example, there are exactly $2^{\binom{n}{2}}$ graphs on n vertices, but it is believed that there is no exact formula for the number of 3-regular graphs on n vertices. Hence an asymptotic formula for the number of 3-regular graphs is useful. (Recall that a graph is 3-regular if every vertex has 3 neighbours.)

In the case of hypergraphs, in addition to the degrees of the vertices, there are other parameters such as the size of the edges and the amount of overlap allowed between edges. A hypergraph is uniform if every edge has the same size. In general, a graph or hypergraph is called *sparse* if it has relatively few edges, and a graph or hypergraph with many edges is called *dense*. There are no precise definitions for the terms sparse and dense, as these terms will vary depending on the context. However, it is still a useful distinction as different methods are often needed for the sparse and dense cases.

The purpose of this thesis is to prove some asymptotic enumeration results for sparse uniform hypergraphs which satisfy certain properties. We will provide three new asymptotic enumeration formulae, and demonstrate some applications. First, we study the approximate number of simple uniform hypergraphs with a given degree sequence which avoid the edges of a given hypergraph. As a corollary, we deduce an asymptotic formula for the probability that a random uniform hypergraph with given degree sequence contains a specified hypergraph as a subhypergraph. This will be applied in the regular setting to compute the average number of perfect matchings and the average number of loose Hamilton cycles in regular hypergraphs with n vertices. As an additional application, we study the expected number of spanning hypertrees in uniform hypergraphs where the degrees of vertices may vary. Our third formula is an asymptotic expression for the number of linear uniform hypergraphs with a given degree sequence, extending a result of [13]. All of these results require certain conditions to hold on the maximum degree and the edge size.

In the rest of this chapter we define some concepts and notation used in this thesis. We also introduce some random models for graphs and hypergraphs, and illustrate the idea of the switching method, which we use in our asymptotic enumeration proofs. Then, in Chapter 2 we discuss the relevant background literature.

In Chapter 3, we compute the asymptotic number of simple hypergraphs with a given degree sequence which avoid a specified set of edges. This result, Theorem 3.1.1, will be proved by using a switching operation designed to remove the forbidden edges one by one. Then we prove Corollary 3.1.2 which gives an asymptotic formula for the probability that a random simple uniform hypergraphs with given degrees contains another hypergraph as subhypergraph.

Chapter 4 contains three applications of Corollary 3.1.2, proved in the preceding chapter. The first application is Corollary 4.1.1, giving an asymptotic formula for the expected number of perfect matchings in regular uniform hypergraphs. Then in Corollary 4.2.1 we provide a similar result for loose Hamilton cycles. Finally, we study the average number of spanning uniform hypertrees in uniform hypergraphs with given degree sequence, leading to Theorem 4.3.1.

In Chapter 5, we estimate the number of linear uniform hypergraphs with specified degrees of its vertices. A hypergraph is linear if each pair of edges intersects in at most one vertex. We work with bipartite graphs, using a well-known correspondence between hypergraphs and bipartite graphs. Our proof applies the switching method again, but on edges of bipartite graphs. Our result, Theorem 5.1.1, extends the formula of Blinovsky and Greenhill [13] to allow a wider range of parameters. This extension is achieved by using a more complicated switching than in [13], involving more edges. Our proof also gives an asymptotic formula for the number of half-regular bipartite graphs with given degrees and girth at least 6, stated in Corollary 5.4.3. A bipartite graph is half-regular if one of the parts is regular. The *girth* of a graph is defined as the length of the shortest cycle in the graph.

To conclude the thesis, in Chapter 6 we summarise our main results and give some possible directions for future work.

The results of Chapter 3, together with Corollary 4.1.1 and Corollary 4.2.1, were published in [2]. The work from Section 4.3 has been submitted for publication [1]. Both papers were collaborations with my supervisor.

1.1 Definitions and notation

A *hypergraph* is defined by a set of vertices $V = \{v_1, v_2, \ldots, v_n\}$ and a multiset E of multisubsets of V. The elements of E are called *hyperedges*. For simplicity, we will use *edges* instead of hyperedges. This definition allows an arbitrary vertex v to occur more than once in a hyperedge, which is known as a *loop* at vertex v. We may also have a repeated edges in a hypergraph where more than one edges has the same set of vertices. A hypergraph is *simple* if it has no loops and no repeated edges. The defintion of hypergraph extends the concept of *graph* whose edges contain precisely two vertices, including the multiplicity. This extension allows many authors to define analogue properties in hypergraphs.

Our research will focus on asymptotic enumeration of simple hypergraphs with specific degrees for its (labelled) vertices. For $n \geq 3$, let $\boldsymbol{k} = \boldsymbol{k}(n) = (k_1, \ldots, k_n)$ be a sequence of nonnegative integers with maximum value $k_{\max}$. For a vertex set $V = \{v_1, \ldots, v_n\}$, we say that a hypergraph H on V has degree sequence $\boldsymbol{k}$ if v_i has degree k_i, that is $k_i = |\{e \in H : v_i \in e\}|$. When each vertex has the same degree in a hypergraph, then we call it a *regular* hypergraph. For a positive integer k, we say that a hypergraph is k-*regular* if every vertex of this hypergraph has degree k. A *uniform* hypergraph is a hypergraph where all its edges contain the same number of vertices. We will write r-uniform hypergraph if we have exactly r vertices in every edge of the hypergraph, accounting the multiplicities of a vertex. Therefore, a 2-uniform hypergraph is known as graph.

We denote by $\mathcal{H}_r(\boldsymbol{k})$ the set of all simple r-uniform hypergraphs on V with degree sequence $\boldsymbol{k}$. For a positive integer k, the regular case of $\mathcal{H}_r(\boldsymbol{k})$ is denoted by $\mathcal{H}_r(k, n)$, where all the vertices have the same degree k. A hypergraph is *linear* if contains no loops and every pair of distinct edges overlap in at most one vertex. This implies that, when all edges of a hypergraph contains at least two vertices that a linear hypergraph is also simple. The notation $\mathcal{L}_r(\boldsymbol{k})$ will be used for the set of all linear hypergraphs in $\mathcal{H}_r(\boldsymbol{k})$.

For a hypergraph H, we say that another hypergraph H' is a *subhypergraph* of H if the vertex set of H' is a subset of the vertex set of H and the set of edges of H' is a subset of the edge set of H. This can be written as $H' \subseteq H$. Under this definition, observe that any subhypergraph of a r-uniform hypergraph must also be r-uniform. When working with graphs we say subgraph, rather than subhypergraph.

Next, we define some interesting substructures which may occur in hypergraphs. When n is divisible by r, a set of n/r disjoint edges in an r-uniform hypergraph H on n vertices, which covers all the vertices of H exactly once is called a *perfect matching* in H. Figure 1.1 shows an example of a 3-uniform hypergraph which contains a perfect matching as a subhypergraph.

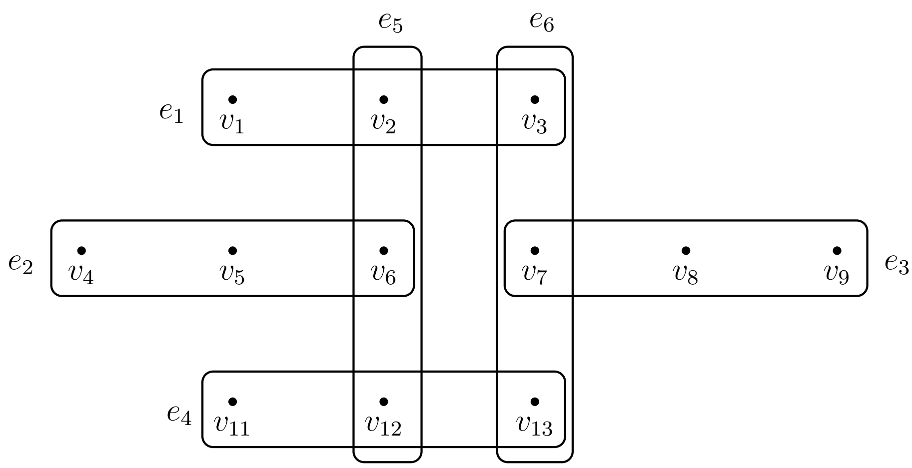

Figure 1.1: A 3-uniform hypergraph with a perfect matching $\{e_1, e_2, e_3, e_4\}$.

Now we consider the concepts of paths and cycles in hypergraphs. There are several choices of definition, which is a bit more complex and wider than in graphs, since the edges of hypergraphs may overlap in more than one vertex. A *Berge path* in H consists of a sequence $v_0, e_1, v_1, e_2, \dots, e_\ell, v_\ell$ where $v_0, v_1, \dots, v_\ell$ are distinct vertices, $e_1, \dots, e_\ell$ are distinct edges and $v_{i-1}, v_i \in e_i$ for all $i = 1, \dots, \ell$. The length of the shortest path between two vertices x and y in hypergraph H is denoted by $\text{dist}_H(x, y)$. A *Berge cycle* of length ℓ is defined similarly, with a sequence $v_0, e_1, v_1, e_2, \dots, v_{\ell-1}, e_\ell$ of alternating distinct vertices and edges, such that v_{i-1}, $v_i \in e_i$ for all $i = 1, \dots, \ell - 1$ and $v_0, v_{\ell-1} \in e_\ell$. These are the most general definitions of paths and cycles in hypergraphs. In particular, a cycle of length 1, also called 1-*cycle* or loop, is an edge which contains a repeated vertex. A 2-*cycle* is given by two edges which intersect in at least two vertices.

Special cases of Berge cycles are s-*overlapping cycles*, where consecutive edges intersect in exactly s vertices. For any integer $\ell \geq 3$, a 1-overlapping cycle with length ℓ is a hypergraph with ℓ distinct edges which can be labelled as $e_1, \ldots, e_\ell$ such that there exists distinct vertices $v_1, \ldots, v_\ell$ with $e_i \cap e_{i+1} = \{v_i\}$ for $i = 1, \ldots, \ell$ (identifying $e_{\ell+1}$ with e_1). We also refer to 1-overlapping cycles as *loose cycles*. A 3-uniform loose cycle with 20 vertices and 10 edges is shown in Figure 1.2.

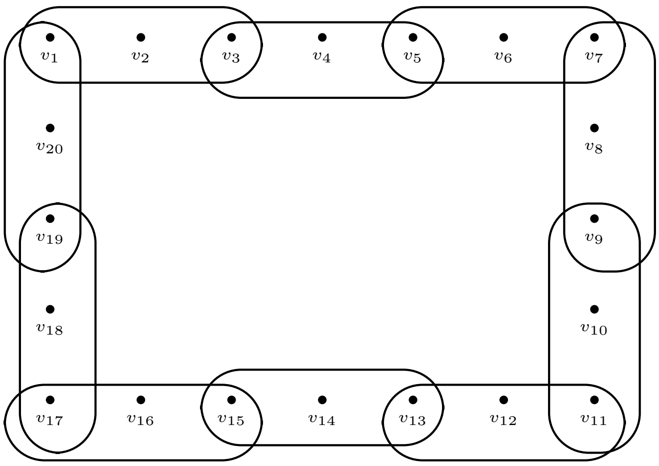

Figure 1.2: A 3-uniform loose cycle

When a cycle forms a subhypergraph of a hypergraph and contains all vertices of the hypergraph, it is known as a *Hamilton cycle*. It is described as *loose* Hamilton cycle if it is 1-overlapping cycle that covers all the vertices of the hypergraph. Precisely, a loose Hamilton cycle is a set of $t = \frac{n}{r-1}$ edges which can be labelled as $e_0, \ldots, e_{t-1}$ such that, for some ordering $v_0, \ldots, v_{n-1}$ of the vertices,

$$e_i = \{v_{i(r-1)}, v_{i(r-1)+1}, \ldots, v_{i(r-1)+(i-1)}\},$$

for $i = 0, \ldots, t-1$. For obvious reasons $r-1$ must be a factor of n in this cycle. Let C be a loose Hamilton cycle with edge set $\{e_1, \ldots, e_t\}$. The degree sequence of

C has t vertices of degree 2 and all remaining vertices of degree 1. Figure 1.3 shows a 4-uniform hypergraph on 30 vertices which contains a loose Hamilton cycle with $30/3 = 10$ edges.

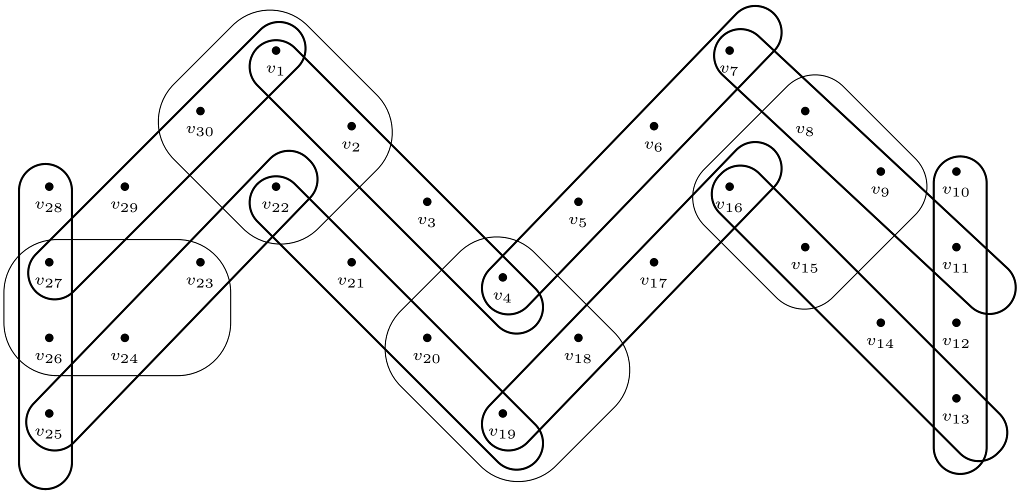

Figure 1.3: A loose Hamilton cycle in a 4-uniform hypergraph

A hypergraph is *connected* if there is a Berge path between every pair of vertices. A *hypertree* is a connected hypergraph which contains no Berge cycles. Under this definition, any hypertree must be linear since any pair of edges which intersect in at least two vertices generates a 2-cycle. A *spanning hypertree* in a hypergraph H is a subhypergraph of H which forms a hypertree containing all vertices of H. In other words, a hypergraph $T = (V(T), E(T))$ is a spanning hypertree in H if T is an acyclic, connected subhypergraph of H with $V(T) = V(H)$. We sometimes abbreviate "r-uniform hypertree" to "r-hypertree". Note that an r-hypertree on n vertices has exactly $\frac{n-1}{r-1}$ edges. Hence when studying spanning r-hypertrees in hypergraphs with n vertices, it is necessary that $n-1$ is divisible by $r-1$. Figure 1.4 demonstrates a 3-uniform hypertree on 17 vertices with 8 edges.

We complete this section with some useful notation. For positive integers a and b we define $[a] = \{1, 2, \dots, a\}$ and $(a)_b$ for the falling factorial $a(a-1)\cdots(a-b+1)$.

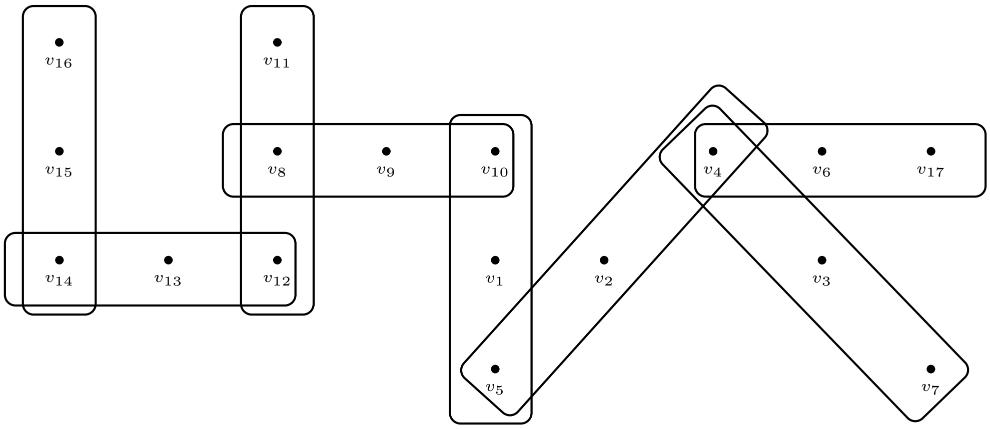

Figure 1.4: A 3-uniform hypertree

For a degree sequence $\boldsymbol{k} = (k_1, \dots, k_n)$, define $M_s(\boldsymbol{k}) = \sum_{i=1}^{n} (k_i)_s$ where s is a positive integer. This implies that

$$M_1(\boldsymbol{k}) = M(\boldsymbol{k}) = \sum_{i=1}^{n} k_i, \quad M_s(\boldsymbol{k}) \leq k_{\max} M_{s-1}(\boldsymbol{k})$$

for $s \geq 2$. For simplicity, we will write M and M_2 respectively, instead of $M(\boldsymbol{k})$ and $M_2(\boldsymbol{k})$, when our calculation involves only one degree sequence $\boldsymbol{k}$. Throughout this thesis, we assume that r divides $M(\boldsymbol{k})$ for the existence of r-uniform hypergraphs with degree sequence $\boldsymbol{k}$.

We use standard asymptotic notation which we review here. For sequences (a_n), (b_n) of real numbers, we write $a_n = O(b_n)$ as $n \to \infty$ if there exist positive constants C and n_0 such that $|a_n| \leq C|b_n|$ for $n \geq n_0$. Also, we write $a_n = o(b_n)$ if $|a_n|/|b_n| \to 0$ as $n \to \infty$. If $a_n = O(b_n)$ then we can also write $b_n = \Omega(a_n)$. Similarly, we can write $b_n = \omega(a_n)$ instead of $a_n = o(b_n)$.

1.2 Random hypergraphs

Random graphs can be used to investigate whether a "typical" graph on n vertices has a desired property.

There are two popular random graph models, denoted by $\mathcal{G}_{n,p}$ and $\mathcal{G}_{n,m}$. In the binomial model $\mathcal{G}_{n,p}$, a graph on n vertices can be generated by joining two distinct vertices with an edge with probability p, independently for each pair of vertices. Here the expected number of edges is $p\binom{n}{2}$. On the other hand, $\mathcal{G}_{n,m}$ is the uniform probability space on the set of all graphs with n vertices and m edges. This can be formed by choosing an m-element subset of the set of all $\binom{n}{2}$ possible edges, uniformly at random. For many graph properties, these two models are equivalent. There are many more details in the textbooks by Bollobás [15], Janson et al. [43] and Frieze and Karoński [31].

In some real-world applications it is useful to be able to specify the number of connections involving each object in the system. This can be modelled by a random graph with a specified degree sequence. Given a sequence $\boldsymbol{k} = (k_1, \ldots, k_n)$ of nonnegative integers with even sum, let $\mathcal{G}_{\boldsymbol{k}}$ denote the uniform probability space over all graphs with degree sequence $\boldsymbol{k}$. A special case is the model $\mathcal{G}_{n,k}$ of uniformly-random k-regular graphs on n vertices. Wormald gives an excellent survey on this model [86]. The edge probabilities in these models are not independent, which makes calculations more difficult compared with the classical models. One solution to this problem is to perform calculations in a related model known as the configuration model, introduced by Bollobás [14].

Some large real-world systems involve relationships between more than two objects. Such systems can be modelled using random hypergraphs. The binomial random hypergraph model $\mathcal{G}_{n,r,p}$ contains all r-uniform hypergraphs on n vertices, where each r-subset of vertices forms as an edge with probability p, independent of all other subsets. In contrast, $\mathcal{G}_{n,r,m}$ is the uniform probability space of all r-uniform hypergraphs on n vertices with m edges. This model is formed by selecting m elements from the set of all $\binom{n}{r}$ possible edges, uniformly. In this thesis we focus on the uniform probability space $\mathcal{G}_{r,\boldsymbol{k}}$ over the set $\mathcal{H}_r(\boldsymbol{k})$. We also consider the special case of random uniform regular hypergraphs $\mathcal{G}_{n,r,k}$, obtained by choosing an element of $\mathcal{H}_r(k, n)$ uniformly at random. To avoid confusion we write all our results in terms of the sets $\mathcal{H}_r(\boldsymbol{k})$, $\mathcal{H}_r(k, n)$, rather than the associated uniform probability spaces.

For sparse hypergraphs with given degree sequences, the most common analysis techniques use either the configuration model or the switching method. In this thesis we use the switching method in our proofs. This method is discussed in the next section.

1.3 Switching method

A classic approach that has been used in studying sparse graphs, and more recently hypergraphs, with specific degrees is known as the switching method. The switching method was introduced by McKay [60] to find the probability of specified subgraphs in random graphs with given degrees, under certain conditions. Since then, it has been used frequently to give asymptotic formulae for various graph classes, see for example [62, 68]. An extension of this method to hypergraphs has been also studied with suitable adjustment to consider the size of edges [12, 13, 23].

The basic idea of the switching method is to define a simple operation that transforms elements of one set into another, in order to estimate the ratio between the sizes of the two sets. Let (A_n) and (B_n) be two sequences of finite sets, indexed by n, and define an operation which takes an element of A_n and transforms it into an element of B_n in several possible ways. Suppose that for every $x \in A_n$ there are $N(x)$ ways to perform this operation from x, giving an element of B_n. Also suppose that every $y \in B_n$ can be created from an element of A_n by this operation in $N'(y)$ ways. If $N(x) = (1 + o(1))N$ and $N'(y) = (1 + o(1))N'$ for all $x \in A_n$ and $y \in B_n$ then by double-counting the number of pairs (x, y) which are related by the operation, we conclude that

$$\frac{|A_n|}{|B_n|} = \frac{N'}{N}(1 + o(1)).$$

This method is often used repeatedly, to remove "defects" one by one. As an example, we could define a switching to remove loops one at a time from non-simple graphs with a given degree sequence $\boldsymbol{k}$. Let S_ℓ be the number of graphs with degree sequence $\boldsymbol{k}$ and ℓ loops, (these are non-simple when $\ell \geq 1$). Suppose that we know a good estimate for $|S_0 \cup S_1 \cup \cdots \cup S_L|$ for some large number L, and we want an estimate for $|S_0|$, the number of (not necessarily simple) graphs with degree sequence $\boldsymbol{k}$ and no loops. Then we can write

$$\frac{|S_0 \cup \ldots \cup S_L|}{|S_0|} = \sum_{j=0}^{L} \prod_{i=0}^{j-1} \frac{|S_{i+1}|}{|S_i|}.$$

If we can obtain sufficiently good estimates for the ratios $|S_{i+1}|/|S_i|$ then we can deduce an approximate formula for $|S_0|$. For more advanced applications of the switching method, see Fack and McKay [27] or Hasheminezhad and McKay [39].

A switching operation may be defined directly on graphs or hypergraphs, replacing a set of target edges, usually to remove a defect. Alternatively, the switching argument may work with the corresponding configuration model. In Chapter 3 we apply the switching method on sparse uniform hypergraphs with given degrees. Here the switching operation is designed to remove forbidden edges one by one. We use the switching method again in Chapter 5, but there we work on edges of bipartite graphs.

Chapter 2

Background

2.1 Enumeration results for graphs and hypergraphs

We will discuss some literature on asymptotic enumeration related to topics we study in this thesis. The obvious starting point is asymptotic enumeration of graphs, since uniform hypergraphs generalise graphs and most previous studies started in the graph case.

2.1.1 Graph enumeration

The asymptotic enumeration of graphs with given degrees has a long history. Some of these results are stated in terms of matrices (possibly symmetric) with given row and column sums, using the adjacency matrix of the graphs. The earliest results focussed on the case of sparse regular graphs, starting with [75, 76]. Bender and Canfield [11] gave an asymptotic formula for the number of k-regular graphs on n vertices, with k a constant.

Next, researchers looked for asymptotic formulae for the number of simple graphs with a given degree sequence, not necessarily regular. Bollobás [14] gave a formula for the number of simple graphs with degree sequence $\boldsymbol{k}$, using a random model known as the configuration model. His argument holds whenever $k_{\max} \leq 2\sqrt{\log n}-1$. As we do not need the configuration model in our research, we do not define it here, but refer the reader to Wormald [86]. McKay [62] used the switching method to show that the same formula holds, with a different error term, when $k_{\max}^4 = o(M)$. Recall that $M = \sum_{i=1}^{n} k_i$. Then McKay and Wormald [68] used a more complicated switching to obtain a more precise formula which holds when $k_{\max}^3 = o(M)$. In particular, for k-regular degrees, their formula holds when $k^2 = o(n)$.

McKay [62] established the first asymptotic enumeration result for the number of simple graphs with a given degree sequence $\boldsymbol{k}$ (not necessarily regular) which avoid a certain set of edges. His result holds when $k_{\max}(k_{\max} + x_{\max}) = o(M^{1/2})$, where $x_{\max}$ is the maximum degree in the graph formed by the forbidden edges. In the k-regular case, McKay's formula holds when $k^3 = o(n)$. He proved that the number of simple graphs with given degree sequence $\boldsymbol{k}$ and contain no edge of X is

$$\frac{M!}{(M/2)!\, 2^{M/2} \prod_{i=1}^{n} k_i!} \exp\left(-\frac{M_2}{2M} - \frac{M_2^2}{4M^2} - \frac{\sum_{v_i v_j \in X} k_i k_j}{M} + O\left(\frac{\hat{\Delta}^2}{M}\right)\right), \quad (2.1.1)$$

when $k_{\max} \geq 1$, $\hat{\Delta} = 2 + k_{\max}(\frac{3}{2}k_{\max} + x_{\max} + 1) \leq \epsilon M$, for $\epsilon < 1/3$. This formula can be used to give an asymptotic expression for the probability that a random graph with given degrees contains a set of specified edges. This can be used to approximate the average number of subgraphs of a given type, such as spanning trees [34]. Bollobás and McKay [16] also provided the formula (2.1.1), with a different error term and under weaker conditions than [62].

McKay performed his calculations in the configuration model, and the expontial factor is the probability that a randomly chosen configuration is simple and contains no pairs corresponding to forbidden edges. Apart from the error term, the terms inside the exponential factor in (2.1.1) have an intuitive explanation as, respectively, the expected number of loops, double pairs, and forbidden pairs in a random configuration. This is a common feature of combinatorial enumeration formulae, where some "bad" events are rare, and hence are asymptotically independent Poisson random variables. If Z is a random variable which is asymptotically Poisson with mean λ then the probability that $Z = 0$ is asymptotically $e^{-\lambda}$. There are similar explanations for many terms that appear within the exponential factors in the enumeration formulae in this thesis.

All results mentioned above involve sparse degree sequences. The first enumeration result for dense degrees was given by McKay and Wormald [67]. They gave an asymptotic enumeration formula for the number of graphs with degree sequence $\boldsymbol{k}$, when $\boldsymbol{k}$ is nearly-regular and all degrees are roughly linear in n. Their proof involves the saddle-point method and complex integration. Barvinok and Hartigan [8] gave a formula which holds for a slightly wider range of dense degrees. Recently, Isaev and McKay [40] provided a new formula which generalises many existing dense enumeration results.

Returning to forbidden edges, in the dense regime, McKay [64] found an asymptotic enumeration formula for the number of simple graphs with given degree sequence $\boldsymbol{k}$ avoiding a certain set of edges X. This formula holds when the average degree is roughly linear, the degree sequence is close to uniform and $|X|$ is roughly linear in n: see [64] for more details. Again, this formula can be used to find the probability that a random graph with degree sequence $\boldsymbol{k}$ contains a given subgraph. See also McKay's survey article [63].

McKay and Wormald [68] observed that the same asymptotic formula holds in both the sparse and dense ranges, and conjectured that this formula should also hold in the gap between these ranges. This conjecture was recently proved by Liebenau and Wormald [57]. Their proof involved a completely new approach, using recurrence relations between sets of graphs with given degrees. This work is also discussed in the survey paper of Wormald [87].

Turning to bipartite graphs, consider a bipartite graph with a given vertex bipartition, where the vertices of the first part have degree sequence $\boldsymbol{a} = (a_1, \ldots, a_m)$ and the vertices of the second part have degree sequence $\boldsymbol{b} = (b_1, \ldots, b_n)$. Then, we must have $\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$. Write $M = \sum_{i=1}^m a_i$. The number of bipartite graphs on the given bipartition with these degrees is equal to the number of 0–1 matrices of order $m \times n$ with row and column sums given by, respectively, $\boldsymbol{a}$ and $\boldsymbol{b}$.

The estimation of the number of bipartite graphs with given degrees begins with the semiregular case. A bipartite graph is *semiregular* when all vertices in the same part have the same degree; that is, $a_i = a$ and $b_j = b$ for all $i \in [m], j \in [n]$ and a, b are positive integers. When a and b are constant, the first result for the number of semiregular bipartite graphs was given by Read [74], who treated the case when all vertices have degree 3. Everett and Stein [26] gave a formula in the semiregular case when a and b are both bounded. An asymptotic formula for arbitrary but bounded degree sequences $\boldsymbol{a}$, $\boldsymbol{b}$ was independently established by Békéssy et al. [9], Bender [10] and Wormald [85]. When the degrees are allowed to grow slowly, an asymptotic formula was provided by Mineev and Pavlov [70] and by Bollobás and McKay [16]. These results, which improved on a previous result due to O'Neil [73, Theorem 2.3], allow the maximum degree to grow as some fractional power of $\log M$. Then, using the switching method, McKay [61] provided the asymptotic number of bipartite graphs with given degree sequences $\boldsymbol{a}$, $\boldsymbol{b}$, when $\max\{a_i, b_j\} = o(M^{1/4})$. Both Bollobás and McKay's result [16, Theorem 5] and McKay's improvement [61] give an asymptotic formula for the number of bipartite graphs with given degrees which avoid a specified set of edges, under some conditions.

Using more complicated switchings, McKay and Wang [66] examined a wider range of degrees in the semiregular case. This was extended to arbitrary $\boldsymbol{a}$, $\boldsymbol{b}$ by Greenhill et al. [37], who proved that when $1 \le a_{\max} b_{\max} = o(M^{2/3})$, the number of $m \times n$ matrices with entries $0, 1$, row sums given by $\boldsymbol{a}$ and column sums given by $\boldsymbol{b}$ is, as $M \to \infty$,

$$\frac{M!}{\prod_{i=1}^m a_i! \prod_{j=1}^n b_j!} \exp\biggl(-\frac{A_2 B_2}{2M^2} - \frac{A_2 B_2}{2M^3} + \frac{A_3 B_3}{3M^3} - \frac{A_2 B_2 (A_2 + B_2)}{4M^4} \\ - \frac{{A_2}^2 B_3 + A_3 {B_2}^2}{2M^4} + \frac{{A_2}^2 {B_2}^2}{2M^5} + O\Bigl(\frac{a_{\max}^3 b_{\max}^3}{M^2} \Bigr) \biggr). \tag{2.1.2}$$

Here $a_{\max}$ and $b_{\max}$ denote the maximum entries of $\boldsymbol{a}$, $\boldsymbol{b}$, respectively, and for $k = 2, 3$,

$$A_k = \sum_{i=1}^m (a_i)_k, \qquad B_k = \sum_{j=1}^n (b_j)_k.$$

A less precise version of this result will help us to estimate the number of linear uniform hypergraphs in Chapter 5. This is sufficient for our argument in this thesis. Furthermore, Greenhill and McKay [36] provided an extension of equation (2.1.2) for all $m \times n$ matrices with nonnegative integer entries rather than 0–1 entries. For results on the asymptotic enumeration of dense bipartite graphs, see [7, 8, 18].

Next we mention a couple of results on the asymptotic enumeration of bipartite graphs with given degrees and with a lower bound on the girth. The *girth* of a graph is the length of the shortest cycle in the graph. Given $g = g(n)$ even, and $k = k(n)$, McKay, Wormald and Wysocka [69, Corollary 3] proved that the probability that a random k-regular bipartite graphs has girth at least g is

$$\exp\left(-\sum_{i=1}^{g/2-1} \frac{(k-1)^{2i}}{2i} + o(1) \right)$$

when $(k-1)^{2g-3} = o(n)$. Similar results were given by Lu and Székely [58], allowing some irregularity in the degrees but assuming that both sides of the bipartition have the same order. In the regular case, their result is not as strong as [69]. We remark that both [58, 69] gave analogous results for random regular graphs, but these are not relevant for our work. A bipartite graph is *half-regular* if the degrees on one side of the vertex bipartition are regular. In Chapter 5, we prove a similar result to [69, Corollary 3] for half-regular bipartite graphs with girth at least 6, as shown in Corollary 5.4.3.

2.1.2 Hypergraph enumeration

There are some analogues of the above asymptotic enumeration results for hypergraphs. However, hypergraphs are less well-studied than graphs. This thesis will provide some asymptotic results for uniform hypergraphs with a given degree sequence and some other conditions. Here we will survey the literature on enumerations of r-uniform hypergraphs with a given degree sequence. In some results the value of r is fixed, while in others r may grow slowly with n.

When $r \geq 3$ is constant, Dudek et al. [23] used a configuration model, in the sparse regime, to show that the number of simple r-uniform k-regular hypergraphs on n vertices is

$$\frac{(kn)!}{(kn/r)!\,(r!\,)^{kn/r}(k!\,)^n} \exp\left(-\tfrac{1}{2}\,(r-1)(k-1) + O((k/n)^{1/2} + k^2/n)\right) \tag{2.1.3}$$

assuming that $k = o(n^{1/2})$ if $r = 3$, or $k = o(n)$ if $r \geq 4$. The relative error is $o(1)$ when $k = o(n^{1/2})$. This generalises McKay's result in [62] for the regular graph case.

Blinovsky and Greenhill [12] proved an analogous result of equation (2.1.3) when $\boldsymbol{k}$ is an irregular degree sequence restricted by $k_{\max}^3 = o(M)$, treating r as constant. The latter result was extended by the same authors in [13, Corollary 2.3] to consider slowly-growing r, as restated below.

Lemma 2.1.1. [13, Corollary 2.3] *For $n \geq 3$, let $r = r(n) \geq 3$ and $\boldsymbol{k} = \boldsymbol{k}(n)$ be the degree sequence such that r divides M for infinitely many n. Suppose that $M \to \infty$ and $r^4 k_{\max}^3 = o(M)$ as $n \to \infty$. Then the number of simple r-uniform hypergraphs with degree sequence $\boldsymbol{k}$ is*

$$\frac{M!}{(M/r)!\,r!^{M/r}\prod_{i=1}^n k_i!} \exp\left(\frac{-(r-1)M_2}{2M} + O\left(\frac{r^4 k_{\max}^3}{M}\right)\right).$$

In Chapter 3, we rely on this formula to estimate the number of simple uniform hypergraphs with given degrees which contain no edges of a specified hypergraph, as stated in Theorem 3.1.1. This result is an analogue of McKay's result for graphs in [62].

In the dense case, Kuperberg et al. [54, Theorem 1.5] gave an asymptotic formula for $\mathcal{H}_r(k, n)$ which holds under conditions which imply that $r \geq c$ and k is bounded below by n^{c-1}, for some sufficiently-large positive constant c. During Kamčev's talk at the Random Structures and Algorithms conference in 2019, Kamčev, Liebenau and Wormald announced an asymptotic enumeration result for hypergraphs which covers the gap between the sparse and dense cases. This formula has just appeared in [49].

In recent work, Espuny Díaz et al. [25] estimated the probability that a random hypergraph from $\mathcal{H}_r(k, n)$ contains a fixed set of edges X. They also gave a formula for the expected number of copies of X in a random hypergraph from $\mathcal{H}_r(k, n)$, when $r \geq 3$ is fixed and k satisfies $k = \omega(1)$ and $k = o(n^{r-1})$.

To the best of our knowledge, there are no other results on asymptotic enumeration of uniform hypergraphs with given degree sequence $\boldsymbol{k}$ with a set of edges which must be included (or avoided): in particular, there are no prior results when $\boldsymbol{k}$ is irregular or $k_{\max} = O(1)$.

Lu and Szekely [58] described a method for proving asymptotic enumeration formulae based on the Lovász Local Lemma. They applied their method to a few applications, including the girth of random bipartite graphs, as mentioned earlier. To the best of our knowledge, their method has not been applied to enumerate hypergraphs with given degrees and other conditions. Since the switching method is designed to be very precise, we have chosen to use the switching method in this thesis.

2.1.3 Linear hypergraph enumeration

Note that a simple graph is a linear hypergraph since the intersection between any two edges contains at most one vertex. Next, we discuss some previous studies on asymptotic enumeration of linear hypergraphs.

The first formula for the number of linear hypergraph with a given degree sequence was proved by Blinovsky and Greenhill [13]. They determined the asymptotic number of r-uniform linear hypergraphs with given degree sequence $\boldsymbol{k}$ provided that $M \to \infty$ and $r^4 k_{\max}^4 (k_{\max} + r) = o(M)$ as $n \to \infty$, which is given by

$$\frac{M!}{(M/r)!\, r!^{M/r} \prod_{i=1}^{n} k_i!} \exp\left(-\frac{(r-1)M_2}{2\,M} - \frac{(r-1)^2\, M_2^2}{4\,M^2} + O\left(\beta\right) \right), \qquad (2.1.4)$$

where $\beta = r^4 k_{\max}^4 (k_{\max} + r)/M$, allowing r to grow with n. In Chapter 5 we extend this result by relaxing the condition to $r^4 k_{\max}^4 = o(M)$, as stated in Theorem 5.1.1.

Next, we mention two further enumeration results on linear hypergraphs, restricted by the number of edges rather than by degree sequence. Recently, McKay and Tian [65] gave an asymptotic formula for the number of uniform linear hypergraphs on n vertices and given number of edges. For $r = r(n) \geq 3$, $N = \binom{n}{r}$ and $m = m(n)$, they proved that the number of r-uniform linear hypergraph with n vertices and m edges is

$$\frac{N^m}{m!} \exp\left(-\frac{(r)_2^2 (m)_2}{4n^2} - \frac{(r)_2^3 (3r^2 - 15r + 20) m^3}{24n^4} + O\left(\frac{r^6 m^2}{n^3} \right) \right) \tag{2.1.5}$$

as $n \to \infty$, provided that $m = o(n^{3/2}/r^3)$. Rather than summing equation (2.1.4) over all possible degree sequences, McKay and Tian [65] applied a direct switching method which involved the analysis of several switchings in hypergraphs which preserve the number of edges. Furthermore, they gave a formula for the probability that a randomly chosen linear r-uniform hypergraph with n vertices and m edges contains a given set of edges, under similar conditions.

When the edges of a hypergraph do not all contain the same number of vertices, it is called a *non-uniform* hypergraph. The first enumeration result for non-uniform linear hypergraphs was given by Hasheminezhad and McKay [38]. They estimated the number of linear hypergraphs on n vertices with given number of edges of each size, such that maximum edge size is bounded by a constant and $m = o(n^{4/3})$.

A (finite) set system consists of a finite set V of elements and a finite collection of subsets of V, whose members are called blocks. This matches the definition of a hypergraph. A set system is a *Steiner system* $S(\ell, r, n)$, if $|V| = n$, every block has size r and each ℓ-subset of V is contained in exactly one block. Similarly, a *partial* $S(\ell, r, n)$ *Steiner system* is a set system where every block has size r and every ℓ-subset of V is contained in at most one block. So an r-uniform linear hypergraph is a partial $S(2, r, n)$ Steiner system. Grable and Phelps [33] proved an asymptotic formula for $\ln s(\ell, r, n)$, where $s(\ell, r, n)$ is the number of partial $S(\ell, r, n)$ Steiner systems and $\ell < r$. This also was proved later by Asratian and Kuzjurin [5] with a simpler approach depending on a previous result on the existence of perfect matchings in hypergraphs, proved by Frankl and Rödl [28].

2.2 Spanning subhypergraphs

2.2.1 Perfect matchings

Let n, k and r be positive integers. Let H be a sparse r-uniform k-regular hypergraph with n vertices. If H has a perfect matching then r must divide n, since the perfect matching has n/r edges. Therefore, the number of vertices n must be even in any graph containing a perfect matching. The number of perfect matchings in a graph is equal to the permanent of the adjacency matrix of the graph.

There is a huge literature on perfect matchings in random graphs, see for example Bollobás [15]. The most relevant result is that of Bollobás and McKay [16], who examined the case where $3 \leq k \leq (\log n)^{1/3}$ and gave an asymptotic expression for the first and second moment of perfect matchings in random k-regular graphs (and for random regular bipartite graphs) on n vertices. Janson [41] investigated the asymptotic distribution of the number of perfect matchings in a random k-regular graph, for $k \geq 3$ fixed, using the small subgraph conditioning method. Similar calculations also appeared in Robinson and Wormald [78].

Now we focus on results related to random hypergraphs. We start with classical models of random hypergraphs: random hypergraphs with a given number of edges, or binomial random hypergraphs where each possible r-subset of vertices is present as an edge with probability p, independently. When $r \geq 3$ is fixed, Schmidt and Shamir [79] proved that if m grows faster than $n^{3/2}$ then the probability that a random r-uniform hypergraph with m edges contains a perfect matching is $1 - o(1)$ as $n \to \infty$. This was developed later by Frieze and Janson [29] with an improved bound of $n^{4/3}$ on the number of edges required for existence of a perfect matching with probability $1 - o(1)$. Kim [50] gave a further improvement to $n^{\frac{6r-4}{5r-3}}$, when

$r \geq 3$ is fixed. In fact, finding the threshold for the property of containing a perfect matching is also known as Shamir's problem. Very recently, Kahn [47] gave an upper bound of $(n/r)\log n$ for this threshold, improving on the result of Johansson et al [45] which gave the threshold up to an unknown constant factor (depending on r). For $q = \binom{n}{r}$ and $V = \{v_1, \dots, v_n\}$, let $e_1, e_2, \dots e_q$ be a random ordering of the set of all of r-subsets of V. Then Kahn [48] also proved that a hypergraph H on V with edge set $\{e_1, \dots, e_T\}$ contains a perfect matching with high probability, where T is the "hitting time" defined as $T = \min\{t : e_1 \cup \cdots \cup e_t = V\}$.

However, the most relevant results for our work involve random uniform hypergraphs with a given degree sequence. The number of perfect matchings in k-regular r-uniform hypergraphs was analysed by Cooper et al. [19] when $r \geq 3$ and $k \geq 2$ are fixed integers. Combining [19, Lemma 3.1] with [19, (6.18)] implies that under these conditions, the expected number of perfect matchings in a random hypergraph in $\mathcal{H}_r(k, n)$ is

$$(1 + o(1))\, e^{(r-1)/2} \sqrt{r} \left(k \left(\frac{k-1}{k} \right)^{(r-1)(k-1)} \right)^{n/r}. \tag{2.2.1}$$

This expected value tends to infinity when $r < \log k / \big((k-1)\log(\frac{k}{k-1})\big) + 1$. To the best of our knowledge, there are no results on the expected number of perfect matchings in r-uniform regular hypergraphs which allow r to grow as a function of n. We provide such a formula when r grows sufficiently slowly, as stated in Corollary 4.1.1. We will prove this result in Section 4.1 by applying Corollary 3.1.2.

2.2.2 Loose Hamilton cycles

We focus on results related to Hamilton cycles in regular graphs and regular uniform hypergraphs. In 1992, Robinson and Wormald [77] proved that a random 3-regular graph has a Hamilton cycle with probability which tends to 1. Their proof introduced an analysis of variance technique which is now known as small subgraph conditioning. Two years later Robinson and Wormald [78] proved that a random k-regular graph contains a Hamilton cycle with probability $1 - o(1)$, for any constant $k \geq 3$. This argument used small subgraph conditioning again, but analysed the number of perfect matchings in random k-regular graphs. Later, Frieze et al. [30] applied the small subgraph conditioning method directly to the number of Hamilton cycles in random k-regular graphs, when $k \geq 3$ is constant. For more information on these results, and on the method used to prove them, see Janson [42] or Wormald [86]. Cooper et al. [20] proved that random k-regular graphs with $C < k < n/C$ contain a Hamilton cycle with probability $1 - o(1)$, for some sufficiently large constant C, and Krivelevich et al. [53] established the same result when $k > n^{1/2} \log n$.

On the other hand, in hypergraphs, the consecutive edges of a Hamilton cycle can intersect in more than one vertex. If successive edges overlap in precisely ℓ vertices then this is known as ℓ-overlapping Hamilton cycles. Here $1 \leq \ell \leq r - 1$. If an r-uniform hypergraph on n vertices contains an ℓ-overlapping Hamilton cycle then we must have $r - \ell$ divides n, since the number of edges of an ℓ-overlapping Hamilton cycle is $n/(r - \ell)$. Then, a loose Hamilton cycle is an 1-overlapping Hamilton cycle. Dudek et al. [24, Theorem 5] gave upper bounds on the degree threshold for existence of ℓ-overlapping Hamilton cycles in $\mathcal{H}_r(k, n)$. They conjectured that for $\ell \geq 2$, these bounds give the correct location for the existence threshold. This conjecture was confirmed by Espuny Díaz et al. [25] when $\ell \geq 2$.

When $r \geq 3$ and $k \geq 2$ are fixed integers, the number of loose Hamilton cycles in a random element of $\mathcal{H}_r(k, n)$ has been studied by Altman et al. [3]. Using the small subgraph conditioning method, Altman et al. confirmed the conjecture of Dudek et al. [24, Conjecture 1] that the degree threshold is constant when $\ell = 1$. Let $\mathcal{H}$ be chosen uniformly at random from $\mathcal{H}_r(k, n)$ and let Y be the number of loose Hamilton cycles in $\mathcal{H}$. In order to complete their argument, Altman et al. [3] needed to know the ratio between $\mathbb{E}(Y)$ and the expected number of loose Hamilton cycles in the corresponding configuration model. This ratio can be deduced from our formula provided in Corollary 4.2.1, which allows r to grow slowly with n. In fact, Altman et al. stated this formula as a conjecture in an earlier draft of their paper (on arXiv).

2.2.3 Spanning hypertrees

The number of spanning trees in a graph G, also called the *complexity* of G, is a very well-studied parameter. Greenhill et al. [34] gave an asymptotic formula for the average number of spanning trees in graphs with a given degree sequence, as long as the degree sequence is sufficiently sparse. This completed a sequence of papers beginning with McKay [59]: see the history described in [34].

In this thesis, we define a hypertree to be a connected hypergraph which contains no Berge cycles, as explained in Section 1.1. However, there are several different definitions of hypertrees in the literature. The classical definition of a tree is a connected graph with no cycle. In graphs, there is only one way to define cycles, since any pair of edges in a simple graph can overlap in at most one vertex. However, this is not the case for hypergraphs where a pair of distinct edges are allowed to intersect in more than one vertex. Consequently, there are various types of cycles in hypergraphs, which gives rise to different definitions of hypertrees and therefore gives different enumerations. Siu [82] gave a family of definitions of hypertrees, parameterised by the amount of overlap allowed between edges. Also, Jégou et al. explored a variety of cycles in [44]. Our definition of hypertrees matches the definition given by Selivanov [80] and Boonyasombat in [17]. It also matches what Siu calls "traditional hypertrees" [82, Section 1.2.1]: the other structures he studies contain 2-cycles, as he allows edges to overlap in more than one vertex.

Goodall and Mier [32] investigated spanning hypertrees in (non-random) 3-uniform hypergraphs, establishing some necessary conditions and some sufficient conditions for the existence of a spanning hypertree. They also proved that any Steiner triple system on n vertices has at least $\Omega((n/6)^{n/12})$ spanning hypertrees [32, Theorem 4]. A Steiner triple system can be viewed as a 3-uniform hypergraph such that every pair of distinct vertices is contained in exactly one edge. Applications of spanning hypertrees include the hypergraph analogue of the Steiner tree problem studied by Warme [84]. As far as we know, there is no prior work on the asymptotic number of spanning hypertrees in random uniform hypergraphs with given degrees. We discuss this enumeration in Section 4.3 and find the average number of spanning hypertrees in a simple uniform hypergraph as stated in Theorem 4.3.1.

Chapter 3

Sparse uniform hypergraphs with forbidden edges

For the graph case, McKay [62] gave an asymptotic formula for the number of simple graphs with a given degree sequence $\boldsymbol{k}$ and avoiding a certain set of edges, when the degree sequence $\boldsymbol{k}$ and the graph formed by the forbidden edges are both sufficiently sparse. Using this formula, it is possible to find the expected number of copies of a given subgraph in a random graph with a given degree sequence.

In this chapter we prove an analogous result for sparse r-uniform hypergraphs with given degrees, allowing slowly-growing r. As a corollary, we obtain an asymptotic formula for the probability that a random r-uniform hypergraph with degree sqeuence $\boldsymbol{k}$ contains a specified subhypergraph. Three applications of this corollary will be given in Chapter 4. All the material in this chapter has been published in [2].

3.1 Main results

Let $X = X(n)$ be a simple r-uniform hypergraph with vertex set V and degree sequence $\boldsymbol{x}$ of non-negative integers and edge set $\{e_1, e_2, \ldots, e_t\}$. By a slight abuse of notation, we also write X to denote its edge set. From now on, we refer to $X = \{e_1, e_2, \ldots, e_t\}$ as the set of *forbidden edges*. Let $\mathcal{H}_r(\boldsymbol{k}, X)$ be the set of all hypergraphs in $\mathcal{H}_r(\boldsymbol{k})$ which contain no edge of X. We estimate the size of $\mathcal{H}_r(\boldsymbol{k}, X)$ as stated below.

Theorem 3.1.1. *For $n \geq 3$, suppose that $r = r(n) \geq 3$. Let $\boldsymbol{k} = \boldsymbol{k}(n) = (k_1, \ldots, k_n)$ be a sequence of non-negative integers with maximum degree $k_{\max}$ and sum $M(\boldsymbol{k})$. We assume that r divides $M(\boldsymbol{k})$ for infinitely many values of n. Let $X = X(n)$ be a* *given simple r-uniform hypergraph with degree sequence $\boldsymbol{x}$ and with t edges. Suppose that $r^4\, k_{\max}^3 = o\left(M(\boldsymbol{k})\right)$ and $\rho = o(1)$, where*

$$\rho = \frac{t\, k_{\max}^3}{M(\boldsymbol{k})^2} + \frac{r\, t\, k_{\max}^4}{M(\boldsymbol{k})^3}.$$

Then the probability that a random hypergraph from $\mathcal{H}_r(\boldsymbol{k})$ contains no edge of X is $\exp\left(O(\rho)\right)$. Therefore, the number of simple r-uniform hypergraphs with degree sequence $\boldsymbol{k}$ containing no edge of X is

$$\begin{aligned}&|\mathcal{H}_r(\boldsymbol{k}, X)|\\ &= \frac{M(\boldsymbol{k})!}{(M(\boldsymbol{k})/r)!\, r!^{M(\boldsymbol{k})/r} \prod_{i=1}^n k_i!} \exp\left(\frac{-(r-1)M_2(\boldsymbol{k})}{2M(\boldsymbol{k})} + O\left(\frac{r^4\, k_{\max}^3}{M(\boldsymbol{k})} + \rho\right)\right).\end{aligned}$$

Our proof of this result uses the switching method, as described in Section 3.3. Observe that the first term of the exponential corresponds to the expected number of loops, as discussed in [13], while the error term is an upper bound of the other bad events including repeated edges or any forbidden edges. Using more complicated switchings, it should be possible to prove an asymptotic enumeration formula for an extended range of parameters, allowing a small but non-vanishing expected number of forbidden edges. The enumeration formula in this case would probably contain a term inside the exponential which is a sum over forbidden edges: note that the formula of McKay (2.1.1) has such a term. However, Theorem 3.1.1 is sufficient for our purposes, so we leave this extension for future work.

As a consequence of Theorem 3.1.1, we can obtain an asymptotic formula for the probability that a random element of $\mathcal{H}_r(\boldsymbol{k})$ contains all edges of X.

Corollary 3.1.2. *For $n \geq 3$ and $r = r(n) \geq 3$, let $\boldsymbol{k}$ and $k_{\max}$ be defined as above. Let $X = X(n)$ be a given simple r-uniform hypergraph with degree sequence $\boldsymbol{x}$ and* *with t edges, where $x_i \le k_i$ for all $i = 1, 2, \ldots, n$. Define*

$$\beta = \frac{r^4 k_{\max}^3}{M(\boldsymbol{k}-\boldsymbol{x})} + \frac{t\, k_{\max}^3}{M(\boldsymbol{k}-\boldsymbol{x})^2} + \frac{r\, t\, k_{\max}^4}{M(\boldsymbol{k}-\boldsymbol{x})^3},$$

and assume that $\beta = o(1)$. Then the probability that a random hypergraph from $\mathcal{H}_r(\boldsymbol{k})$ contains every edge of X is

$$\frac{(M(\boldsymbol{k})/r)_t\, r!^t \prod_{i=1}^n (k_i)_{x_i}}{(M(\boldsymbol{k}))_{rt}} \exp\left(\frac{r-1}{2}\left(\frac{M_2(\boldsymbol{k})}{M(\boldsymbol{k})} - \frac{M_2(\boldsymbol{k}-\boldsymbol{x})}{M(\boldsymbol{k}-\boldsymbol{x})}\right) + O\left(\beta\right)\right).$$

Proof. For a given r-uniform hypergraph X, the number of hypergraphs with degree sequence $\boldsymbol{k}$ which contain every edge of X is equal to the number of hypergraphs with degree sequence $\boldsymbol{k}-\boldsymbol{x}$ which contain no edge of X. Therefore, the probability that a random hypergraph $\mathcal{H} \in \mathcal{H}_r(\boldsymbol{k})$ contains X is

$$\mathbb{P}(X \subseteq \mathcal{H}) = \frac{|\mathcal{H}_r(\boldsymbol{k}-\boldsymbol{x}, X)|}{|\mathcal{H}_r(\boldsymbol{k})|}.$$

This probability can be computed using Theorem 3.1.1, leading to the stated expression with error term given by

$$\frac{r^4 k_{\max}^3}{M(\boldsymbol{k})} + \frac{r^4 d_{\max}^3}{M(\boldsymbol{k}-\boldsymbol{x})} + \frac{t\, d_{\max}^3}{M(\boldsymbol{k}-\boldsymbol{x})^2} + \frac{r\, t\, d_{\max}^4}{M(\boldsymbol{k}-\boldsymbol{x})^3} = O(\beta),$$

where $d_{\max} = \max\{k_j - x_j \,:\, j = 1, \ldots, n\}$. □

Corollary 3.1.2 can be used to give an asymptotic formula for the number of copies of a given hypergraph in a random element of $\mathcal{H}_r(\boldsymbol{k})$. In Chapter 4 we present three applications of this corollary, giving asymptotic expressions for the expected numbers of perfect matchings, loose Hamilton cycles and spanning hypertrees in random hypergraphs from $\mathcal{H}_r(\boldsymbol{k})$, under certain conditions.

3.2 Structure of our argument

We now outline our argument, and then describe the structure for the rest of this this chapter. Recall that $X = \{e_1, \dots, e_t\}$. For $i = 1, \dots, t$, let $\mathcal{F}_i \subseteq \mathcal{H}_r(\boldsymbol{k})$ be the set of hypergraphs in $\mathcal{H}_r(\boldsymbol{k})$ which contain the edge e_i. Define $\mathcal{F} = \cup_{i=1}^t \mathcal{F}_i$ and observe that $\mathcal{F}^c$ is the set of hypergraphs in $\mathcal{H}_r(\boldsymbol{k})$ which contain no edges of X. Define $\xi_i = |\mathcal{F}_i|/|\mathcal{F}_i^c|$ for $i = 1, \dots, t$. Then, for a random hypergraph from $\mathcal{H}_r(\boldsymbol{k})$,

$$\mathbb{P}(\mathcal{F}) \le \sum_{i=1}^{t} \mathbb{P}(\mathcal{F}_i) \le \sum_{i=1}^{t} \xi_i.$$

If $\sum_{i=1}^t \xi_i = o(1)$ then

$$1 \ge \mathbb{P}(\mathcal{F}^c) \ge 1 - \sum_{i=1}^{t} \xi_i = 1 - o(1),$$

and hence

$$\mathbb{P}(\mathcal{F}^c) = \exp\left(- \sum_{i=1}^{t} \xi_i + O\left(\left(\sum_{i=1}^{t} \xi_i \right)^2 \right) \right). \tag{3.2.1}$$

In Section 3.3 we use the method of switchings to obtain an upper bound on ξ_i. The proof of Theorem 3.1.1 is completed in Section 3.4, using Lemma 2.1.1.

The structure of our argument produces an error term of the form $O(tQ)$, where Q is a uniform upper bound on ξ_i for $i = 1, \dots, t$. Hence the conditions of our theorem also involve t, while the condition of McKay's theorem in the graph case involves $x_{\max}$. It is possible that a different argument could lead to an asymptotic enumeration formula for $|\mathcal{H}_r(\boldsymbol{k}, X)|$ which is expressed in terms of $x_{\max}$ instead of t, but we have not explored this here.

3.3 The switching on forbidden edges

Let $e_i \in X$ be given. We will now define and analyse a switching operation in order to obtain an upper bound on ξ_i.

3.3.1 Forward x-switching

Suppose that G^* is a hypergraph in $\mathcal{F}_i$. Define $S^* = S^*(G^*, i)$ to be the set of all 6-tuples $(z_1, z_2, y_1, y_2, f_1, f_2)$ defined as follows:

- z_1, z_2, y_1, y_2 are distinct vertices from V,
- e_i, f_1, f_2 are distinct edges of G^*, and
- $z_1, z_2 \in e_i$ and $y_j \in f_j$ for $j = 1, 2$.

Let G be a hypergraph resulting from a forward switching operation on G^* determined by the 6-tuple $(z_1, z_2, y_1, y_2, f_1, f_2)$. That is,

$$G = (G^* \setminus \{e_i, f_1, f_2\}) \cup \{g, g_1, g_2\},$$

where $g = (e_i \setminus \{z_1, z_2\}) \cup \{y_1, y_2\}$ and $g_j = (f_j \setminus \{y_j\}) \cup \{z_j\}$, for $j = 1, 2$. This switching is illustrated in Figure 3.1, following the arrow from left to right.

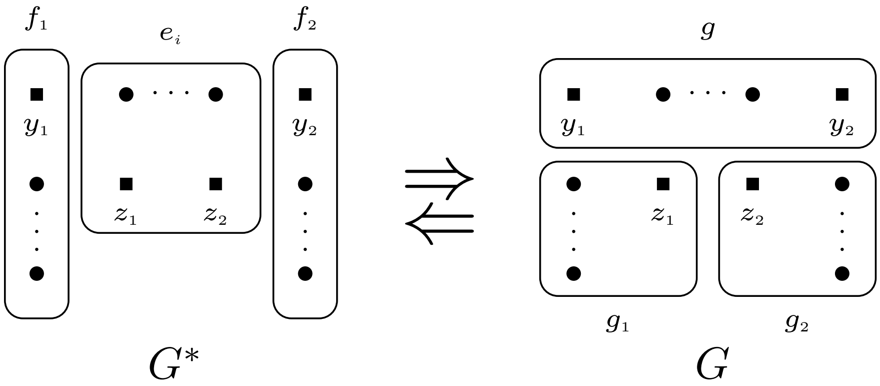

Figure 3.1: The x-switchings on forbidden edges

We say that the forward switching given by $(z_1, z_2, y_1, y_2, f_1, f_2)$ is *legal* if $G \in \mathcal{F}_i^c$, otherwise it is illegal. The next lemma describes illegal forward switchings from G^*.

Lemma 3.3.1. *Let $G^* \in \mathcal{F}_i$. Suppose that the* 6*-tuple $(z_1, z_2, y_1, y_2, f_1, f_2) \in S^*$ results in an illegal forward switching from G^*. Then at least one of the following holds:*

(I) *At least one of z_j, y_j belongs to both edges e_i and f_j, for some $j \in \{1, 2\}$.*

(II) *There is an edge $e \in G^* \setminus \{e_i, f_1, f_2\}$ such that either*

(a) $e \cap e_i = e_i \setminus \{z_1, z_2\}$ *and* $e \cap f_j = \{y_j\}$ *for* $j = 1, 2$*, or*

(b) $e \cap f_j = f_j \setminus \{y_j\}$ *and* $e \cap e_i = \{z_j\}$*, for some* $j \in \{1, 2\}$*.*

(III) *For some* $j \in \{1, 2\}$*,* $f_j \setminus \{y_j\} = e_i \setminus \{z_j\}$*.*

Proof. Suppose that $(z_1, z_2, y_1, y_2, f_1, f_2)$ is a 6-tuple in S^* which gives an illegal switching on $G^* \in \mathcal{F}_i$. This means that the resulting hypergraph G does not belong to $\mathcal{F}_i^c$. Then we have at least one of the following situations:

- G **contains a loop.** This implies that at least one new loop has been created accidentally at one of the vertices z_j, y_j for some $j \in \{1, 2\}$. If g_j contains a loop at z_j for some $j \in \{1, 2\}$, then we have $z_j \in f_j \cap e_i$ in G^*. Therefore, (I) holds. Similarly, if g has a loop at y_j for some $j \in \{1, 2\}$ then $y_j \in f_j \cap e_i$ in G^*, so (I) holds.
- G **contains a repeated edge.** Then the repeated edge must involve one of the new edges g, g_1, g_2, since $G^* \in \mathcal{F}_i$ is simple. Suppose that g has multiplicity greater than one in G. Then g also belongs to $G^* \setminus \{e_i, f_1, f_2\}$, as an edge of multiplicity 1. Hence, $g \setminus \{y_1, y_2\} = e_i \setminus \{z_1, z_2\}$. In addition, g intersects both f_1, f_2 in y_1, y_2, respectively. Hence (II)(a) holds. Similarly, if g_j is a multiple edge in G for some $j \in \{1, 2\}$ then g_j also belongs to $G^* \setminus \{e_i, f_1, f_2\}$, and (II)(b) holds.
- G **contains the edge** e_i**.** Since G^* is simple and z_1, z_2, y_1, y_2 are distinct vertices, either $g_1 = e_i$ or $g_2 = e_i$. Then $g_j \setminus \{z_j\}$ is the same set as $e_i \setminus \{z_j\}$, for some $j \in \{1, 2\}$. From the definition of g_j we also have $g_j \setminus \{z_j\} = f_j \setminus \{y_j\}$. Therefore (III) holds.

This completes the proof. $\square$

Next, we analyse forward switchings.

Lemma 3.3.2. *Let G^* be a hypergraph in $\mathcal{F}_i$ and let $S^* = S^*(G^*, i)$ be the set of 6-tuples $(z_1, z_2, y_1, y_2, f_1, f_2)$ defined earlier. If $r^4\, k_{\max}^3 = o(M(\boldsymbol{k}))$ then the number of 6-tuples in S^* which determine a legal switching is*

$$r(r-1)M(\boldsymbol{k})^2 \left(1 + O\left(\frac{k_{\max}^2 + r\, k_{\max}}{M(\boldsymbol{k})}\right)\right).$$

Proof. From the definition of S^*, it is obvious that the number of 6-tuples which determine a legal forward switching on G^* is bounded above by $|S^*|$. To find a lower bound on this number, we will subtract from $|S^*|$ an estimate for the number of 6-tuples which result in an illegal switching. These illegal 6-tuples are described in Lemma 3.3.1.

First, we will find an asymptotic expression for $|S^*|$. There are $r(r-1)$ choices for (z_1, z_2), as a pair of distinct vertices from e_i, and at most $M(\boldsymbol{k})^2$ choices for (y_1, y_2, f_1, f_2). Therefore,

$$|S^*| \leq r(r-1)M(\boldsymbol{k})^2.$$

Now we find a lower bound for $|S^*|$. First we need to choose an edge $f_1 \neq e_i$ and a vertex $y_1 \in f_1$ such that $y_1 \notin \{z_1, z_2\}$. The number of ways to choose (y_1, f_1) is at least

$$r\left(\frac{M(\boldsymbol{k})}{r} - 1\right) - 2k_{\max} = M(\boldsymbol{k})\left(1 + O\left(\frac{r + k_{\max}}{M(\boldsymbol{k})}\right)\right).$$

Next, the number of choices for (y_2, f_2) such that $f_2 \notin \{e_i, f_1\}$, $y_2 \in f_2$ and $y_2 \notin \{z_1, z_2, y_1\}$ is at least $M(\boldsymbol{k})\left(1 + O\left(\frac{r+k_{\max}}{M(\boldsymbol{k})}\right)\right)$.

Combining the bounds of $|S^*|$, we have

$$|S^*| = r(r-1)M(\boldsymbol{k})^2 \left(1 + O\left(\frac{r + k_{\max}}{M(\boldsymbol{k})}\right)\right). \tag{3.3.1}$$

Now, we estimate an upper bound for the number of 6-tuples in S^* which satisfy some property in Lemma 3.3.1.

For (I), suppose that $y_j \in e_i \cap f_j$ for some $j \in \{1,2\}$. There are $r(r-1)$ ways to choose (z_1, z_2) and at most $(r-2)\, k_{\max}\, M(\boldsymbol{k})$ choices for (y_1, y_2, f_1, f_2) satisfying this condition. Similarly, if $z_j \in e_i \cap f_j$ for some $j \in \{1,2\}$ then we have $r(r-1)$ choices for (z_1, z_2) and at most $(r-1)\, k_{\max}\, M(\boldsymbol{k})$ choices for (y_1, y_2, f_1, f_2). Therefore, the number of 6-tuples in S^* satisfying (I) is at most

$$2r(r-1)^2 k_{\max} M(\boldsymbol{k}).$$

For (II)(a), suppose that there exists an edge $e \in G^* \setminus \{e_i, f_1, f_2\}$ such that $e \cap e_i = e_i \setminus \{z_1, z_2\}$ and $e \cap f_j = \{y_j\}$ for $j = 1, 2$. There are $r(r-1)$ choices for (z_1, z_2) as distinct vertices in e_i. Then there are at most $k_{\max}$ choices for e and two ways to choose (y_1, y_2), as these are the two vertices in $e \setminus e_i$. We also have at most $k_{\max}^2$ choices for (f_1, f_2) as incident edges for y_1, y_2, respectively. Therefore, the number of 6-tuples in S^* satisfying (II)(a) is at most

$$2r(r-1)\, k_{\max}^3.$$

For (II)(b), suppose that there exists an edge $e \in G^* \setminus \{e_i, f_1, f_2\}$ such that $e \cap f_j = f_j \setminus \{y_j\}$ and $e \cap e_i = z_j$ for some $j \in \{1,2\}$. Then the number of choices for the 6-tuple satisfying (II)(b) is at most

$$r(r-1)\, k_{\max}^2 M(\boldsymbol{k}).$$

For (III), suppose that $f_j \setminus \{y_j\} = e_i \setminus \{z_j\}$ for some $j \in \{1,2\}$. Arguing as above, the number of 6-tuples in S^* satisfying this condition is at most

$$r(r-1)\, k_{\max}\, M(\boldsymbol{k}).$$

Combining these cases shows that the number of 6-tuples in S^* which give rise to an illegal switching is at most

$$\begin{aligned}
&2r(r-1)^2 k_{\max} M + 2r(r-1)\, k_{\max}^3 + r(r-1)(k_{\max}+1)\, k_{\max} M(\boldsymbol{k}) \\
&= r(r-1) M(\boldsymbol{k})^2\, O\left(\frac{(r+k_{\max})k_{\max}}{M(\boldsymbol{k})} + \frac{k_{\max}^3}{M(\boldsymbol{k})^2}\right) \\
&= r(r-1) M^2\, O\left(\frac{k_{\max}^2 + r\, k_{\max}}{M(\boldsymbol{k})}\right).
\end{aligned}$$

Subtracting this from (3.3.1) completes the proof. □

3.3.2 Reverse x-switching

Let $G \in \mathcal{F}_i^c$ be chosen at random and $S = S(G,i)$ be the set of all 6-tuples $(z_1, z_2, y_1, y_2, g_1, g_2)$ defined as follows:

- z_1, z_2, y_1, y_2 are distinct vertices in V,
- g_1, g_2 are distinct edges of G,
- $g_j \cap e_i = \{z_j\}$ for $j = 1, 2$, and
- there is an edge $g \in G$ which contains y_1, y_2 such that $g \cap e_i = e_i \setminus \{z_1, z_2\}$.

A reverse switching on G operating by the 6-tuple $(z_1, z_2, y_1, y_2, g_1, g_2)$ results in a hypergraph G^* defined by

$$G^* = (G \setminus \{g, g_1, g_2\}) \cup \{e_i, f_1, f_2\},$$

where $f_j = (g_j \setminus \{z_j\}) \cup \{y_j\}$, for $j = 1, 2$. This reverse switching is illustrated in Figure 3.1 by reversing the arrow. We say that the reverse switching is *legal* if $G^* \in \mathcal{F}_i$.

Every 6-tuple which gives rise to a legal reverse switching belongs to S. Therefore, it is sufficient to obtain an upper bound on $|S|$ in order to upper-bound the number of legal reverse switchings. Since $z_1, z_2 \in e_i$, there are at most $r(r-1)\, k_{\max}^2$ choices for (z_1, z_2, g_1, g_2) such that $z_j \in g_j$ for $j = 1, 2$. Also we have at most $2k_{\max}$ choices for (y_1, y_2) such that these vertices belong to an edge which intersects with

e_i in exactly $r-2$ vertices. Therefore, the number of legal reverse switchings which can be performed on G is at most

$$2r(r-1)k_{\max}^3. \tag{3.3.2}$$

Now, we can complete the proof of our main result of this chapter.

3.4 The proof of Theorem 3.1.1

Proof. We conclude from Lemma 3.3.2 and (3.3.2) that

$$\begin{aligned}\xi_i = \frac{|\mathcal{F}_i|}{|\mathcal{F}_i^c|} &\le \frac{2r\,(r-1)\,k_{\max}^3}{r(r-1)M^2\ \left(1+O\left(\left(k_{\max}^2 + r\,k_{\max}\right)/M\right)\right)}\\ &= O\left(\frac{k_{\max}^3}{M(\boldsymbol{k})^2} + \frac{r\,k_{\max}^4}{M(\boldsymbol{k})^3}\right).\end{aligned} \tag{3.4.1}$$

The assumptions of Theorem 3.1.1 imply that

$$\sum_{i=1}^{t} \xi_i = O\left(\frac{t\,k_{\max}^3}{M(\boldsymbol{k})^2} + \frac{r\,t\,k_{\max}^4}{M(\boldsymbol{k})^3}\right) = o(1).$$

Therefore, by (3.2.1) and (3.4.1),

$$\mathbb{P}(\mathcal{F}^c) = \exp\left(O\left(\frac{t\,k_{\max}^3}{M(\boldsymbol{k})^2} + \frac{r\,t\,k_{\max}^4}{M(\boldsymbol{k})^3}\right)\right). \tag{3.4.2}$$

We complete the proof by multiplying (3.4.2) by the value of $|\mathcal{H}_r(\boldsymbol{k})|$ given by Lemma 2.1.1. $\square$

Chapter 4

Applications

This chapter will discuss some applications of Corollary 3.1.2. All hypergraphs considered in this chapter will be simple. The chapter is divided into three sections, starting with estimating the expected number of perfect matchings in random k-regular r-uniform hypergraphs, when r and k grow sufficiently slowly with n. The next application finds the expected number of loose Hamilton cycles, again for regular uniform hypergraphs with slowly-growing k and r. These two applications, presented in Sections 4.1 and 4.2, have been published in [2] (together with the results from Chapter 3).

The final application of Corollary 3.1.2 is finding an asymptotic formula for the expected number of spanning hypertrees in a random r-uniform hypergraph with degree sequence $\boldsymbol{k}$, when the maximum degree is not too large and r may grow slowly. This result, presented in Section 4.3, has been submitted to a journal [1].

4.1 Perfect matchings

In the following corollary, we provide an asymptotic formula for the expected number of perfect matchings in a random element of $\mathcal{H}_r(k, n)$ when k, r grow sufficiently slowly as $n \to \infty$. Our formula matches the result of Cooper et al. [19] when k, r are constant, restated in (2.2.1), up to the error term.

Corollary 4.1.1. *For a positive integer $n \geq 3$, let $r = r(n) \geq 3$ be such that r divides n for infinitely many values of n. Let $k = k(n) \geq 2$ and let Z denote the* *number of perfect matchings in a hypergraph chosen randomly from $\mathcal{H}_r(k,n)$. Then, when $r^4 k^2 = o(n)$,*

$$\mathbb{E}(Z) = \sqrt{r}\left(k\left(\frac{k-1}{k}\right)^{(r-1)(k-1)}\right)^{n/r} \exp\left(\frac{r-1}{2} + O\left(\frac{r^4k^2}{n}\right)\right).$$

Proof. Let $\mathcal{H}$ be chosen uniformly at random from $\mathcal{H}_r(k,n)$. Then

$$\mathbb{E}(Z) = \sum_X \mathbb{P}(X \subseteq \mathcal{H}), \tag{4.1.1}$$

where the sum is over all possible perfect matchings X. Now let X be a fixed perfect matching with $t = n/r$ edges. Since X has degree sequence $\boldsymbol{x} = (1,1,\ldots,1)$, we have $\boldsymbol{k} - \boldsymbol{x} = (k-1,\ldots,k-1)$ and

$$M(\boldsymbol{k}-\boldsymbol{x}) = (k-1)n, \qquad M_2(\boldsymbol{k}-\boldsymbol{x}) = (k-1)(k-2)n.$$

Then, by Corollary 3.1.2,

$$\begin{aligned}&\mathbb{P}(X \subseteq \mathcal{H})\\ &= \frac{(kn/r)_t\, r!^t\ k^n}{(kn)_n} \exp\left(\frac{r-1}{2}\left(\frac{k\,(k-1)\,n}{k\,n} - \frac{(k-1)\,(k-2)\,n}{(k-1)\,n}\right) + O\left(\beta\right)\right),\end{aligned}$$

where

$$\beta = \frac{r^4k^3}{(k-1)n} + \frac{t\,k^3}{(k-1)^2n^2} + \frac{r\,t\,k^4}{(k-1)^3n^3} = O\left(\frac{r^4\,k^2}{n}\right).$$

The number of perfect matchings in the complete r-uniform hypergraph on n vertices is

$$\frac{n!}{(n/r)!\ r!^{n/r}}.$$

Hence by symmetry, using (4.1.1), we obtain

$$\mathbb{E}(Z) = \frac{n!}{(n/r)!\ r!^{n/r}} \frac{(\frac{kn}{r})!\ (kn-n)!\ r!^t\ k^n}{(\frac{kn}{r}-t)!\,(kn)!} \exp\left(\frac{r-1}{2} + O\left(\frac{r^4k^2}{n}\right)\right).$$

The factorial terms in this formula can be expanded by applying Stirling's formula, giving error term $O(r/n)$ which is absorbed by the stated error term. This completes the proof. $\square$

It may be possible to extend this argument to calculate the second moment $\mathbb{E}(Z^2)$. However, this would be more complicated due to the fact that there are many options for the union of two perfect matchings, even when they are disjoint. Calculating the second moment may lead to concentration results in some cases, though we note that when k and r are fixed constants, Z is not concentrated, at least in the configuration model. See [19, Lemma 3.1].

4.2 Loose Hamilton cycles

We use Corollary 3.1.2 to calculate the expected number of loose Hamilton cycles in a random element of $\mathcal{H}_r(k,n)$ when k and r grow sufficiently slowly with $n \to \infty$, as stated in Corollary 4.2.1. This result has recently been used by Altman et al. [3] in their analysis of the asymptotic distribution of the number of loose Hamilton cycles in $\mathcal{H}_r(k,n)$, when r and k are constants.

Corollary 4.2.1. *For $n \geq 3$, let $r = r(n) \geq 3$, $k = k(n) \geq 2$ and assume that r divides M and $r-1$ divides n for infinitely many values of n. Let $\mathcal{H}$ be chosen uniformly at random from $\mathcal{H}_r(k,n)$ and let Y be the number of loose Hamilton cycles in $\mathcal{H}$. If $r^4 k^2 = o(n)$ then*

$$\mathbb{E}(Y) = \sqrt{\frac{\pi}{2n}}(r-1)\left((k-1)(r-1)\left(\frac{rk-k-r}{rk-k}\right)^{(r-1)(rk-r-k)/r}\right)^{n/(r-1)}$$
$$\times \exp\left(\frac{(r-1)\,(rk-r-2)}{2(rk-r-k)} + O\left(\frac{r^4k^2}{n}\right)\right).$$

Proof. Let X be a fixed loose Hamilton cycle with $t = \frac{n}{r-1}$ edges and degree sequence $\boldsymbol{x}$, where t vertices have degree 2 in $\boldsymbol{x}$ and the remaining $n-t$ vertices have degree

1. With $\boldsymbol{k} = (k, \ldots, k)$, we have

$$D = (rk - r - k)\, t, \qquad D_2 = (k-2)_2\, t + (r-2)(k-1)_2\, t.$$

By Corollary 3.1.2, the probability that a random hypergraph in $\mathcal{H}_r(k, n)$ contains X is

$$\frac{(kn/r)_t\, r!^t \prod_{i=1}^{n} (k)_{x_i}}{(kn)_{rt}} \exp\left(\frac{r-1}{2} \left(\frac{M_2(\boldsymbol{k})}{M(\boldsymbol{k})} - \frac{M_2(\boldsymbol{k}-\boldsymbol{x})}{M(\boldsymbol{k}-\boldsymbol{x})} \right) + O\left(\beta\right) \right),$$

where

$$\beta = \frac{r^4 k^3}{(rk-r-k)\, t} + \frac{k^3}{(rk-r-k)^2\, t} + \frac{r\, k^4}{(rk-r-k)^3\, t^2} = O\left(\frac{r^4 k^2}{n} \right),$$

using the fact that $rk/(rk - r - k) = O(1)$. Next, we have

$$\tfrac{1}{2} \left(\frac{M_2(\boldsymbol{k})}{M(\boldsymbol{k})} - \frac{M_2(\boldsymbol{k}-\boldsymbol{x})}{M(\boldsymbol{k}-\boldsymbol{x})} \right) = \frac{(r-1)\,(rk-r-2)}{2(rk-r-k)}.$$

The number of loose Hamilton cycles in the complete r-uniform hypergraph on n vertices is

$$\frac{(r-1)\; n!}{2n\,(r-2)!^t}.$$

Therefore, by symmetry,

$$\begin{aligned} \mathbb{E}(Y) = \frac{(r-1)\; n!}{2n\,(r-2)!^t}\, &\frac{(kn/r)!\; r!^t\; (kn-rt)! \prod_{i=1}^{n} (k)_{x_i}}{(kn/r-t)!\; (kn)!} \\ &\times \exp\left(\frac{(r-1)\,(rk-r-2)}{2(rk-k-r)} + O\left(\frac{r^4 k^2}{n} \right) \right). \end{aligned} \tag{4.2.1}$$

Observe that $\prod_{i=1}^{n}(k)_{x_i} = (k(k-1))^t\, k^{(r-2)t}$. The factor outside the exponential can be estimated using Stirling's formula, giving

$$\frac{(r-1)\,n!\;(k(k-1))^t\;k^{(r-2)t} r!^t\;(kt(r-1)-rt)!\;((kt(r-1)/r)!}{2n\,(r-2)!^t\;(kt(r-1)/r-t)!\;(kt(r-1))!}$$

$$= \sqrt{\frac{\pi}{2n}}(r-1)\left((k-1)\,(r-1)\;\left(\frac{rk-k-r}{rk-k}\right)^{(r-1)(rk-k-r)/r}\right)^t$$
$$\times \exp\left(O\left(\frac{1}{n}+\frac{r}{kn-rt}\right)\right).$$

The proof is completed by combining this expression with (4.2.1), since the error term from (4.2.1) dominates. $\square$

4.3 Spanning hypertrees

The aim of this section is to estimate the average number of spanning hypertrees in r-uniform hypergraphs with a given degree sequence $\boldsymbol{k}$, when r and the maximum degree are not too large.

We say that a sequence of n positive integers $\boldsymbol{x} = (x_1, \ldots, x_n)$ is a *suitable* degree sequence for a hypertree in $\mathcal{H}_r(\boldsymbol{k})$ if $x_i \le k_i$ for all $i \in [n]$ and $\sum_{i=1}^{n} x_i = rt$ where $t = \frac{n-1}{r-1}$ is the number of edges in a hypertree on $[n]$. Denote by $\mathcal{T}$ the set of all r-hypertrees on n vertices. Then

$$|\mathcal{T}|= \frac{(n-1)!\;n^{t-1}}{t!\;(r-1)!^t},$$

generalising Cayley's formula. This result was given by Selivanov [80], see also [52]. Alternative proofs using generalisations of Prüfer codes were given in [55, 81, 83]. A more general result was proved by Siu [82, Theorem 2.1] using a different definition of hypertrees, where edges are added consecutively and a new edge may overlap a preceding edge in d vertices. Our definition of hypertree corresponds to the case $d = 1$.

For a suitable degree sequence $\boldsymbol{x}$, define

$$\mathcal{T}_{\boldsymbol{x}} = \{T \in \mathcal{T} : T \text{ has degree sequence } \boldsymbol{x}\}.$$

Bacher [6, Theorem 1.1] proved that

$$|\mathcal{T}_{\boldsymbol{x}}| = \frac{(r-1)(n-2)!}{(r-1)!^t \prod_{i=1}^{n}(x_i-1)!}. \tag{4.3.1}$$

This generalises a formula given by Moon [71] in the case of graphs.

Denote by $\tau_{\boldsymbol{k}}^{(r)}$ the number of spanning r-hypertrees in a hypergraph $\mathcal{H}$ chosen uniformly at random from $\mathcal{H}_r(\boldsymbol{k})$. If $\boldsymbol{k}$ is a regular degree sequence with $k_j = k$ for all $j \in [n]$ then we write $\tau_{n,k}^{(r)}$ for $\tau_{\boldsymbol{k}}^{(r)}$. Define

$$k = \frac{1}{n}\sum_{j=1}^{n} k_i, \qquad \hat{k} = \left(\prod_{i=1}^{n} k_i\right)^{1/n}$$

and write

$$F^{(r)}(k,\hat{k}) = \frac{(k-1)^{\frac{1}{2}}(r-1)}{n(kr-k-r)^{\frac{r+1}{2(r-1)}}} \left(\frac{\hat{k}\,(r-1)^{k/r}(k-1)^{k-1}}{k^{\frac{kr-k}{r}}(kr-k-r)^{\frac{kr-k-r}{r(r-1)}}}\right)^n.$$

Our result for $\tau_{\boldsymbol{k}}^{(r)}$ is stated below.

Theorem 4.3.1. *For $n \geq 3$, let $r = r(n) \geq 3$ be an integer number, and let $\boldsymbol{k} = \boldsymbol{k}(n) = (k_1, \ldots, k_n)$ be a sequence of positive integers with maximum $k_{\max}$. Assume that r divides kn and $r-1$ divides $n-1$ for infinitely many values of n, and perform asymptotics with respect to n only along these values. If $r^4\, k_{\max}^3 = o(M)$ then the average number of spanning hypertrees in a simple r-uniform hypergraph* *with degree sequence $\boldsymbol{k}$ is*

$$\mathbb{E}\,\tau_{\boldsymbol{k}}^{(r)} = F^{(r)}(k,\hat{k})\ \exp\left(\frac{kr-r-1}{2(k-1)} - \frac{kr-r-2k+1}{2k(k-1)^2 n}\sum_{i=1}^{n}(k_i-k)^2 + O\left(\frac{r^5\,k_{\max}^3}{(kr-k-r)n}\right)\right).$$

We note that this theorem holds only when $r \geq 3$ and does not capture the correct asymptotic expression in the case of graphs given by Greenhill et al. [34]: when $r = 2$, the factor $F^{(2)}(k,\hat{k})$ is correct but the exponential factor is different. This reflects the fact that simplicity for hypergraphs is not equivalent to conditioning on the event "no 1-cycles and no 2-cycles", as is the case for graphs.

Furthermore, we remark that the conclusion of Theorem 4.3.1 also holds if some entries of $\boldsymbol{k}$ equal zero, as then both $\tau_{\boldsymbol{k}}^{(r)}$ and $F^{(r)}(k,\hat{k})$ equal zero.

For k-regular r-uniform hypergraphs, we immediately obtain the following corollary.

Corollary 4.3.2. *For $n \geq 3$, let $r = r(n) \geq 3$ and $k = k(n)$ be positive integers. Assume that r divides kn and $r-1$ divides $n-1$ for infinitely many values of n, and perform asymptotics with respect to n only along these values. If $r^4\,k^2 = o(n)$ then the average number of spanning hypertrees in an r-uniform k-regular hypergraph is*

$$\mathbb{E}\,\tau_{n,k}^{(r)} = F^{(r)}(k,k)\ \exp\left(\frac{kr-r-1}{2(k-1)} + O\left(\frac{r^5\,k^3}{(kr-k-r)n}\right)\right).$$

This corollary was used by Greenhill et al. [35] to complete their analysis of the asymptotic distribution of the number of spanning trees in $\mathcal{H}_r(k,n)$, when r and k are fixed constants.

We follow the approach used by Greenhill et al. [34] in the graph case. For a given r-uniform hypertree T on vertex set $V = \{v_1,\ldots,v_n\}$, we can apply Corollary 3.1.2 to find the probability that a random element of $\mathcal{H}_r(\boldsymbol{k})$ contains T. By summing over all hypertrees with a given degree sequence $\boldsymbol{x}$, we obtain the expected number of spanning hypertrees with degree sequence $\boldsymbol{x}$ in a random element of $\mathcal{H}_r(\boldsymbol{k})$. Finally, by summing over all suitable degree sequences we complete the proof of Theorem 4.3.1.

Observe that the asymptotic formula given in Corollary 3.1.2 depends only on r, $\boldsymbol{k}$ and $\boldsymbol{x}$ (up to the stated error term), and not on the specific edges of X. In contrast, the corresponding formula of McKay [62, Theorem 4.6] which was used by Greenhill et al. [34] in their enumeration of the average number of spanning trees in graphs with given degrees, has terms which depend on the edges of X. This leads to differences in the calculation in the hypergraph case, as we do not have to average over all trees with a given degree sequence as in [34].

4.3.1 The proof of Theorem 4.3.1

Recall that k is the average of the elements of $\boldsymbol{k}$, $M(\boldsymbol{k})$ is the sum of entries of $\boldsymbol{k}$, and $M_2(\boldsymbol{k}) = \sum_{i=1}^n (k_i)_2$. Suppose that r divides $M(\boldsymbol{k})$ for infinitely many values of n and take n to infinity along these values. Suppose that $\boldsymbol{x} = (x_1, \ldots, x_n)$ is a suitable degree sequence.

Corollary 4.3.3. *Let $n \geq 3$, $r = r(n) \geq 3$ be integers and $\boldsymbol{k} = \boldsymbol{k}(n)$ be a sequence of positive integers. Let T be an r-hypertree with degree sequence $\boldsymbol{x}$ and $t = \frac{n-1}{r-1}$ edges, where $x_i \leq k_i$ for all $i \in [n]$. Define*

$$\lambda_0 = \frac{r-1}{2kn} \sum_{i=1}^n (k_i)_2, \qquad \lambda(\boldsymbol{x}) = \frac{(r-1)^2}{2(kr-k-r)n + 2r} \sum_{i=1}^n (k_i - x_i)_2.$$

If $r^4\, k_{\max}^3 = o(M)$ then the probability that a random hypergraph from $\mathcal{H}_r(\boldsymbol{k})$ contains T is

$$\mathbb{P}(T) = \frac{(kn/r)_t\, r!^t \prod_{i=1}^n (k_i)_{x_i}}{(kn)_{rt}} \exp\left(\lambda_0 - \lambda(\boldsymbol{x}) + O\left(\frac{r^5 k_{\max}^3}{(kr-k-r)n}\right)\right).$$

Proof. This follows by direct application of Corollary 3.1.2, using the fact that $M(\boldsymbol{k}) = kn$ and $M(\boldsymbol{k} - \boldsymbol{x}) = kn - rt$. □

Define $\tau_{\boldsymbol{k}}^{(r)}(\boldsymbol{x})$ as the number of r-hypertrees with degree sequence $\boldsymbol{x}$ in a random hypergraph in $\mathcal{H}_r(\boldsymbol{k})$. Hence, using Corollary 4.3.3 and linearity of expectation, we have

$$\begin{aligned}\mathbb{E}\,\tau_{\boldsymbol{k}}^{(r)}(\boldsymbol{x}) &= \sum_{T\in\mathcal{T}_{\boldsymbol{x}}} \mathbb{P}(T)\\ &= \frac{(kn/r)_t\, r!^t\, \prod_{i=1}^{n}(k_i)_{x_i}\,|\mathcal{T}_{\boldsymbol{x}}|}{(kn)_{rt}} \exp\left(\lambda_0 - \lambda(\boldsymbol{x}) + O\left(\frac{r^5 k_{\max}^3}{(kr-k-r)n}\right)\right)\end{aligned}$$

since the formula from Corollary 4.3.3 depends only on $\boldsymbol{x}$ and not on the edges of T. Substituting from (4.3.1) gives

$$\begin{aligned}&\mathbb{E}\,\tau_{\boldsymbol{k}}^{(r)}(\boldsymbol{x})\\ &= \frac{(kn/r)_t\, r^t\,(r-1)\,(n-2)!}{(kn)_{rt}}\left(\prod_{i=1}^{n}\frac{(k_i)_{x_i}}{(x_i-1)!}\right)\exp\left(\lambda_0-\lambda(\boldsymbol{x})+O\left(\frac{r^5k_{\max}^3}{(kr-k-r)n}\right)\right).\end{aligned}$$

Now, we multiply and divide by $\binom{(k-1)n}{t-1}$, then by rearranging the result and taking the sum over all possible suitable degree sequences $\boldsymbol{x}$, we obtain

$$\begin{aligned}&\mathbb{E}\,\tau_{\boldsymbol{k}}^{(r)}\\ &= D_{\boldsymbol{k}}^{(r)}\sum_{\boldsymbol{x}}\left\{\frac{\prod_{i=1}^{n}\binom{k_i-1}{x_i-1}}{\binom{(k-1)n}{t-1}}\exp\left(g(\boldsymbol{x})+O\left(\frac{r^5k_{\max}^3}{(kr-k-r)n}\right)\right)\right\},\end{aligned} \tag{4.3.2}$$

where $g(\boldsymbol{x}) = \lambda_0 - \lambda(\boldsymbol{x})$ and

$$D_{\boldsymbol{k}}^{(r)} = \frac{(kn/r)_t\, r^t\,(r-1)\,(n-2)!\ \hat{k}^n}{(kn)_{rt}}\binom{(k-1)n}{t-1}.$$

Next, we work on $D_{\boldsymbol{k}}^{(r)}$. By definition of t,

$$\begin{aligned}D_{\boldsymbol{k}}^{(r)} &= \frac{r^t\,\hat{k}^n\, n!\ (kn/r)!\ ((k-1)n)!}{n\,(\frac{kn-rt}{r})!\ (kn)!\quad t!}\\ &= \sqrt{\frac{k-1}{t(kn-rt)}}\ \frac{\hat{k}^n\, n^{kn/r}\,(k-1)^{(k-1)n}}{k^{(kr-k)n/r}\,(kn-rt)^{kn/r-t}\ t^t}\exp\left(O\left(\frac{r}{kn-rt}\right)\right)\end{aligned}$$

using Stirling's formula. Since $kn - rt = t(kr - k - r)\, e^{O(k/t)}$ and $t^{-1} = \frac{r-1}{n}\, e^{O(1/n)}$, it follows that

$$D_{\boldsymbol{k}}^{(r)} = F^{(r)}(k, \hat{k})\ \exp\left(O\left(\frac{r+k}{kn - rt}\right)\right). \tag{4.3.3}$$

Hence, since the error term in (4.3.3) is dominated by the error term in (4.3.2),

$$\mathbb{E}\,\tau_{\boldsymbol{k}}^{(r)} = F^{(r)}(k, \hat{k}) \sum_{\boldsymbol{x}} \frac{\prod_{i=1}^{n} \binom{k_i - 1}{x_i - 1}}{\binom{(k-1)n}{t-1}}\ \exp\left(g(\boldsymbol{x}) + O\left(\frac{r^5 k_{\max}^3}{(kr - k - r)n}\right)\right).$$

We only need to find the sum over $\boldsymbol{x}$ in this expression in order to prove the main result. This sum can be estimated using a similar approach to [34].

First, a slight generalisation of [34, Lemma 5.1] is stated below, in our notation.

Lemma 4.3.4. [34, Lemma 5.1] *Partition $[(k-1)n]$ into n sets $A_1, \ldots, A_n$, where $|A_i| = k_i - 1$ for $i = 1, \ldots, n$. Let C be a subset of $[(k-1)n]$ of size $t - 1$, chosen uniformly at random. Define a random vector $\boldsymbol{X} = \boldsymbol{X}(C) = (X_1, \cdots, X_n)$ by $X_j = |A_j \cap C| + 1$. Then*

$$\mathbb{E}\tau_{\boldsymbol{k}}^{(r)} = F^{(r)}(k, \hat{k})\ \mathbb{E} \exp\left(g(\boldsymbol{X}) + O\left(\frac{r^5 k_{\max}^3}{(kr - k - r)n}\right)\right).$$

The expectation of $\exp(g(\boldsymbol{X}))$ can be easily determined by computing $e^{\mathbb{E}g(\boldsymbol{X})}$ after proving that $\mathbb{E}(e^{g(\boldsymbol{X})}) \sim e^{\mathbb{E}g(\boldsymbol{X})}$. This can be done with the assistance of [34, Corollary 2.2], restated below.

Lemma 4.3.5. [34, Corollary 2.2] *Let $\binom{[N]}{s}$ be the set of s-subsets of $\{1, \ldots, N\}$ and let $h : \binom{[N]}{s} \to \mathbb{R}$ be given. Let C be a uniformly random element of $\binom{[N]}{s}$. Suppose that, for any $A, A' \in \binom{[N]}{s}$ with $s - 1$ elements in common, there exists $\alpha \geq 0$ such that*

$$|h(A) - h(A')| \leq \alpha.$$

Then

$$\mathbb{E}\exp(h(C)) = \exp\left(\mathbb{E}h(C) + K\right), \tag{4.3.4}$$

where K is a real constant such that $0 \le K \le \frac{1}{8}\min\{s, N-s\}\alpha^2$. Furthermore, for any real $z > 0$,

$$\Pr(|h(C) - \mathbb{E}h(C)| \ge z) \le \exp\left(\frac{-2z^2}{\min\{s, N-s\}\alpha^2}\right).$$

Two suitable degree sequences $\boldsymbol{x}, \boldsymbol{x}^{'}$ are *adjacent* if they are different in two entries i, j such that $x'_i = x_i + 1$, $x'_j = x_j - 1$. These sequences, respectively, correspond to two sets $A, A' \in \binom{[(k-1)n]}{t-1}$ with $t-2$ vertices in common.

Lemma 4.3.6.

$$\mathbb{E}\, e^{g(\boldsymbol{X})} = \exp\left(\mathbb{E}\, g(\boldsymbol{X}) + O\left(\frac{r^3\, k_{\max}^2}{(kr-k-r)^2\, n}\right)\right).$$

Proof. For adjacent suitable degree sequences $\boldsymbol{x}, \boldsymbol{x}'$ and from definition of $g(\boldsymbol{x})$ we have

$$\begin{aligned} |g(\boldsymbol{x}) - g(\boldsymbol{x}')| &= |\lambda(\boldsymbol{x}) - \lambda(\boldsymbol{x}')| \\ &= \frac{(r-1)^2 |2(k_i - x_i - 1) - (k_j - x_j)|}{2(kr-k-r)n + 2r} \\ &= O\left(\frac{r^2\, k_{\max}}{(kr-k-r)n}\right). \end{aligned}$$

Therefore we can apply Lemma 4.3.5 where

$$h(C) = g(\boldsymbol{X}(C)), \quad N = (k-1)\,n, \quad s = t-1 \text{ and } \alpha = O\Big(\frac{r^2\, k_{\max}}{(kr-k-r)n}\Big).$$

Since $N - s = (kr - k - r)t + 1 > s$, we have

$$K = O\left(\frac{t\, r^4 k_{\max}^2}{(kr-k-r)^2\, n^2}\right) = O\left(\frac{r^3\, k_{\max}^2}{(kr-k-r)^2\, n}\right).$$

This completes the proof. $\square$

The distribution of Lemma 4.3.4 is called a multivariate hypergeometric distribution with parameters $(t-1, \boldsymbol{k})$ as defined in [46, equation (39.1)]. Therefore, for non-negative integers a, b and using [46, equation (39.6)] we can compute the expectation of $(X_j - 1)_a$ as

$$\mathbb{E}\left((X_j - 1)_a\right) = \frac{(t-1)_a}{((k-1)n)_a}\,(k_j - 1)_a. \tag{4.3.5}$$

We use this expression to estimate $\mathbb{E}\left(g(\boldsymbol{X})\right)$ as follows.

Lemma 4.3.7.

$$\mathbb{E}\, g(\boldsymbol{X}) = \frac{kr - r - 1}{2(k-1)} - \frac{kr - r - 2k + 1}{2k(k-1)^2 n}\sum_{i=1}^{n}(k_i - k)^2 + O\left(\frac{r\, k_{\max}}{kn}\right).$$

Proof. Recall $g(\boldsymbol{x}) = \lambda_0 - \lambda(\boldsymbol{x})$, where λ_0 and $\lambda(\boldsymbol{x})$ are defined in Corollary 4.3.3. We restate $\lambda(\boldsymbol{X})$ as

$$\lambda(\boldsymbol{X}) = \frac{r-1}{2(kn - rt)}\sum_{i=1}^{n}\Big((k_i - 1)_2 - 2(k_i - 2)(X_i - 1) + (X_i - 1)_2\Big). \tag{4.3.6}$$

Applying (4.3.5) on the expected value of the summand in (4.3.6) implies

$$(k_i - 1)_2 - 2(k_i - 1)_2 \frac{t-1}{(k-1)n} + (k_i - 1)_2 \frac{(t-1)_2}{((k-1)n)_2}. \tag{4.3.7}$$

Taking out a common factor and using the identity $t-1=rt-n$, (4.3.7) can be rewritten as

$$\begin{aligned}&\frac{(k_i-1)_2}{((k-1)n)_2}\Big(((k-1)n)_2-2(rt-n)(kn-n-1)+(rt-n)_2\Big)\\&=\frac{(kn-rt)(kn-rt-1)\,(k_i-1)_2}{((k-1)n)_2}.\end{aligned}$$

Substuting this into (4.3.6), the expected value of $\lambda(\boldsymbol{X})$ is

$$\frac{(r-1)(kn-rt-1)}{2((k-1)n)_2}\sum_{i=1}^{n}(k_i-1)_2=\frac{kr-k-r}{2(k-1)^2n}\sum_{i=1}^{n}(k_i-1)_2+O\left(\frac{rk_{\max}}{kn}\right).$$

As a result, the expectation of $g(\boldsymbol{X})$ is

$$\mathbb{E}g(\boldsymbol{X})=\frac{r-1}{2kn}\sum_{i=1}^{n}(k_i)_2-\frac{kr-k-r}{2(k-1)^2n}\sum_{i=1}^{n}(k_i-1)_2+O\left(\frac{rk_{\max}}{kn}\right). \tag{4.3.8}$$

The first sum in this equation can be written as $k(k-1)n+\sum_{i=1}^{n}(k_i-k)^2$, while the second sum is $(k-1)(k-2)n+\sum_{i=1}^{n}(k_i-k)^2$. Substituting these into (4.3.8) and simplifying the result will complete the proof of this lemma. □

Finally, we can combine these results to prove the main result of this section.

Proof of Theorem 4.3.1. Substitution from Lemma 4.3.6 and Lemma 4.3.7 into the expression of Lemma 4.3.4 proves the required result, with combined error term

$$O\left(\frac{r^5k_{\max}^3}{(kr-k-r)n}+\frac{r^3k_{\max}^2}{(kr-k-r)^2n}+\frac{rk_{\max}}{kn}\right)=O\left(\frac{r^5k_{\max}^3}{(kr-k-r)n}\right)$$

as stated. □

Chapter 5

Linear uniform hypergraphs with given degrees

Recall that $\mathcal{H}_r(\boldsymbol{k})$ is the set of r-uniform simple hypergraphs with given degree sequence $\boldsymbol{k}$ and $\mathcal{L}_r(\boldsymbol{k})$ is the set of all linear hypergraphs in $\mathcal{H}_r(\boldsymbol{k})$. Blinovsky and Greenhill [13] gave an asymptotic formula for the number of linear r-uniform hypergraphs with degree sequence $\boldsymbol{k}$, which holds as long as $(r + k_{\max})r^4k_{\max}^4 = o(M)$. Our aim in this chapter is to make a (modest) improvement in the range of applicability of this formula, as detailed in Theorem 5.1.1 below. This improved range of r and $k_{\max}$ makes the asymptotic formula more useful.

5.1 Improved linear hypergraph enumeration result

In this chapter we will prove the following result.

Theorem 5.1.1. *Let $n \geq 3$ and $r = r(n) \geq 3$ be integers. Define $\boldsymbol{k} = \boldsymbol{k}(n) = (k_1, \ldots, k_n)$ to be a sequence of nonnegative integers with maximum value $k_{\max}$. Denote by M the sum of enteries of $\boldsymbol{k}$ which must be divisible by r. Suppose that $M \to \infty$ and $r^4k_{\max}^4 = o(M)$ for sufficiently large n tends to infinity. Then*

$$|\mathcal{L}_r(\boldsymbol{k})| = \frac{M!}{(M/r)!\, r!^{M/r} \prod_{i=1}^{n} k_i!} \exp\left(-\frac{(r-1)M_2}{2\,M} - \frac{(r-1)^2\, M_2^2}{4\,M^2} + O\left(\frac{r^4k_{\max}^4}{M}\right)\right).$$

Our formula is the same as that obtained by Blinovsky and Greenhill [13], apart from the error term which is shown in (2.1.4). The proof of this theorem follows mainly the strategies of [13]. However, our switching operation involves more edges than the switching used in [13] which leads to a looser constraint on r and $k_{\max}$, making our formula more widely applicable. Specifically, the switching in [13] aims to replace only two edges of the 4-cycle while we completely destroy the 4-cycle in our switching by replacing all its edges with other 4 edges. This allows us to ease the condition as stated in Theorem 5.1.1.

Here is the structure for the rest of this chapter. In the next section, we explain how to use bipartite graphs to represent hypergraphs in $\mathcal{H}_r(\boldsymbol{k})$. We define a set $\mathcal{B}_r^+(\boldsymbol{k})$ of bipartite graphs, corresponding to hypergraphs which avoid some bad substructures. Our set $\mathcal{B}_r^+(\boldsymbol{k})$ is slightly larger than the corresponding set in [13]. This will be followed by Lemma 5.2.2 which shows that $\mathcal{B}_r^+(\boldsymbol{k})$ contains almost all bipartite graphs which represent elements of $\mathcal{H}_r(\boldsymbol{k})$. Then Section 5.3 will describe a switching operation and analyse both directions of the switching. At the end of this chapter we apply a summation lemma from [37, Corollary 4.5] to find an asymptotic expression for the ratio of $|\mathcal{L}_r(\boldsymbol{k})|$ and $|\mathcal{B}_r^+(\boldsymbol{k})|$. Combining these expressions will complete the proof of our main result.

5.2 Representation of hypergraphs by bipartite graphs

This section explains how to transfer any uniform hypergraph into a bipartite graph and vice versa. We can do this correlation by interpreting the incidence matrix of a hypergraph as adjacency matrix of bipartite graph. This leads us to prove our main result with the help of some previous results for bipartite graphs. It will also visualise the intersections between the edges of a hypergraph in its corresponding bipartite graph, which then make it easier to apply a switching method on bipartite graphs with certain properties.

For a hypergraph $H = (V, E)$ and an edge e in E, we define a *link* in e as a 2-element multisubset of e. A *loop* is a link with repeated vertex. The *multiplicity* of a link is the number of edges in E which contains the link. For instance, if the link $\{x, y\}$ occurs in exactly 5 edges, then this link has multiplicity five. We will call any link with multiplicity two a *double* link.

The incidence matrix of any hypergraph H is given by a matrix $A_H = (a_{ij})$ of size $n \times (M/r)$, where a_{ij} take value 1 if the corresponding vertex of ith row belongs to the edge representing the jth column, otherwise $a_{ij} = 0$. Notice that the matrix depends on the labelling of the edges of H. In particular, if H is simple then the distinct edges result in distinct columns of the incidence matrix. In this case, there are exactly $(M/r)!$ distinct incidence matrices for the same simple hypergraph H. The incidence matrix of a hypergraph H (with respect to a given labelling of the edges of H) can be viewed as biadjacency matrix of a simple bipartite graph, say $B = B(H) = (V(B), E(B))$ with bipartition set $V(B) = V \cup E$ and edge set given by

$$E(B) = \{ve : \, v \in V \text{ and the edge } e \in H \text{ contains } v\}.$$

We refer to a vertex of V in B as a *left vertex* and a vertex of E in B as a *right vertex*. We will also write $H(B)$ for the hypergraph which corresponds to the bipartite graph B. For a complete bipartite graph $K_{a,b}$, we say that $K_{a,b}$ is a *subgraph of* B, or that B contains a *copy* of $K_{a,b}$, if there are a left vertices and b right vertices which induce a complete subgraph of B. Throughout this chapter, we will refer to a copy of $K_{2,2}$ in B as 4-*cycle*. In this way we can represent b edges of $H(B)$ which intersect in a vertices by a subgraph $K_{a,b}$ of B. Hence, double links and triple links in H correspond respectively to subgraphs $K_{2,2}$ and $K_{2,3}$ in B. An example of 5-uniform hypergraph and its representative of bipartite graph is demonstrated in Figure 5.1. In this example, the edges e_1 and e_4 have a double link $\{v_1, v_2\}$ which corresponds to a copy of $K_{2,2}$ on $\{v_1, v_2\} \cup \{e_1, e_4\}$, while the intersection of three vertices between the edges e_2 and e_3, is shown as a copy of $K_{3,2}$ in the corresponding bipartite graph. In particular, a hypergraph is linear if and only if its corresponding bipartite graph contains no copy of $K_{2,2}$.

Next, denote by $\mathcal{B}_r(\boldsymbol{k})$ the set of all (simple) bipartite graphs $B = B(H)$ for every $H \in \mathcal{H}_r(\boldsymbol{k})$. Hence, the degree sequence of the left vertices of B is given by $\boldsymbol{k}$ and each right vertex has degree r. Let $\mathcal{B}_r^{(0)}(\boldsymbol{k})$ contains the elements of $\mathcal{B}_r(\boldsymbol{k})$ such

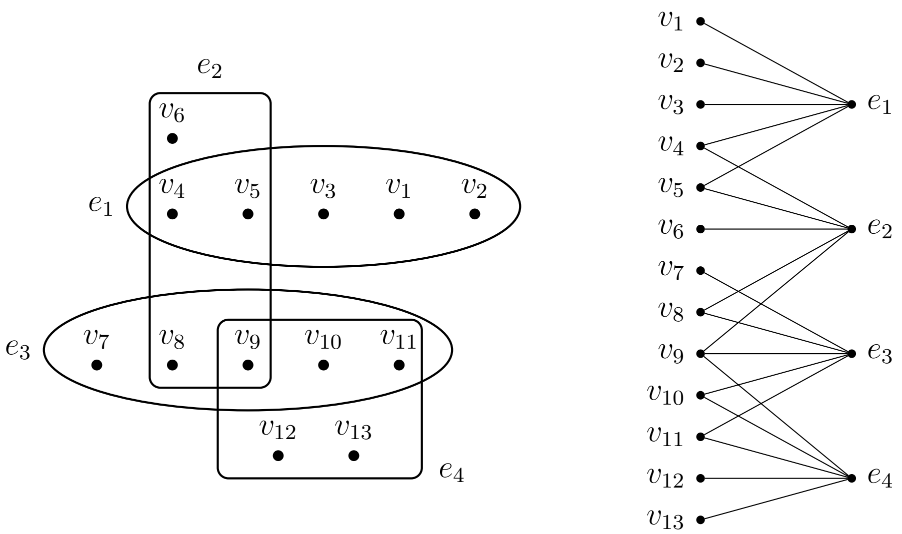

Figure 5.1: A hypergraph and its corresponding bipartite graph.

that no two right vertices are adjacent to the same set of left vertices. This gives us simple hypergraphs. Since there are $(M/r)!$ incidence matrices that represent a given simple hypergraph, then for each hypergraph H in $\mathcal{H}_r(\boldsymbol{k})$ we have $(M/r)!$ corresponding bipartite graphs in $\mathcal{B}_r^{(0)}(\boldsymbol{k})$. This implies that

$$(M/r)!\ |\mathcal{H}_r(\boldsymbol{k})| = |\mathcal{B}_r^{(0)}(\boldsymbol{k})|. \tag{5.2.1}$$

The number of simple bipartite graphs has been computed by Greenhill, McKay and Wang [37]. We conclude from [37, Theorem 1.3] that

$$|\mathcal{B}_r(\boldsymbol{k})| = \frac{M!}{(r!\,)^{M/r}\prod_{j=1}^{n} k_j!}\ \exp\left(-\frac{(r-1)\,M_2}{2M} + O\left(\frac{r^2 k_{\max}^2}{M}\right)\right). \tag{5.2.2}$$

In fact the formula in [37, Theorem 1.3], restated in (2.1.2), has more precise terms in the exponential function. We only keep the main term in the exponent, as all remaining terms are bounded above by $O\,(r^2\ k_{\max}^2/M)$.

We will define a subset of bipartite graphs in $\mathcal{B}_r(\boldsymbol{k})$ which correspond to hypergraphs in $\mathcal{H}_r(\boldsymbol{k})$ with a limited number of double links and other restrictions.

Define

$$N = 5 \max \left\{ \lceil \log M \rceil, \left\lceil 2(r-1)^2 \, M_2^2 / M^2 \right\rceil \right\}. \tag{5.2.3}$$

Let $\mathcal{B}_r^+(\boldsymbol{k})$ be a set of bipartite graphs $B \in \mathcal{B}_r(\boldsymbol{k})$ satisfying the following properties. We also translate the property into the hypergraph setting.

(i) B has no copy of $K_{3,2}$. This means that any pair of edges in $H(B)$ intersect in at most two vertices.

(ii) B contains no copy of $K_{2,3}$. Hence, there is no set of three edges in $H(B)$ which share two vertices.

(iii) All 4-cycles in B are edge-disjoint. (That is, two distinct 4-cycles in B can share a right vertex or a left vertex but not an edge). This implies that if $\{v, w\}$ is a double link in e then e contains no other double link of the form $\{v, u\}$ or $\{w, u\}$.

(iv) The number of 4-cycles in B is at most N. This will produce a hypergraph $H(B)$ with at most N double links.

Observe that when $r \geq 3$, the set $\mathcal{B}_r^+(\boldsymbol{k})$ is contained in $\mathcal{B}_r^{(0)}(\boldsymbol{k})$ since all bipartite graphs in $\mathcal{B}_r^+(\boldsymbol{k})$ satisfy property (i), and hence the corresponding hypergraphs have no repeated edges. For a given bipartite graph L, an upper bound on the probability that a random graph in $\mathcal{B}_r(\boldsymbol{k})$ contains L as subgraph has been provided by McKay [60]. The following lemma is a special case of [60, Theorem 3.5(a)] which is obtained by substituting J by L and setting $H = \emptyset$ in the notation of [60]. The following version of [60, Theorem 3.5(a)] has also been stated in [13, Lemma 2.1].

Lemma 5.2.1. [60, Theorem 3.5(a)] *Let $\mathcal{B}(\boldsymbol{g})$ be the set of simple bipartite graphs with vertex bipartition given by $\{a_1, \ldots, a_n\} \cup \{b_1, \ldots, b_m\}$ and degree sequence*

$$\boldsymbol{g} = (g_1, \ldots, g_n; g_1', \ldots, g_m'),$$

where $deg(a_i) = g_i$ *for* $i = 1, \ldots, n$ *and* $deg(b_j) = g'_j$ *for* $j = 1, \ldots, m$. *Let* L *be a subgraph of the complete bipartite graph on this vertex bipartition. Denote by* $\mathcal{B}(\boldsymbol{g}, L)$ *the set of bipartite graphs in* $\mathcal{B}(\boldsymbol{g})$ *which contain* L *as a subgraph where* $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_n; \ell'_1, \ldots, \ell'_m)$ *is the degree sequence of* L. *Write* $E_{\boldsymbol{g}} = \sum_{i=1}^{n} g_i$ *and* $E_{\boldsymbol{\ell}} = \sum_{i=1}^{n} \ell_i$ *and let* $g_{\max}$ *and* $\ell_{\max}$ *denote the maximum degree in* $\boldsymbol{g}$ *and* $\boldsymbol{\ell}$, *respectively. Define*

$$\Gamma = 2g_{\max}(g_{\max} + \ell_{\max} - 1) + 2.$$

If $E_{\boldsymbol{g}} - \Gamma \geq E_{\boldsymbol{\ell}}$ *then*

$$\frac{|\mathcal{B}(\boldsymbol{g}, L)|}{|\mathcal{B}(\boldsymbol{g})|} \leq \frac{\prod_{i=1}^{n}(g_i)_{\ell_i} \prod_{j=1}^{m}(g'_j)_{\ell_{j'}}}{(E_{\boldsymbol{g}} - \Gamma)_{E_{\boldsymbol{\ell}}}}.$$

This inequality will be applied frequently in the rest of this chapter. The following lemma confirms that the size of $\mathcal{B}_r^+(\boldsymbol{k})$ is asymptotically equal to the number of bipartite graphs in $\mathcal{B}_r(\boldsymbol{k})$.

Lemma 5.2.2. *Under the conditions of Theorem 5.1.1 we have*

$$|\mathcal{B}_r^+(\boldsymbol{k})| = \left(1 + O\Big(\frac{r^4\, k_{\max}^4}{M}\Big)\right) |\mathcal{B}_r(\boldsymbol{k})|.$$

Proof. Suppose that B is chosen uniformly at random from $\mathcal{B}_r(\boldsymbol{k})$. We will apply Lemma 5.2.1 to find the probability that each of the defining properties of $\mathcal{B}_r^+(\boldsymbol{k})$ holds in B. Hence from the defnition of $\mathcal{B}_r(\boldsymbol{k})$, B has degree sequence given by $\boldsymbol{g} = (k_1, \ldots, k_n; r, \ldots, r)$ and $E_{\boldsymbol{g}} = M$. We will define L as a subgraph of bipartite graph in $\mathcal{B}_r(\boldsymbol{k})$ which illustrates each property of $\mathcal{B}_r^+(\boldsymbol{k})$. In this case we have $g_{\max} = \max\{r, k_{\max}\}$ and $\ell_{\max} \leq g_{\max}$. Therefore from the definition of Γ, it can be bounded above by

$$\Gamma = O\left(g_{\max}^2\right) = O\left(r^2 + k_{\max}^2\right).$$

For property (i): we choose three distinct left vertices labelled by $v_{j_1}, v_{j_2}, v_{j_3}$ and two distinct right vertices e_{i_1} and e_{i_2}. Lemma 5.2.1 indicates that the probability that B contains $L = K_{3,2}$ on the vertices $\{v_{j_1}, v_{j_2}, v_{j_3}\} \cup \{e_{i_1}, e_{i_2}\}$ is bounded above by

$$\frac{r^2(r-1)^2(r-2)^2\,(k_{j_1})_2\,(k_{j_2})_2\,(k_{j_3})_2}{(M - O(r^2 + k_{\max}^2))_6}.$$

Since $r^4 k_{\max}^4 = o(M)$ holds, then

$$(M - O(r^2 + k_{\max}^2))_6 = M^6(1 + o(1)).$$

The number of choices for $\{e_{i_1}, e_{i_2}\}$ is $\binom{M/r}{2}$. Hence the expected number of $K_{3,2}$ in B is at most

$$\begin{aligned}(1+o(1))\binom{M/r}{2}(r^6/M^6)\sum_{\{j_1,j_2,j_3\}}(k_{j_1})_2\,(k_{j_2})_2\,(k_{j_3})_2 &= O\left(\frac{(M/r)^2\,M_2^3\,r^6}{M^6}\right)\\ &= O\left(\frac{r^4\,k_{\max}^3}{M}\right),\end{aligned}$$

where the sum is over all possible choices of the set of three left vertices $\{v_{j_1}, v_{j_2}, v_{j_3}\}$. Therefore, property (i) fails in B with probability $O\left(r^4\,k_{\max}^3/M\right)$. Now, for property (ii), we choose two distinct left vertices v_{j_1}, v_{j_2} and three right vertices with distinct labels $e_{i_1}, e_{i_2}, e_{i_3}$. Arguing as above, applying Lemma 5.2.1 shows the probability that B contains a copy of $L = K_{2,3}$ on $\{v_{j_1}, v_{j_2}\} \cup \{e_{i_1}, e_{i_2}, e_{i_3}\}$ is at most

$$\frac{r^3(r-1)^3\,(k_{j_1})_3\,(k_{j_2})_3}{(M - O(r^2 + k_{\max}^2))_6}.$$

Then, the expected number of $K_{2,3}$ in B is at most

$$\begin{aligned}(1+o(1))\binom{M/r}{3}(r^6/M^6)\sum_{\{j_1,j_2\}}(k_{j_1})_3\,(k_{j_2})_3 &= O\left(\frac{(M/r)^3\,M_3^2\,r^6}{M^6}\right)\\ &= O\left(\frac{r^3\,k_{\max}^4}{M}\right).\end{aligned}$$

This proves that property (ii) fails in B with probability $O\left(r^3\,k_{\max}^4/M\right)$.

Now consider L to be a bipartite graph which consists of two 4-cycles with one edge in common. Let $\{v_{j_1}, v_{j_2}, v_{j_3}\} \cup \{e_{i_1}, e_{i_2}, e_{i_3}\}$ be a fixed set of 3 distinct left vertices and 3 distinct right vertices, and assume without loss of generality that $v_{i_2}e_{i_2}$ is the common edge of the two 4-cycles. Applying Lemma 5.2.1 again shows the upper bound of the probability that B has a copy of L on the chosen vertex set is bounded above by

$$\frac{r^3(r-1)^3(r-2)\,(k_{j_1})_2\,(k_{j_2})_3(k_{j_3})_2}{(M-O(r^2+k_{\max}^2))_7}.$$

In this case, the expected number of copies of L in B is at most

$$\begin{aligned}(1+o(1))\binom{M/r}{3}(r^7/M^7)\sum_{\{j_1,j_2,j_3\}}(k_{j_1})_2\,(k_{j_2})_3\,(k_{j_3})_2 &= O\left(\frac{(M/r)^3\,M_2^2\,M_3\,r^7}{M^7}\right)\\ &= O\left(\frac{r^4\,k_{\max}^4}{M}\right),\end{aligned}$$

which indicates that (ii) fails in B with probability $O\,(r^4\,k_{\max}^4/M)$.

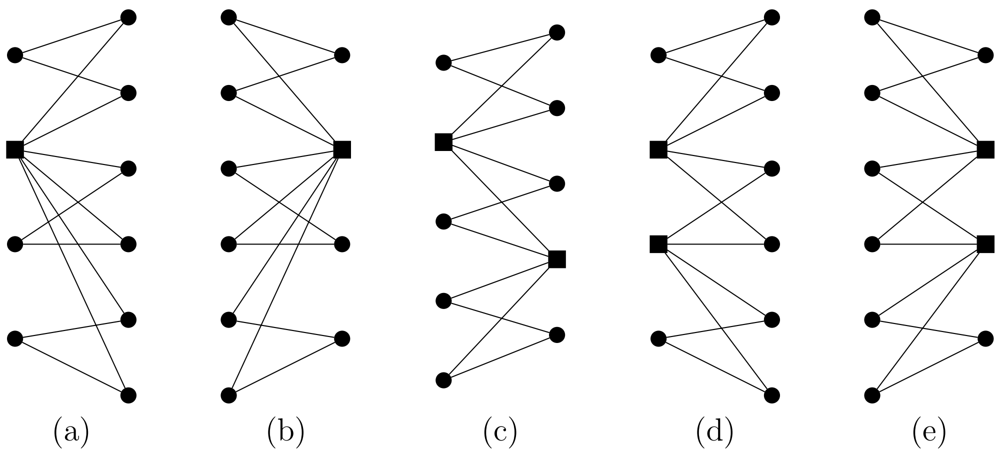

Figure 5.2: Some subgraphs which are rare in $\mathcal{B}_r(\boldsymbol{k})$.

To prove that property (iv) holds with high probability, first we show that some clusters of three 4-cycles, shown in Figure 5.2, are sufficiently rare. Repeating the above arguments, Lemma 5.2.1 can be applied to find the probability for L defined here by three distinct 4-cycles with precisely one left vertex in common, as shown in Figure 5.2 (a). In this case, we deduce the expected number of copies of L in B is at most

$$\begin{aligned}\frac{(M/r)^6\, M_2^3\, M_6\, r^{12}}{(M - O(r^2 + k_{\max}^2))^{12}} &= O\left(\frac{r^6\, k_{\max}^8}{M^2}\right)\\ &= O\left(\frac{r^2\, k_{\max}^4}{M}\right)\end{aligned}$$

since $r^4\, k_{\max}^4 = o(M)$. For L depicted in Figure 5.2 (b), the expected number of copies L where three 4-cycles overlap only at a right vertex is at most

$$\begin{aligned}\frac{(M/r)^4\, M_2^6\, r^{12}}{(M - O(r^2 + k_{\max}^2))^{12}} &= O\left(\frac{r^8\, k_{\max}^6}{M^2}\right)\\ &= O\left(\frac{r^4\, k_{\max}^2}{M}\right).\end{aligned}$$

Next, suppose that L consists of three 4-cycles with 5 left distinct vertices and 5 right distinct vertices, where one of the 4-cycles meets the second 4-cycle only at a left vertex and shares precisely one right vertex with the third 4-cycle, as shown in Figure 5.2 (c). Then the expected number of copies of L in B is bounded above by

$$\begin{aligned}\frac{(M/r)^5\, M_2^4\, M_4\, r^{12}}{(M - O(r^2 + k_{\max}^2))^{12}} &= O\left(\frac{r^7\, k_{\max}^7}{M^2}\right)\\ &= O\left(\frac{r^3\, k_{\max}^3}{M}\right).\end{aligned}$$

Finally, the expected number of copies of the graph depicted in Figure 5.2(d) in B is at most

$$(1 + o(1))\, \frac{(M/r)^6\, M_2^2\, M_4^2\, r^{12}}{M^{12}} = O\left(\frac{r^6\, k_{\max}^8}{M^2}\right) = O\left(\frac{r^2\, k_{\max}^4}{M}\right)$$

while the expected number of copies of the graph in Figure 5.2(e) is bounded above by

$$(1 + o(1))\, \frac{(M/r)^4\, M_2^6\, r^{12}}{M^{12}} = O\left(\frac{r^8\, k_{\max}^6}{M^2}\right) = O\left(\frac{r^4\, k_{\max}^2}{M}\right).$$

Overall, the above arguments show that the probability that any graph from Figure 5.2 occurs in B is $O\left(r^4 k_{\max}^4/M\right)$.

Now we can prove that with sufficiently high probability, the number of 4-cycles in B is at most N. Combining all above arguments, we know that with probability $1-O(r^4 k_{\max}^4/M)$, each 4-cycle in B is either vertex-disjoint from all other 4-cycles, or shares a vertex with at most one other 4-cycle. We say that a pair of 4-cycles is a *left-fused pair* if the two 4-cycles intersect in exactly one left vertex, and are otherwise vertex-disjoint from all other 4-cycles. Similarly, a pair of 4-cycles is a *right-fused pair* if the two 4-cycles intersect in exactly one right vertex, and are otherwise vertex-disjoint from all other 4-cycles. Observe that there are 4 automorphisms for each (labelled) disjoint 4-cycle and each of the fused pairs has 8 automorphisms. Again, by applying Lemma 5.2.1 we will show that B does not contain too many 4-cycles which are disjoint, or in left-fused or right-fused pairs.

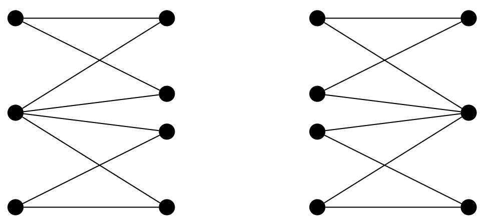

Figure 5.3: Left-fused pair (on the left) and right-fused pair of 4-cycles

Let $Q_1 = \max\{\lceil \log M \rceil, \lceil 2(r-1)^2\, M_2^2/M^2 \rceil\}$ and $d_1 = Q_1 + 1$. We will bound the expected number of sets of d_1 disjoint 4-cycles in B, using Lemma 5.2.1. Let $(j_1, j_2, \ldots, j_{2d_1}) \in [n]^{2d_1}$ and choose a $2d_1$-tuple of edge labels $(i_1, \ldots, i_{2d_1}) \in [M/r]^{2d_1}$. Then the probability that B has a 4-cycle on $\{v_{j_{2s-1}}, v_{j_{2s}}\} \cup \{e_{i_{2s-1}}, e_{i_{2s}}\}$ for $s = 1, 2, \ldots, d_1$, by Lemma 5.2.1, is at most

$$\prod_{s=1}^{2d_1} (k_{j_s})_2 \; O\left(\left(\frac{r(r-1)}{M^2}\right)^{2d_1}\right).$$

To find the expected number of sets of d_1 disjoint 4-cycles we must sum over all choices of $(j_1, \ldots, j_{2d_1})$ and $(i_1, \ldots, i_{2d_1})$, then divide by $4^{d_1}\, d_1!$ to account for symmetries, giving an upper bound of

$$\begin{aligned} O\left(\frac{1}{d_1!}\left(\frac{(r-1)^2\, M_2^2}{4\, M^2}\right)^{d_1}\right) &= O\left(\left(\frac{e\,(r-1)^2\, M_2^2}{4\, d_1\, M^2}\right)^{d_1}\right) \\ &= O\left((e/8)^{d_1}\right) \\ &= O(1/M). \end{aligned}$$

Next, define $Q_2 = \max\{\lceil \log M \rceil, \lceil (r-1)^4\, M_2^2/M^4 \rceil\}$ and $d_2 = Q_2 + 1$. We will bound the expected number of sets of d_2 vertex-disjoint left-fused pairs of 4-cycles. Using Lemma 5.2.1 and arguing as above, this expectation is bounded above by

$$\begin{aligned} O\left(\frac{1}{d_2!}\left(\frac{(r-1)^4\, M_2^2 M_4}{8\, M^4}\right)^{d_2}\right) &= O\left(\left(\frac{e\,(r-1)^4\, M_2^2\, M_4}{8\, d_2\, M^4}\right)^{d_2}\right) \\ &= O\left((e/8)^{d_2}\right) \\ &= O(1/M). \end{aligned}$$

Similarly, define $d_3 = Q_3 + 1$ where $Q_3 = \max\{\lceil \log M \rceil, \lceil (r-1)^3\, (r-2)_2\, M_2^4/M^5 \rceil\}$ and let $d_3 = Q_3 + 1$. Arguing as above, the expected number of sets of d_3 vertex-disjoint right-fused pairs of 4-cycles is at most

$$\begin{aligned} O\left(\frac{1}{d_3!}\left(\frac{(r-1)^3\,(r-1)^2\, M_2^4}{8\, M^5}\right)^{d_3}\right) &= O\left(\left(\frac{e\,(r-1)^3\,(r-2)^2\, M_2^4\, M_4}{8\, d_3\, M^5}\right)^{d_3}\right) \\ &= O\left(\left(\frac{e}{8}\right)^{d_3}\right) \\ &= O(1/M). \end{aligned}$$

Therefore, with probability $1 - O(1/M)$, the number of 4-cycles in B which overlap at most one other 4-cycle is at most $Q_1 + 2Q_2 + 2Q_3 \leq 5Q_1 = N$. This completes the proof since we have shown that with probability $1 - O(r^4 k_{\max}^4/M)$, no other 4-cycles are present in B. □

5.3 Switching for 4-cycles

From the definition of $\mathcal{B}_r^+(\boldsymbol{k})$, the linear hypergraphs in $\mathcal{L}_r(\boldsymbol{k})$ can be represented by bipartite graphs in $\mathcal{B}_r^+(\boldsymbol{k})$ which have no 4-cycles For this reason, we define $\mathcal{C}_d = \mathcal{C}_d(r, \boldsymbol{k})$ to be the set of bipartite graphs in $\mathcal{B}_r^+(\boldsymbol{k})$ with precisely d distinct 4-cycles, where $d \in \{0, 1, \ldots N\}$. This enables us to express $|\mathcal{B}_r^+(\boldsymbol{k})|$ as

$$|\mathcal{B}_r^+(\boldsymbol{k})| = \sum_{d=0}^{N} |\mathcal{C}_d|. \tag{5.3.1}$$

Observe that $\mathcal{C}_0$ is the set of bipartite graphs which represent linear hypergraphs in $\mathcal{L}_r(\boldsymbol{k})$. Then (5.2.1) implies that

$$|\mathcal{L}_r(\boldsymbol{k})| = \frac{|\mathcal{C}_0|}{(M/r)!}. \tag{5.3.2}$$

In this section we will use a switching operation to estimate the ratio $|\mathcal{C}_d|/|\mathcal{C}_{d-1}|$. This ratio, with the summation lemma, will be used to estimate $|\mathcal{C}_0|$ as shown at the end of this chapter.

Now we describe a switching approach as follows. First, we define some notation that we need in both directions of the switching. Let T be a 12-tuple of distinct vertices,

$$T = (u_1, u_2, w_1, w_2, w_3, w_4, f_1, f_2, g_1, g_2, g_3, g_4)$$

such that $u_1, u_2, w_1, w_2, w_3, w_4$ are left vertices and $f_1, f_2, g_1, g_2, g_3, g_4$ are right vertices of the vertex set V in the target bipartite graph. Then two sets F, F' of 8 distinct edges whose vertices are determined by the vertices of T, are defined as

$$F = \{u_1f_1,\ u_1f_2,\ u_2f_1,\ u_2f_2,\ w_1g_1,\ w_2g_2,\ w_3g_3,\ w_4g_4\}, \tag{5.3.3}$$

$$F' = \{w_1 f_1,\ w_3 f_1,\ w_2 f_2,\ w_4 f_2,\ u_1 g_1,\ u_1 g_3,\ u_2 g_2,\ u_2 g_4\}. \tag{5.3.4}$$

Define a *forward d-switching* from $G \in \mathcal{C}_d$ determined by T such that

- $u_1 f_1 u_2 f_2 u_1$ is a 4-cycle in G,
- G contains the edges $w_1 g_1$, $w_2 g_2, w_3 g_3$ and $w_4 g_4$,
- the edges in F' are all absent in G.

This switching operation applied on G will create a new bipartite graph denoted by G', which has same vertex set as G and edge set given by

$$E(G') = \Big(E(G) \cup F'\Big) - F.$$

The forward d-switching from G to G' is depicted in Figure 5.4, following the arrow from left to right.

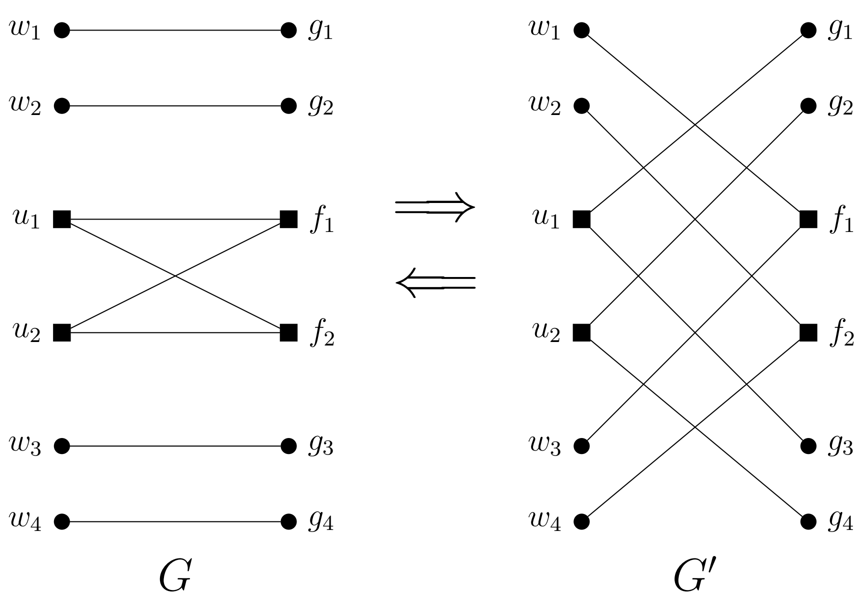

Figure 5.4: A d-switching which is designed to remove a 4-cycle

Note that in the corresponding hypergraphs of G and G', the forward switching replaces the edges $f_1, f_2, g_1, g_2, g_3, g_4$ with six new edges $f_1', f_2', g_1', g_2', g_3', g_4'$ defined by

$$f_1' = (f_1 \setminus \{u_1, u_2\}) \cup \{w_1, w_3\},\ f_2' = (f_2 \setminus \{u_1, u_2\}) \cup \{w_2, w_4\}$$

and

$$g_i' = (g_i \setminus \{w_i\}) \cup \{u_1\},\ g_j' = (g_j \setminus \{w_j\}) \cup \{u_2\},$$

for $i = 1, 3$ and $j = 2, 4$.

The forward d-switching from G determined by T is *legal* if the corresponding bipartite graph G' belongs to $\mathcal{C}_{d-1}$, otherwise we say it is an *illegal* forward d-switching.

The following lemma provides the cases which result in illegal forward switching.

Lemma 5.3.1. *Let $d \leq N$ be a positive integer and $G \in \mathcal{C}_d$. Suppose that the 12-tuple $T = (u_1, u_2, w_1, w_2, w_3, w_4, f_1, f_2, g_1, g_2, g_3, g_4)$ leads to an illegal forward d-switching from G. Then at least one of the following properties holds:*

(I) *The edge $w_i g_i$ belongs to a 4-cycle in G for some $i \in \{1, 2, 3, 4\}$.*

(II) *For some $i \in \{1, 2\}$, G satisfies at least one of the following:*

$$\mathrm{dist}_G(u_1, g_{2i-1}) = 3,\ \mathrm{dist}_G(u_2, g_{2i}) = 3,\ \mathrm{dist}_G(w_{2i-1}, f_1) = 3 \ \textit{or}\ \mathrm{dist}_G(w_{2i}, f_2) = 3.$$

(III) $\mathrm{dist}_G(w_1, w_3) = 2$ *or* $\mathrm{dist}_G(w_2, w_4) = 2$.

(IV) $\mathrm{dist}_G(g_1, g_3) = 2$ *or* $\mathrm{dist}_G(g_2, g_4) = 2$.

(V) *For some $i \in \{1, 2\}$, G contains at least one of the following 2-sets of edges:*

$$\{w_{2i}f_1,\ w_1 f_2\}, \quad \{w_{2i}f_1,\ w_3 f_2\}, \quad \{u_1 g_{2i},\ u_2 g_1\} \quad \textit{or} \quad \{u_1 g_{2i},\ u_2 g_3\}.$$

Proof. Fix $G \in \mathcal{C}_d$ and suppose that the 12-tuple T determines an illegal forward switching from G into a new graph denoted by G' as described above. Then either G' belongs to $\mathcal{B}_r^+(\boldsymbol{k})$ but is not in $\mathcal{C}_{d-1}$, or G' is not an element of $\mathcal{B}_r^+(\boldsymbol{k})$. The forward switching is designed to remove the 4-cycle $u_1 f_1 u_2 f_2 u_1$, and any other 4-cycle which is removed by the switching will be referred to as an *additional* or *undesired* 4-cycle. If the number of 4-cycles is reduced by more than one, under this forward switching, then at least one edge of an undesired 4-cycle (or set of undesired 4-cycles) has also been deleted by the switching. In this case we say that the undesired 4-cycle has been *destroyed*. (Up to 3 edges of this 4-cycle may remain in G', so it does not mean that the entire undesired 4-cycle has been removed.)

First suppose that $G' \in \mathcal{B}_r^+(\boldsymbol{k})$ but not in $\mathcal{C}_{d-1}$. One of the possible reasons that G' does not belong to $\mathcal{C}_{d-1}$ is that G' has less than $d-1$ distinct 4-cycles. This happens when the switching has destroyed at least one additional 4-cycle. Since $G' \in \mathcal{B}_r^+(\boldsymbol{k})$, then the additional 4-cycle that was destroyed must contain $w_i g_i$ for some $i \in \{1,2,3,4\}$. Hence (I) holds.

Another possible reason for G' not being in $\mathcal{C}_{d-1}$ is that the number of 4-cycles after the switching is at least d. This indicates that at least one new 4-cycle has been created in G', which must contain at least one of the new edges of F' defined in (5.3.4). Consider the case where the new 4-cycle contains precisely one edge of F' and let $i \in \{1,2\}$. If a new 4-cycle contains $u_1 g_{2i-1}$ then $\text{dist}_G(u_1, g_{2i-1}) = 3$, so (II) holds. Also if $u_2 g_{2i}$ occurs in the new 4-cycle in G', then $\text{dist}_G(u_2, g_{2i}) = 3$ so (II) holds. Similarly, if precisely one edge of $w_{2i-1} f_1$ or $w_{2i} f_2$ is the only edge of F' which belongs to the new 4-cycle then respectively $\text{dist}_G(w_{2i-1}, f_1) = 3$ or $\text{dist}_G(w_{2i}, f_2) = 3$, which satisfies (II).

Now we will demonstrate the case when the new 4-cycle in G' contains at least two edges of F'. If both $w_1 f_1$ and $w_3 f_1$ are edges of a 4-cycle in G', then $\text{dist}_G(w_1, w_3) = 2$ which leads to property (III). We also obtain $\text{dist}_G(w_2, w_4) = 2$ when the new 4-cycle contains $w_2 f_2$ and $w_4 f_2$, so (III) holds in this case. Similarly if either $\{u_1 g_1, u_1 g_3\}$ or $\{u_2 g_2, u_2 g_4\}$ is a set of two edges of F' appearing in the new 4-cycle then we have respectively $\text{dist}_G(g_1, g_3) = 2$ or $\text{dist}_G(g_2, g_4) = 2$, hence (IV) holds.

If both edges $w_1 f_1$ and $w_2 f_2$ occur in a 4-cycle in G' then $w_1 f_2$ and $w_2 f_1$ are already edges of G, so (V) holds . Also if $\{w_1 f_1,\, w_4 f_2\}$, $\{w_2 f_2,\, w_3 f_1\}$ or $\{w_3 f_1,\, w_4 f_2\}$ is contained in a 4-cycle in G', then respectively G already contains $\{w_1 f_2,\, w_4 f_1\}$, $\{w_2 f_1,\, w_3 f_2\}$ or $\{w_3 f_2,\, w_4 f_1\}$. Therefore, (V) holds. Repeating similar arguments, if $\{u_1 g_1,\, u_2 g_2\}$, $\{u_1 g_1,\, u_2 g_4\}$, $\{u_2 g_2,\, u_1 g_3\}$ or $\{u_1 g_3,\, u_2 g_4\}$ belongs to the new 4-cycle created by the switching then, respectively, $\{u_1 g_2,\, u_2 g_1\}$, $\{u_1 g_4,\, u_2 g_1\}$, $\{u_2 g_3,\, u_1 g_2\}$ or $\{u_1 g_4,\, u_2 g_3\}$ is contained already in G which shows that property (V) holds.

Next, suppose that G' does not satisfy at least one of the properties of $\mathcal{B}_r^+(\boldsymbol{k})$. Then at least one 4-cycle has been created by the switching. Arguing as above, at least one of the properties (II)–(IV) must hold. This completes the proof. $\square$

In the next lemma we will analyse the number of legal forward switchings.

Lemma 5.3.2. *Assume that the conditions of Theorem 5.1.1 hold and let $d \in \{1, 2, \ldots, N\}$. For each $G \in \mathcal{C}_d$, the number of 12-tuples T which result in a legal forward switching from G is*

$$4dM^4 \left(1 + O\Big(\frac{d + r^2 k_{\max}^2}{M}\Big)\right).$$

Proof. Fix $d \in \{1, 2, \ldots, N\}$ and let G be a arbitrary bipartite graph from $\mathcal{C}_d$. Define S to be the set of all 12-tuples $T = (u_1, u_2, f_1, f_2, w_1, w_2, w_3, w_4, g_1, g_2, g_3, g_4)$ such that

- $u_1, u_2, w_1, w_2, w_3, w_4$ are distinct vertices of $\{v_1, \ldots, v_n\}$ and $f_1, f_2, g_1, g_2, g_3, g_4$ are distinct vertices of $\{e_1, \ldots, e_{M/r}\}$,
- G has a 4-cycle on $u_1 f_1 u_2 f_2 u_1$,
- $w_i g_i$ is an edge of G which does not belong to any 4-cycle, for $i = 1, 2, 3, 4$,
- G does not contain any edge of F'.

So, the number of 12-tuples giving a legal forward d-switching from G is bounded above by $|S|$. To estimate this number, we first compute the cardinality of S. There are $4d$ choices for a 4-tuple (u_1, u_2, f_1, f_2) such that there is a 4-cycle on $\{u_1, u_2, f_1, f_2\}$, and at most M ways to choose each edge $w_i g_i$, for all $i = 1, 2, 3, 4$. Then,

$$|S| \leq 4dM^4.$$

For the lower bound of $|S|$, after choosing the 4-tuple (u_1, u_2, f_1, f_2) representing the 4-cycle in $4d$ ways, we have at least

$$M - 2k_{\max} - 2r - 4d - 2rk_{\max} = M\left(1 - O\Big(\frac{d + rk_{\max}}{M}\Big)\right)$$

ways to choose w_1g_1 such that $w_1 \notin \{u_1, u_2\}$, $g_1 \notin \{f_1, f_2\}$, the edge w_1g_1 is not contained in a 4-cycle in G, and the vertices w_1 and g_1 are respectively not neighbours of f_1 and u_1. Next, we need to choose w_2g_2 provided that w_2 is distinct from u_1, u_2, w_1, g_2 is distinct from f_1, f_2, g_1, the edge w_2g_2 is not contained in a 4-cycle in G, and the edges w_2f_2, u_2g_2 are not in G. In this way, the choices of w_2g_2 is at most

$$M - 3k_{\max} - 3r - 4d - 2rk_{\max} = M\left(1 - O\Big(\frac{d + rk_{\max}}{M}\Big)\right).$$

Similarly, the number of choices of w_3g_3 such that $w_3 \notin \{u_1, u_2, w_1, w_2\}$, $g_3 \notin \{f_1, f_2, g_1, g_2\}$, w_3g_3 is not part of any 4-cycle in G, and the edges w_3f_1, u_1g_3 are not in G, is at least

$$M - 4k_{\max} - 4r - 4d - 2rk_{\max} = M\left(1 - O\Big(\frac{d + rk_{\max}}{M}\Big)\right).$$

The last edge we need to choose is w_4g_4 such that $w_4 \notin \{u_1, u_2, w_1, w_2, w_3\}$, $g_4 \notin \{f_1, f_2, g_1, g_2, g_3\}$, w_4g_4 does not belong to a 4-cycle in G and the edges w_4f_2, u_2g_4 are not in G. Therefore, the number of choosing w_4g_4 is bounded above by

$$M - 5k_{\max} - 5r - 4d - 2rk_{\max} = M\left(1 - O\Big(\frac{d + rk_{\max}}{M}\Big)\right).$$

Combining the upper and lower bounds of $|S|$ shows that

$$|S| = 4\,d\,M^4\left(1 + O\Big(\frac{d + rk_{\max}}{M}\Big)\right). \tag{5.3.5}$$

This determines the upper bound of the number of tuples in legal forward switching. Subtracting from $|S|$ the number of T in the illegal case will give the lower bound of the 12-tuples T in the legal forward d-switchings. The upper bound in illegal case can be obtained by computing the upper bound of the number of T when G satisfies each properties (II)-(V) of Lemma 5.3.1. Observe that property (I) has been ignored here because no element of S satisfies this property. For condition (II), the number of 12-tuples in S such that $\mathrm{dist}_G(u_1, g_{2i-1}) = 3$ for some $i \in \{1, 2\}$ is at most

$$2 \cdot 4d\, M^3\, r^2\, k_{\max}^2 = 4dM^4\, O\left(\frac{r^2 k_{\max}^2}{M}\right),$$

where the factor 2 in this equation is to cover all values of i. We also have the same upper bound for the number of 12-tuples in S with $\mathrm{dist}_G(u_2, g_{2i}) = 3$ for some $i \in \{1, 2\}$. Similarly, the number of 12-tuples in S which satisfy $\mathrm{dist}_G(w_{2i-1}, f_1) = 3$ or $\mathrm{dist}_G(w_{2i}, f_2) = 3$, for some $i \in \{1, 2\}$, is bounded above by

$$4 \cdot 4d\, M^3\, r^2\, k_{\max}^2 = 4dM^4\, O\left(\frac{r^2 k_{\max}^2}{M}\right).$$

Now suppose G satisfies (III). Then $\mathrm{dist}_G(w_1, w_3) = 2$ or $\mathrm{dist}_G(w_2, w_4) = 2$. In this case, we can choose T in at most

$$2 \cdot 4dM^2\, M_2\, r\, k_{\max} = 4dM^4\, O\left(\frac{r k_{\max}^2}{M}\right)$$

ways. Similarly, when property (IV) holds in G, then either $\mathrm{dist}_G(g_1, g_3) = 2$ or $\mathrm{dist}_G(g_2, g_4) = 2$. Therefore, the number of 12-tuples in S which satisfy (IV) is at most

$$2 \cdot 4dM^2 M_2 r^2 = 4dM^4\, O\left(\frac{r^2\, k_{\max}}{M}\right).$$

For condition (V), the number of 12-tuples in S such that both $w_{2i}f_1$ and $w_1 f_2$ belongs to G for some $i \in \{1, 2\}$ is at most

$$2 \cdot 4dM^2\, r^2 k_{\max}^2 = 4dM^4\, O(r^2 k_{\max}^2 / M^2).$$

The same bound holds when replacing $\{w_{2i}, f_1,\ w_1 f_2\}$ by one of the pairs $\{w_3 f_2, w_{2i} f_1\}$, $\{u_2 g_1, u_1 g_{2i}\}$ or $\{u_2 g_3, u_1 g_{2i}\}$. This proves that, using Lemma 5.3.1, the number of

12-tuples T in S which gives an illegal forward switching in G is bounded above by

$$4dM^4\, O\left(\frac{r^2\, k_{\max}^2}{M}\right).$$

We complete the proof by subtracting this formula from (5.3.5). $\square$

Recall the edge sets F, F' defined in (5.3.3) and (5.3.4). Now we define a *reverse d-switching* from $G' \in \mathcal{C}_{d-1}$ which is designed to increase the number of 4-cycles by one. It is determined by a 12-tuple

$$T = (u_1, u_2, w_1, w_2, w_3, w_4, f_1, f_2, g_1, g_2, g_3, g_4)$$

of distinct vertices such that $u_1, u_2, w_1, w_2, w_3, w_4$ are left vertices and f_1, f_2, g_1, g_2, g_3, g_4 are right vertices with $F' \subseteq G'$ and $F \cap E(G') = \emptyset$. The reverse switching on G' determined by T will create a new bipartite graph G which has the same vertex set as G', and edge set defined by

$$E(G) = (E(G') \cup F) - F'.$$

The reverse switching from G' determined by T is called *legal* if the corresponding bipartite graph G belongs to $\mathcal{C}_d$, otherwise we say it is *illegal*.

The next lemma provides the conditions that lead to illegal reverse d-switching.

Lemma 5.3.3. *Let $d \leq N$ be a positive integer and let $G' \in \mathcal{C}_{d-1}$. If the reverse d-switching from G' defined by the 12-tuple T is illegal then at least one of the following properties holds:*

(I′) *At least one edge of F' belongs to a 4-cycle in G'.*

(II′) *For some $i \in [4]$ and $j, \ell \in \{1,2\}$, we have $\mathrm{dist}_{G'}(u_j, f_\ell) = 3$ or $\mathrm{dist}_{G'}(w_i, g_i) = 3$.*

(III′) *For some $i, j \in [4]$, $i \neq j$, the edges $w_i g_j$ and $w_j g_i$ are present in G'.*

(IV$'$) *At least one of the following pairs of edges are contained in G', for some $i \in \{1,2\}$: $\{u_1g_{2i}, w_{2i}f_1\}$ or $\{u_2g_{2i-1}, w_{2i-1}f_2\}$.*

(V$'$) $\mathrm{dist}_{G'}(u_1,u_2) = 2$ *or* $\mathrm{dist}_{G'}(f_1,f_2) = 2$.

Proof. Fix $G' \in \mathcal{C}_{d-1}$ and suppose that the 12-tuple T determines an illegal reverse switching which maps G' into a new graph G. Then either G belongs to $\mathcal{B}_r^+(\boldsymbol{k})$ but $G \notin \mathcal{C}_d$, or G is not an element of $\mathcal{B}_r^+(\boldsymbol{k})$.

First suppose that $G \in \mathcal{B}_r^+(\boldsymbol{k})$ and $G \notin \mathcal{C}_d$ which involves two cases. We begin with the case when G contains at most $d-1$ of 4-cycles, then the reverse d-switching has destroyed at least one 4-cycle. This 4-cycle must contain at least one edge of F', hence (I$'$) holds. Next, suppose that $G \notin \mathcal{C}_d$ because G has more than d of 4-cycles. This implies that the reverse switching creates at least one 4-cycle which is not equal to the desired 4-cycle on $\{u_1,u_2\} \cup \{f_1,f_2\}$. Then the unwanted 4-cycle must contain at least one edge of F, otherwise leads to contradiction with G being in $\mathcal{C}_d$. If, for $i \in \{1,2,3,4\}$, w_ig_i is the only edge of F' that occurs in a 4-cycle, then $\mathrm{dist}_{G'}(w_i,g_i) = 3$, hence (II$'$) holds. If a 4-cycle in G contains both edges w_ig_i and w_jg_j for some $i,j \in [4]$ with $i \neq j$, then w_ig_j and w_jg_i are already edges of G', so (III$'$) holds. We have assumed that $G \in \mathcal{B}_r^+(\boldsymbol{k})$, and so the edges u_1f_1, u_1f_2, u_2f_1, u_2f_2 are not involved in the additional 4-cycle. Hence we have covered all possibilities in this case.

Next, suppose that G is not contained in $\mathcal{B}_r^+(\boldsymbol{k})$. Then at least one of the properties of $\mathcal{B}_r^+(\boldsymbol{k})$ fails to hold in G. If the reverse switching has created a copy of $K_{3,2}$ or $K_{2,3}$ in G then this subgraph must contain at least one edge of F, otherwise this contradicts the fact that G' belongs to $\mathcal{C}_{d-1}$. Similar argument as above proves that $\mathrm{dist}_{G'}(w_i,g_i) = 3$ when w_ig_i is the only edge of F that occurs in $K_{3,2}$ (or $K_{2,3}$) in G, for some $i \in \{1,2,3,4,\}$. Then (II$'$) holds. If the copy of $K_{3,2}$ (or $K_{2,3}$) contains the 4-cycle on $\{u_1,u_2,f_1,f_2\}$ then $\mathrm{dist}_{G'}(f_1,f_2) = 2$ (or $\mathrm{dist}_{G'}(u_1,u_2) = 2$), so (V$'$) holds. If the copy of $K_{3,2}$ (or $K_{2,3}$) contains both edges w_ig_i and w_jg_j,for some $i,j \in \{1,2,3,4\}$, $i \neq j$, then (III$'$) holds.

If G contains two distinct 4-cycles which are not edge-disjoint, then the common edge must belong to F. We have $\mathrm{dist}_{G'}(w_i, g_i) = 3$ if $w_i g_i$ is in the intersection of two 4-cycles in G, which implies (III′). Similarly, if $u_i f_j$ belongs to two distinct 4-cycles for some $i, j \in \{1, 2\}$, then $\mathrm{dist}_{G'}(u_i, f_j) = 3$, hence (II′) holds. If G has a new 4-cycle which contains both edges $w_{2i}g_{2i}$, $u_1 f_1$, then $u_1 g_{2i}$ and $w_{2i} f_1$ are already edges of G' for some $i \in \{1, 2\}$. Therefore, (IV′) holds. Similarly, if there is a 4-cycle in G which contains both edges $w_{2i-1}g_{2i-1}$, $u_2 f_2$, for some $i \in \{1, 2\}$, then $u_2 g_{2i-1}$ and $w_{2i-1} f_2$ are already edges of G', hence (IV′) holds.

Finally, suppose that G satisfies properties (i)-(iii) of $\mathcal{B}_r^+(\boldsymbol{k})$ but contains more than N distinct 4-cycles. Then the reverse switching has introduced more than one new 4-cycles. We can treat these cases using the same arguments as above (when G contained more than d distinct 4-cycles). This completes the proof. □

The next lemma provides the number of 12-tuples T which gives legal reverse d-switchings.

Lemma 5.3.4. *Assume that the conditions of Theorem 5.1.1 hold and let $d \in \{1, 2, \ldots, N\}$. For each $G' \in \mathcal{C}_{d-1}$, the number of 12-tuples T which result in a legal reverse switching from G' is*

$$(r-1)^2 \, M^2 \, M_2^2 \, \left(1 + O\Big(\frac{dk_{\max} + r^2 k_{\max}^3}{M_2}\Big)\right).$$

Proof. We begin by finding the size of a certain set of 12-tuples which we use as an upper bound of the number of 12-tuples describing legal reverse switchings. Denote by S' the set of 12-tuples $T = (u_1, u_2, w_1, w_2, w_3, w_4, f_1, f_2, g_1, g_2, g_3, g_4)$ which determine a reverse switching from G' such that $F' \subseteq G'$, no edge of F' belongs to a 4-cycle and no edge of F is present in G'. Hence, the number of 12-tuples which describe legal reverse switchings is bounded above by $|S'|$. To estimate $|S'|$, we have at most M_2 choices for (u_1, g_1, g_3), at most M_2 choices for (u_2, g_2, g_4), at most $(r-1)M$ choices for (w_1, f_1, w_3) and at most $(r-1)M$ choices for (w_2, f_2, w_4). Therefore

$$|S'| \leq (r-1)^2 \, M^2 \, M_2^2.$$

To determine the lower bound of $|S'|$, we need a careful calculation for choosing its elements as follows. There are at least $(r-1)M - 2(r-1)(d-1)$ choices for (w_1, f_1, w_3), such that $w_1 f_1$ and $w_3 f_1$ are edges which does not belong to any 4-cycle in G'. Next, we have at least

$$\begin{aligned}
&(r-1)M - 2(r-1)(d-1) - 4(r-1)k_{\max} - (r-2)_2 \\
&= (r-1)M\left(1 + O\left(\frac{d + r + k_{\max}}{M}\right)\right)
\end{aligned}$$

choices for (w_2, f_2, w_4) such that $w_2 f_2$ and $w_4 f_2$ are two edges not in a 4-cycle, $w_2,\, w_4 \notin \{w_1, w_3\}$ and $f_1,\, f_2$ are distinct . The 3-tuple (u_1, g_1, g_3) where $u_1 g_1$ and $u_1 g_3$ are not in a 4-cycle in G', $u_1 \notin \{w_1, w_2, w_3, w_4\}$, $g_1, g_3 \notin \{f_1, f_2\}$ and the edges $w_1 g_1$, $w_3 g_3$, $u_1 f_1$, $u_1 f_2$ are all absent, can be chosen in at most

$$\begin{aligned}
&M_2 \;-\; 2(d-1)k_{\max} \;-\; 4\,k_{\max}^2 \;-\; 4(r-2)k_{\max} - 4(r-1)k_{\max}^2 \\
&= M_2\left(1 + O\left(\frac{dk_{\max} + rk_{\max}^2}{M_2}\right)\right).
\end{aligned}$$

Similarly, we have at least

$$M_2\left(1 + O\left(\frac{dk_{\max} + rk_{\max}^2}{M_2}\right)\right)$$

choices for 3-tuple (u_2, g_2, g_4) satisfying $u_2 g_2$ and $u_2 g_4$ are not in a 4-cycle in G', $u_2 \notin \{u_1, w_1, w_2, w_3, w_4\}$, $g_2, g_4 \notin \{g_1, g_3, f_1, f_2\}$ and the edges $w_2 g_2$, $w_4 g_4$, $u_2 f_1$, $u_2 f_2$ are all absent. Hence, combining with the upper bound on $|S'|$ proved earlier,

$$|S'| = (r-1)^2 \, M^2 \, M_2^2 \, \left(1 + O\Big(\frac{dk_{\max} + rk_{\max}^2}{M_2}\Big)\right). \tag{5.3.6}$$

This is also an upper bound for the the number of 12-tuples T which give a legal reverse switching. The lower bound can be estimated by subtracting the upper bound of the number of tuples in S' which gives illegal cases. We deduce this upper bound by estimating the number of T satisfying the properties of Lemma 5.3.3. Notice that there is no element in S' satisfies property (I′).

First suppose that G' satisfies (II′). If $\mathrm{dist}_{G'}(w_i, g_i) = 3$ for some $i \in \{1, 2, 3, 4\}$ then the 12-tuple T can be chosen in at most

$$4 \cdot (r-1)MM_2^2 r^3 k_{\max}^2 = (r-1)^2 M^2 M_2^2 \; O\left(\frac{r^2 k_{\max}^2}{M}\right).$$

The multiplication by 4 in this expression is to consider all possible options for i. We also achieve the same upper bound when $\mathrm{dist}_{G'}(u_j, f_\ell) = 3$ for $j, \ell \in \{1, 2\}$. Next, for condition (III′), if $w_i g_j$ and $w_j g_i$ belong to G', for some $i, j \in [4], i \neq j$, then we have at most

$$(r-1)^2\, M^2 M_2^2 \; O\left(\frac{r k_{\max}^2}{M}\right)$$

choices for T. When property (IV′) holds, there are at most

$$M_3 M_2 (r-1)^4 k_{\max} = (r-1)^2 M^2 M_2^2 \; O\left(\frac{r^2 k_{\max}^2}{M^2}\right)$$

choices of T such that $u_2 g_{2i-1}$ and $w_{2i-1} f_2$ are edges in G' for some $i \in \{1, 2\}$. Similarly we have same upper bound for T when G' contains $u_1 g_{2i}$ and $w_{2i} f_1$ for some $i \in \{1, 2\}$. Finally, for condition (V′), the upper bound for the number of T when $\mathrm{dist}_{G'}(u_1, u_2) = 2$ holds is

$$(r-1)^2\, M^2 M_2^2 \; O\left(\frac{r k_{\max}^3}{M_2}\right).$$

Also when $\mathrm{dist}_{G'}(f_1, f_2) = 2$ we have at most

$$(r-1)^2\, M^2 M_2^2 \; O\left(\frac{r^2 k_{\max}}{M}\right)$$

ways to choose the 12-tuple T. Now we combine these terms and subtract them from $|S'|$ to obtain the lower bound on the number of legal 12-tuples in S'. Therefore, the number of 12-tuples $T \in S'$ which describe a legal reverse d-switching from G' is

$$\begin{aligned}
&(r-1)^2\, M^2\, M_2^2\, \left(1 + O\Big(\frac{dk_{\max} + rk_{\max}^2}{M_2} + \frac{r^2\, k_{\max}^2}{M} + \frac{r\, k_{\max}^3}{M_2}\Big)\right) \\
&= (r-1)^2\, M^2\, M_2^2\, \left(1 + O\Big(\frac{dk_{\max} + r^2 k_{\max}^3}{M_2}\Big)\right),
\end{aligned}$$

as claimed. □

5.4 Completing our argument

In this section we demonstrate the proof of Theorem 5.1.1. We first need to provide an expression that relates $\sum_{d=0}^{N} |\mathcal{C}_d|$ and $|\mathcal{C}_0|$. This can be found by applying a summation lemma stated below. The proof of Lemma 5.4.2 will depend on the ratio $|\mathcal{C}_d|/|\mathcal{C}_{d-1}|$, for $d \in [N]$, given by

$$\frac{|\mathcal{C}_d|}{|\mathcal{C}_{d-1}|} = \frac{(r-1)^2 M_2^2}{4d\, M^2}\, \left(1 + O\Big(\frac{d\, k_{\max} + r^2\, k_{\max}^3}{M_2}\Big)\right), \tag{5.4.1}$$

which is derived from Lemma 5.3.2 and Lemma 5.3.4.

We now need to combine the ratios from (5.4.1). To do this we will use a summation lemma from [37], in the slightly different restated form which can be found in [13].

Lemma 5.4.1. [37, Corollary 4.5] *Let $N \geq 2$ be an integer and, for $1 \leq i \leq N$, let $A(i)$ and $C(i)$ be given real numbers such that $A(i) \geq 0$ and $A(i) - (i-1)C(i) \geq 0$. Define*

$$A_1 = \min_{i=1,\dots,N} A(i), \quad A_2 = \max_{i=1,\dots,N} A(i)$$

$$C_1 = \min_{i=1,\dots,N} C(i), \quad C_2 = \max_{i=1,\dots,N} C(i).$$

Suppose that there exists a real number $\hat{c}$ *with* $0 < \hat{c} < \frac{1}{3}$ *such that*

$$\max\{A_2/N,\ |C_1|, |C_2|\} \le \hat{c}.$$

Define $n_0, \ldots, n_N$ *by* $n_0 = 1$ *and* $n_i = \frac{1}{i}\,(A(i) - (i-1)C(i))\, n_{i-1}$ *for* $1 \le i \le N$. *Then*

$$\Sigma_1 \le \sum_{i=0}^{N} n_i \le \Sigma_2$$

where

$$\Sigma_1 = \exp\left(A_1 - \tfrac{1}{2}A_1C_2\right) - (2e\hat{c})^N,$$
$$\Sigma_2 = \exp\left(A_2 - \tfrac{1}{2}A_2C_1 + \tfrac{1}{2}A_2C_1^2\right) + (2e\hat{c})^N.$$

We will prove the following result.

Lemma 5.4.2. *Under the assumptions of Theorem 5.1.1,*

$$\sum_{d=0}^{N} \frac{|\mathcal{C}_d|}{|\mathcal{C}_0|} = \exp\left(\frac{(r-1)^2 M_2^2}{4\,M^2} + O\left(\frac{r^4\, k_{\max}^4}{M}\right)\right).$$

Proof. Define d' to be the first value of $d \le N$ for which $\mathcal{C}_d = \emptyset$, or $d' = N+1$ if no such value of d exists. By the assumptions of Theorem 5.1.1, the relative error in Lemma 5.3.2 is $o(1)$ while the main term $4dM^4$ tends to infinity. Therefore the number of legal forward switchings from any $G \in \mathcal{C}_d$ is always positive, and each such switching produces an element of $\mathcal{C}_{d-1}$. Therefore, if $\mathcal{C}_d$ is nonempty then so is $\mathcal{C}_{d-1}$, for any $d \ge 1$. This implies that $\mathcal{C}_d = \emptyset$ for $d' \le d \le N$. For $1 \le d < d'$, the ratio in (5.4.1) can be written as

$$\begin{aligned} \frac{|\mathcal{C}_d|}{|\mathcal{C}_0|} &= \frac{1}{d}\,\frac{|\mathcal{C}_{d-1}|}{|\mathcal{C}_0|}\ \left(\frac{(r-1)^2M_2^2}{4\,M^2} + O\Big(\frac{(r-1)^2M_2^2}{4\,M^2}\cdot\frac{(d\,k_{\max}\ + r^2\,k_{\max}^3)}{M_2}\Big)\right) \\ &= \frac{1}{d}\,\frac{|\mathcal{C}_{d-1}|}{|\mathcal{C}_0|}\ \left(A(d) - \ (d-1)\,C(d)\right), \end{aligned} \tag{5.4.2}$$

where α_d is some uniformly bounded function of d, and

$$A(d) = \frac{(r-1)^2 M_2^2 - \alpha_d\, r^4\, k_{\max}^3\, M_2}{4\, M^2}, \quad C(d) = \frac{\alpha_d\, r^2\, k_{\max} M_2}{4\, M^2}$$

Since $\mathcal{C}_d = \emptyset$ for $d' \leq d \leq N$, then equation (5.4.2) is also true in this case by taking $A(d) = C(d) = 0$.

We now check that all conditions of Lemma 5.4.1 hold. For $d' \leq d \leq N$, it is obvious that $A(d) = 0$ and $A(d) - (d-1)C(d) = 0$, from the definition of $A(d)$ and $C(d)$. Next, suppose that $1 \leq d \leq d'$. If $\alpha_d \geq 0$, then we conclude from $|\mathcal{C}_{d-1}| > 0$, $|\mathcal{C}_d| \geq 0$, and equation (5.4.2) that

$$A(d) \geq A(d) - (d-1)\, C(d) \geq 0.$$

Also, the definition of $A(d)$ shows that $A(d) \geq 0$ when $\alpha_d < 0$. Consider A_1, A_2, C_1 and C_2 as in the statement of Lemma 5.4.1 and let $A \in [A_1, A_2]$ and $C \in [C_1, C_2]$. Choose $\hat{c} = \frac{1}{7}$ and by the assumptions of Theorem 5.1.1, we have

$$A = \frac{(r-1)^2 M_2^2}{4\, M^2} + o(1), \quad C = O\left(\frac{r^2\, k_{\max}^2}{M}\right) = o(1).$$

Then by using the definition of N as stated in (5.2.3), it is easy to check that $\max\{A/N,\, C\} \leq \hat{c}$. The definition of C demonstrates that $|C_1| = |C_2| = o(1) \leq \hat{c}$.

Hence, Lemma 5.4.1 can be applied. Here we have

$$\begin{aligned} A_2 - \tfrac{1}{2} A_2 C_1 + \tfrac{1}{2} A_2 C_1^2 &= \frac{(r-1)^2 M_2^2}{4\, M^2} + O\left(\frac{r^4 k_{\max}^4}{M} + \frac{(r-1)^2 M_2^2}{4\, M^2} \cdot \frac{r^2\, k_{\max}^2}{4\, M}\right) \\ &= \frac{(r-1)^2 M_2^2}{4\, M^2} + O\left(\frac{r^4 k_{\max}^4}{M}\right). \end{aligned}$$

Here the first error term comes from A_2 and the second error term covers the terms involving $A_2 C_1$ and $A_2 C_1^2$, using the fact that $C_1 = o(1)$. Applying Lemma 5.4.1

provides the upper bound on the sum of $|\mathcal{C}_d|$ given by

$$\sum_{d=0}^{N} \frac{|\mathcal{C}_d|}{|\mathcal{C}_0|} \leq \exp\left(\frac{(r-1)^2 M_2^2}{4\,M^2} + O\left(\frac{r^4\,k_{\max}^4}{M}\right)\right) + (2e\hat{c})^N.$$

The value of N determines that $(2e\hat{c})^N \leq (2e/7)^{5\log M} \leq 1/M$, which implies that

$$\sum_{d=0}^{N} \frac{|\mathcal{C}_d|}{|\mathcal{C}_0|} \leq \exp\left(\frac{(r-1)^2 M_2^2}{4\,M^2} + O\left(\frac{r^4\,k_{\max}^4}{M}\right)\right). \tag{5.4.3}$$

If $d' = N+1$ then $A_1 = \frac{(r-1)^2 M_2^2}{4\,M^2} + O(r^4 k_{\max}^4/M)$, so the lower bound from Lemma 5.4.1 matches the upper bound from (5.4.3), within the stated error term.

It remains to find a lower bound when $1 \leq d' \leq N$. The definition of d' indicates that $\mathcal{C}_{d'} = \emptyset$ and $\mathcal{C}_{d'-1} \neq \emptyset$. Hence the number of legal reverse switchings from $\mathcal{C}_{d'-1}$ must be 0. Then using Lemma 5.3.4 shows that

$$M_2 = O\Big(d'\,k_{\max} + r^2\,k_{\max}^3\Big).$$

Since $d' \leq N = O\left(r^2\,k_{\max}^2 + \log M\right)$ then

$$M_2 = O\Big(k_{\max}(r^2\,k_{\max}^2 + \log M)\Big).$$

Then the upper bound from (5.4.3) satisfies

$$\begin{aligned}
&\exp\left(\frac{(r-1)^2 M_2^2}{M^2} + O\left(\frac{r^4 k_{\max}^4}{M}\right)\right)\\
&= \exp\left(O\left(\frac{r^2 k_{\max}^2 (r^2\,k_{\max}^2 + \log M)^2}{M^2} + \frac{r^4 k_{\max}^4}{M}\right)\right)\\
&= \exp\left(O\left(\frac{r^4\,k_{\max}^4}{M}\right)\right),
\end{aligned}$$

and we can use the trivial lower bound

$$\sum_{d=0}^{N} \frac{|\mathcal{C}_d|}{|\mathcal{C}_0|} \geq \frac{|\mathcal{C}_0|}{|\mathcal{C}_0|} = 1 = \exp\left(O\left(\frac{r^4\,k_{\max}^4}{M}\right)\right).$$

This shows that the lemma holds when $d' \in \{1, 2, \ldots, N\}$, completing the proof. $\square$

Finally, we can prove the main result of this chapter.

Proof of Theorem 5.1.1.

From (5.2.2), Lemma 5.2.2 and (5.3.1), we have

$$\begin{aligned}
\sum_{d=0}^{N} |\mathcal{C}_d| &= |\mathcal{B}_r^+(\boldsymbol{k})| \\
&= \left(1 + O\Big(\frac{r^4\, k_{\max}^4}{M}\Big)\right) |\mathcal{B}_r(\boldsymbol{k})| \\
&= \frac{M!}{(r!\,)^{M/r} \prod_{j=1}^{n} k_j!} \exp\left(-\frac{(r-1)\, M_2}{2M} + O\left(\frac{r^4 k_{\max}^4}{M}\right)\right).
\end{aligned}$$

Substitution from Lemma 5.4.2 into (5.3.2) gives

$$|\mathcal{L}_r(\boldsymbol{k})| = \frac{\sum_{d=0}^{N} |\mathcal{C}_d|}{(M/r)!} \exp\left(-\frac{(r-1)^2 M_2^2}{4\, M^2} + O\left(\frac{r^4\, k_{\max}^4}{M}\right)\right).$$

Then combining these equations implies that

$$\begin{aligned}
&|\mathcal{L}_r(\boldsymbol{k})| \\
&= \frac{M!}{(r!\,)^{M/r}\, (M/r)! \prod_{j=1}^{n} k_j!} \exp\left(-\frac{(r-1)\, M_2}{2M} - \frac{(r-1)^2 M_2^2}{4\, M^2} + O\left(\frac{r^4\, k_{\max}^4}{M}\right)\right),
\end{aligned}$$

which completes the proof. $\square$

To complete this chapter, we state a corollary which follows immediately from our calculations. This generalises a special case of the result of McKay et al. [69, Corollary 3] from regular bipartite graphs to half-regular bipartite graphs.

Corollary 5.4.3. *Suppose that $\boldsymbol{k}$, r and M satisfy the assumptions of Theorem 5.1.1. The probability that a random element of $\mathcal{B}_r(\boldsymbol{k})$ has girth at least 6 is given by*

$$\exp\left(-\frac{(r-1)^2\, M_2^2}{4\, M^2} + O\left(\frac{r^4\, k_{\max}^4}{M}\right)\right).$$

Proof. A simple bipartite graph has girth at least 6 if and only if it has no 4-cycles. The result follows by combining Lemma 5.2.2, (5.3.1) and Lemma 5.4.2. $\square$

Chapter 6

Conclusion

In this chapter we summarise our results discussed in this thesis and other topics for future work.

The aim of this research was to prove some asymptotic results for sparse uniform hypergraphs with given degree sequences. In all of our results we require that the maximum degree and the edge size do not grow too quickly. In Chapter 3 we provided our first result, Theorem 3.1.1, giving a formula for the approximate number of simple uniform hypergraphs with specified degrees which contain no edge of X, where X is a specified set of edges satisfying some conditions. Then we derived another result, Corollary 3.1.2, for the probability that a given hypergraph occurs as a subhypergraph in a random hypergraph with specified degree sequence $\boldsymbol{k}$. In Chapter 4, we discussed three applications of Corollary 3.1.2, determining the average numbers of three spanning subhypergraphs in uniform hypergraphs under various restrictions. These subhypergraphs are perfect matchings, loose Hamilton cycles and spanning hypertrees. Finally, in Chapter 5, we studied the asymptotic enumeration of linear uniform hypergraphs with a given degree sequence, extending a previous enumeration by Blinovsky and Greenhill [13]. The presentation of linear hypergraphs in terms of bipartite graphs and the switching used in Chapter 5 also enabled us to provide the probability that a random bipartite graphs has girth at least 6, under certain conditions.

There are still many other open questions regarding asymptotic enumeration for sparse hypergraphs.

Regarding the enumeration of hypergraphs which avoid a given set of edges, it may be possible to improve the expression of Theorem 3.1.1 using a more complex switching with more edges. This may allow the result to apply to larger maximum degree and edge size, and provide further significant terms. Consequently, an improved formula for $|\mathcal{H}_r(\boldsymbol{k}, X)|$ would also lead to improvements to Corollary 3.1.2, allowing wider applications of this result.

Considering linear hypergraphs, we are also interested in studying the number of linear hypergraphs with a given degree sequence avoiding a specified set of edges: that is, an analogue of Theorem 3.1.1. This would provide as a formula for the probability that a random linear hypergraph with given degrees contains a specified subhypergraph. Such a formula could then be applied to the same spanning subhypergraphs as studied in Chapter 4. Specifically, we want to study the expected number of spanning uniform hypertrees in random linear hypergraphs with given degrees. It would be interesting to compare such a result with the case of simple hypergraphs. Since all simple graphs are linear, it is possible that the asymptotic formula for the expected number of spanning hypertrees in simple linear uniform hypergraphs will more closely match the formula for graphs. This is a direction for future work.

Furthermore, it is possible to generalise all the enumeration results in this thesis to non-uniform hypergraphs, where the size of the edges may vary. The bipartite setting used in Chapter 5 is the most natural model for this generalisation, as considering different number of vertices in each edge of a hypergraph will correspond to bipartite graphs with irregular degrees for both vertex partition sets. In particular, this would generalise Corollary 5.4.3 from half-regular bipartite graphs to the irregular case. It should also be possible to apply the switching method to produce a more general result about girth in random bipartite graphs with given degrees: specifically, a formula for the probability that the girth is at least g. Such a result would extend the result of McKay, Wormald and Wysocka [69] for regular bipartite graphs.

Finally, we remark that there are real-world applications of directed hypergraphs, also called dihypergraphs. Here each edge has some specified in-vertices and some specified out-vertices. In particular, there are applications of dihypergraphs in the area of metabolic networks see for example [21]. To the best of my knowledge, there are no asymptotic enumeration results for dihypergraphs, which may be another topic for future work.

References

[1] H. S. Aldosari and C. Greenhill. The average number of spanning hypertrees in sparse uniform hypergraphs. *Discrete Mathematics, (to appear)* *arXiv:1907.04993*.

[2] H. S. Aldosari and C. Greenhill. Enumerating sparse uniform hypergraphs with given degree sequence and forbidden edges. *European Journal of Combinatorics*, 77:68–77, 2019.

[3] D. Altman, C. Greenhill, M. Isaev, and R. Ramadurai. A threshold result for loose Hamiltonicity in random regular uniform hypergraphs. *Journal of Combinatorial Theory, Series B*, 142:307–373, 2020.

[4] F. Amato, V. Moscato, A. Picariello, and G. Sperlí. Multimedia social network modeling: a proposal. In *2016 IEEE Tenth International Conference on Semantic Computing (ICSC)*, pages 448–453. IEEE, 2016.

[5] A. S. Asratian and N. Kuzjurin. On the number of partial Steiner systems. *Journal of Combinatorial Designs*, 8(5):347–352, 2000.

[6] R. Bacher. On the enumeration of labelled hypertrees and of labelled bipartite trees. *arXiv:1102.2708*, 2011.

[7] A. Barvinok. Matrices with prescribed row and column sums. *Linear Algebra and its Applications*, 436(4):820–844, 2012.

[8] A. Barvinok and J. A. Hartigan. The number of graphs and a random graph with a given degree sequence. *Random Structures & Algorithms*, 42(3):301–348, 2013.

[9] A. Békéssy, P. Békéssy, and J. Komlós. Asymptotic enumeration of regular matrices. *Studia Scientiarum Mathematicarum Hungarica*, 7:343–353, 1972.

[10] E. A. Bender. The asymptotic number of non-negative integer matrices with given row and column sums. *Discrete Mathematics*, 10:217–223, 1974.

[11] E. A. Bender and E. R. Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A*, 24(3):296–307, 1978.

[12] V. Blinovsky and C. Greenhill. Asymptotic enumeration of sparse uniform hypergraphs with given degrees. *European Journal of Combinatorics*, 51:287–296, 2016.

[13] V. Blinovsky and C. Greenhill. Asymptotic enumeration of sparse uniform linear hypergraphs with given degrees. *The Electronic Journal of Combinatorics*, 23(3):#P3.17, 2016.

[14] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[15] B. Bollobás. *Random Graphs*. Cambridge University Press, Cambridge, 2001.

[16] B. Bollobás and B. D. McKay. The number of matchings in random regular graphs and bipartite graphs. *Journal of Combinatorial Theory, Series B*, 41:80–91, 1986.

[17] V. Boonyasombat. Degree sequences of connected hypergraphs and hypertrees. In *Graph Theory Singapore 1983*, pages 236–247. Springer, 1984.

[18] E. R. Canfield, C. Greenhill, and B. D. McKay. Asymptotic enumeration of dense 0–1 matrices with specified line sums. *Journal of Combinatorial Theory, Series A*, 115(1):32–66, 2008.

[19] C. Cooper, A. Frieze, M. Molloy, and B. Reed. Perfect matchings in random r-regular, s-uniform hypergraphs. *Combinatorics, Probability and Computing*, 5:1–14, 1996.

[20] C. Cooper, A. Frieze, and B. Reed. Random regular graphs of non-constant degree: connectivity and Hamiltonicity. *Combinatorics, Probability and Computing*, 11(3):249–261, 2002.

[21] L. Cottret and F. Jourdan. Graph methods for the investigation of metabolic networks in parasitology. *Parasitology*, 137(9):1393, 2010.

[22] A. Ducournau, A. Bretto, S. Rital, and B. Laget. A reductive approach to hypergraph clustering: An application to image segmentation. *Pattern Recognition*, 45(7):2788–2803, 2012.

[23] A. Dudek, A. Frieze, A. Ruciński, and M. Šileikis. Approximate counting of regular hypergraphs. *Information Processing Letters*, 113(19-21):785–788, 2013.

[24] A. Dudek, A. Frieze, A. Ruciński, and M. Šileikis. Embedding the Erdős–Rényi hypergraph into the random regular hypergraph and Hamiltonicity. *Journal of Combinatorial Theory, Series B*, 122:719–740, 2017.

[25] A. Espuny Díaz, F. Joos, D. Kühn, and D. Osthus. Edge correlations in random regular hypergraphs and applications to subgraph testing. *SIAM Journal on Discrete Mathematics*, 33(4):1837–1863, 2019.

[26] C. Everett and P. Stein. The asymptotic number of integer stochastic matrices. *Discrete Mathematics*, 1(1):55–72, 1971.

[27] V. Fack and B. D. McKay. A generalized switching method for combinatorial estimation. *Australasian Journal of Combinatorics*, 39:141–154, 2007.

[28] P. Frankl and V. Rödl. Near perfect coverings in graphs and hypergraphs. *European Journal of Combinatorics*, 6(4):317–326, 1985.

[29] A. Frieze and S. Janson. Perfect matchings in random s-uniform hypergraphs. *Random Structures & Algorithms*, 7(1):41–57, 1995.

[30] A. Frieze, M. Jerrum, M. Molloy, R. Robinson, and N. Wormald. Generating and counting Hamilton cycles in random regular graphs. *Journal of Algorithms*, 21(1):176–198, 1996.

[31] A. Frieze and M. Karoński. *Introduction to Random Graphs*. Cambridge University Press, Cambridge, 2016.

[32] A. Goodall and A. de Mier. Spanning trees of 3-uniform hypergraphs. *Advances in Applied Mathematics*, 47(4):840–868, 2011.

[33] D. A. Grable and K. T. Phelps. Random methods in design theory: a survey. *Journal of Combinatorial Designs*, 4(4):255–273, 1996.

[34] C. Greenhill, M. Isaev, M. Kwan, and B. D. McKay. The average number of spanning trees in sparse graphs with given degrees. *European Journal of Combinatorics*, 63:6–25, 2017.

[35] C. Greenhill, M. Isaev, and G. Liang. Spanning trees in random regular uniform hypergraphs. *arXiv:2005.07350*, 2020.

[36] C. Greenhill and B. D. McKay. Asymptotic enumeration of sparse nonnegative integer matrices with specified row and column sums. *Advances in Applied Mathematics*, 41(4):459–481, 2008.

[37] C. Greenhill, B. D. McKay, and X. Wang. Asymptotic enumeration of sparse 0-1 matrices with irregular row and column sums. *Journal of Combinatorial Theory, Series A* 113(2):291–324, 2006.

[38] M. Hasheminezhad and B. D. McKay. Asymptotic enumeration of non-uniform linear hypergraphs. *Discussiones Mathematicae Graph Theory*, to appear.

[39] M. Hasheminezhad and B. D. McKay. Combinatorial estimates by the switching method. *Contemporary Mathematics*, 531:209–221, 2010.

[40] M. Isaev and B. D. McKay. Complex martingales and asymptotic enumeration. *Random Structures & Algorithms*, 52(4):617–661, 2018.

[41] S. Janson. The numbers of spanning trees, Hamilton cycles and perfect matchings in a random graph. *Combinatorics, Probability and Computing*, 3(1):97–126, 1994.

[42] S. Janson. Random regular graphs: asymptotic distributions and contiguity. *Combinatorics, Probability and Computing*, 4(4):369–405, 1995.

[43] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs*. John Wiley & Sons, New York, 2000.

[44] P. Jégou and S. N. Ndiaye. On the notion of cycles in hypergraphs. *Discrete Mathematics*, 309(23-24):6535–6543, 2009.

[45] A. Johansson, J. Kahn, and V. Vu. Factors in random graphs. *Random Structures & Algorithms*, 33(1):1–28, 2008.

[46] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Discrete Multivariate Distributions*. John Wiley & Sons, New York, 1997.

[47] J. Kahn. Asymptotics for Shamir's Problem. *arXiv:1909.06834*, 2019.

[48] J. Kahn. Hitting times for Shamir's Problem. *arXiv:2008.01605*, 2020.

[49] N. Kamčev, A. Liebenau, and N. Wormald. Asymptotic enumeration of hypergraphs by degree sequence. *arXiv:2008.07757*, 2020.

[50] J. H. Kim. Perfect matchings in random uniform hypergraphs. *Random Structures & Algorithms*, 23(2):111–132, 2003.

[51] S. Klamt, U.-U. Haus, and F. Theis. Hypergraphs and cellular networks. *PLoS Computational Biology*, 5(5):e1000385, 2009.

[52] A. Kolchin. On the number of hyperforests. *Journal of Mathematical Sciences*, 76(2):2250–2258, 1995.

[53] M. Krivelevich, B. Sudakov, V. H. Vu, and N. C. Wormald. Random regular graphs of high degree. *Random Structures & Algorithms*, 18(4):346–363, 2001.

[54] G. Kuperberg, S. Lovett, and R. Peled. Probabilistic existence of regular combinatorial structures. *Geometric and Functional Analysis*, 27(4):919–972, 2017.

[55] C. Lavault. A note on Prüfer-like coding and counting forests of uniform hypertrees. *arXiv:1110.0204*, 2011.

[56] J. M. Levine, J. Bascompte, P. B. Adler, and S. Allesina. Beyond pairwise mechanisms of species coexistence in complex communities. *Nature*, 546:56–64, 2017.

[57] A. Liebenau and N. Wormald. Asymptotic enumeration of graphs by degree sequence, and the degree sequence of a random graph. *arXiv:1702.08373*, 2019.

[58] L. Lu and L. A. Székely. A new asymptotic enumeration technique: the Lovász local lemma. *arXiv:0905.3983*, 2009.

[59] B. D. McKay. Spanning trees in random regular graphs. *Proceedings of the Third Carribean Conference on Combinatorics and Computing*, pages 139–143, 1981.

[60] B. D. McKay. Subgraphs of random graphs with specified degrees. *Congressus Numerantium*, 33:213–223, 1981.

[61] B. D. McKay. Asymptotics for 0-1 matrices with prescribed line sums. *Enumeration and Design*, pages 225–238, 1984.

[62] B. D. McKay. Asymptotics for symmetric 0-1 matrices with prescribed row sums. *Ars Combinatoria*, 19A:15–25, 1985.

[63] B. D. McKay. Subgraphs of random graphs with specified degrees. In *Proceedings of the International Congress of Mathematicians*, volume IV, pages 2489–2501, Hyderabad, 2010. World Scientific.

[64] B. D. McKay. Subgraphs of dense random graphs with specified degrees. *Combinatorics, Probability and Computing*, 20(3):413–433, 2011.

[65] B. D. McKay and F. Tian. Asymptotic enumeration of linear hypergraphs with given number of vertices and edges. *Advances in Applied Mathematics*, 115:102000, 2020.

[66] B. D. McKay and X. Wang. Asymptotic enumeration of 0-1 matrices with equal row sums and equal column sums. *Linear Algebra and its Applications*, 373:273–287, 2003.

[67] B. D. McKay and N. C. Wormald. Asymptotic enumeration by degree sequence of graphs of high degree. *European Journal of Combinatorics*, 11(6):565–580, 1990.

[68] 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.

[69] B. D. McKay, N. C. Wormald, and B. Wysocka. Short cycles in random regular graphs. *The Electronic Journal of Combinatorics*, page #R66, 2004.

[70] M. P. Mineev and A. I. Pavlov. On the number of (0,1)-matrices with prescribed sums of rows and columns. *Doklady Akademii Nauk*, 230(2):271–274, 1976.

[71] J. W. Moon. *Counting Labelled Trees*, volume 1 of *Canadian Mathematical Monographs*. Canadian Mathematical Congress, Montreal, 1970.

[72] T. Morimae, Y. Takeuchi, and M. Hayashi. Verification of hypergraph states. *Physical Review A*, 96(6):062321, 2017.

[73] P. E. O'Neil. Asymptotics and random matrices with row-sum and column-sum restrictions. *Bulletin of the American Mathematical Society*, 75:1276–1282, 1969.

[74] R. C. Read. *Some enumeration problems in graph theory*. PhD thesis, University of London, 1958.

[75] R. C. Read. The enumeration of locally restricted graphs (I). *Journal of the London Mathematical Society*, 34:417–436, 1959.

[76] R. C. Read. The enumeration of locally restricted graphs (II). *Journal of the London Mathematical Society*, 35:344–351, 1960.

[77] R. W. Robinson and N. C. Wormald. Almost all cubic graphs are Hamiltonian. *Random Structures & Algorithms*, 3(2):117–125, 1992.

[78] R. W. Robinson and N. C. Wormald. Almost all regular graphs are Hamiltonian. *Random Structures & Algorithms*, 5(2):363–374, 1994.

[79] J. Schmidt and E. Shamir. A threshold for perfect matchings in random d-pure hypergraphs. *Discrete Mathematics*, 45(2-3):287–295, 1983.

[80] B. I. Selivanov. Enumeration of homogeneous hypergraphs with a simple cycle structure. *Kombinatornyĭ Analiz*, 2:60–67, 1972.

[81] S. Shannigrahi and S. P. Pal. Efficient Prüfer-like coding and counting labelled hypertrees. *Algorithmica*, 54(2):208–225, 2009.

[82] W.-C. Siu. *Hypertrees in d-uniform Hypergraphs*. PhD thesis, Michigan State University, 2002.

[83] S. Sivasubramanian. Spanning trees in complete uniform hypergraphs and a connection to r-extended Shi hyperplane arrangements. *arXiv:math/0605083*, 2006.

[84] D. M. Warme. *Spanning trees in hypergraphs with applications to Steiner trees.* PhD thesis, University of Virginia, 1998.

[85] N. C. Wormald. *Some problems in the enumeration of labelled graphs.* PhD thesis, University of Newcastle, 1979.

[86] N. C. Wormald. Models of random regular graphs. *London Mathematical Society Lecture Note Series*, pages 239–298, 1999.

[87] N. C. Wormald. Asymptotic enumeration of graphs with given degree sequence. In *Proceedings of the International Congress of Mathematicians*, volume 4, pages 3263–3282, Rio de Janiero, 2018. World Scientific.

[88] Y. Zhu, X. Zhu, M. Kim, D. Kaufer, and G. Wu. A novel dynamic hypergraph inference framework for computer assisted diagnosis of neuro-diseases. In *International Conference on Information Processing in Medical Imaging*, pages 158–169. Springer, 2017.