SPANNING TREES IN RANDOM REGULAR UNIFORM HYPERGRAPHS

Gary Liang

Supervisor: A/Prof. Catherine Greenhill

School of Mathematics and Statistics
UNSW Sydney

October 2018

Submitted in partial fulfillment of the requirements of the degree of
Bachelor of Science (Advanced Mathematics) with Honours

Plagiarism statement

I declare that this thesis is my own work, except where acknowledged, and has not been submitted for academic credit elsewhere.

I acknowledge that the assessor of this thesis may, for the purpose of assessing it:

- Reproduce it and provide a copy to another member of the University; and/or,
- Communicate a copy of it to a plagiarism checking service (which may then retain a copy of it on its database for the purpose of future plagiarism checking).

I certify that I have read and understood the University Rules in respect of Student Academic Misconduct, and am aware of any potential plagiarism penalties which may apply.

By signing this declaration I am agreeing to the statements and conditions above.

Signed: ______________________________ Date: __________________

Acknowledgements

If I have seen further it is by standing on the shoulders of giants.

– Isaac Newton

In December 2017, I was travelling in Europe with my good friend Sebastian Jux. I was still recovering from a debilitating illness and a series of other setbacks which made me question my life choices. In my frustration, I had applied for a job as a mathematics teacher, and received an offer to teach in one of the top selective schools in Guangzhou, China. The school gave me 24 hours to think about whether I would take the job. I decided resolutely I was going to give up my Honours year, because I was sick of Sydney and my life back home.

As fate would have it, I receive an unprompted email from Catherine Greenhill with an idea for a project for the year. In a moment, everything changes. I make a 180 degree turn, and decide that, for the first time in my life, I need to take a year to spend all of my time concentrating on mathematics. Flash forward to today and it has been one of the best decisions I have ever made.

So first and foremost, the greatest thanks must go to my supervisor, Catherine Greenhill. I say with no exaggeration that she has changed the course of my life. I am grateful for her ongoing support and patience, and I am motivated to pursue further study in mathematics.

I would like to thank my high school mathematics teacher, Mr Paul Bigelow, for inspiring me to study mathematics in the first place.

I would like to thank the academic staff at UNSW. I have learnt a tremendous amount. I would like to thank Thomas Britz, Peter Brown, Ian Doust, Catherine Greenhill (again), David Harvey and Denis Potapov for their style of teaching.

I would like to thank the friends I spent time with this year: Mark, Tanya, Alan, Steph and Henry for making my final year at university thoroughly enjoyable; Ben and Kieran for convincing me to take a two-week holiday in Korea and Japan a month before my thesis due date; Jerry, Justin and Kevin for their unwavering support and genuine care; James for his fashion advice. I have come out of this year with friendships stronger than ever before.

I would like to thank the friends I have made through my mathematics degree. Learning mathematics at UNSW has been a remarkably humbling experience. There have been a countless number of times where someone has helped me solve a problem or understand a concept. I have no doubt my peers will achieve amazing things. Special thanks to Marley for proofreading my thesis.

Last but not least, I would like to thank my family. None of this could be done without them. I am especially grateful for the last moments I spent with my grandfather this year. I hope I made you proud.

Gary Liang, October 2018.

Abstract

In 1992, Robinson and Wormald introduced the small subgraph conditioning method when they proved that almost all cubic graphs are Hamiltonian. Since then, there have been many applications of the method to derive the asymptotic distribution of random variables which count certain large subgraphs in a random regular graph. In particular, Greenhill, Kwan and Wind used this method to determine the asymptotic distribution of the number of spanning trees in a random cubic graph. They conjectured that a similar result holds more generally for a random r-regular graph, for any fixed $r \geq 4$.

We give an overview of random regular graphs, the configuration model and the small subgraph conditioning method. We also describe some interesting enumeration results on trees. In doing so, we lay the groundwork for the first piece of original work in this thesis: proving that the conjecture of Greenhill, Kwan and Wind holds true for $r \geq 4$.

Hypergraphs are a generalisation of graphs where an edge is allowed to contain more than 2 vertices, and a uniform hypergraph is one where the size of each edge is the same. We generalise the aforementioned problem and ask the question: How many spanning trees are there in a random regular uniform hypergraph? We derive a threshold result which allows us to asymptotically determine the existence of spanning trees. We also make a conjecture for the asymptotic distribution of the number of spanning trees in a random regular uniform hypergraph, and use the small subgraph conditioning method to prove this conjecture in the case that every edge contains 3 vertices.

Contents

CHAPTER 1

Introduction

Do not go where the path may lead, go instead where there is no path and leave a trail.

– Ralph Waldo Emerson

In 1959, Erdős and Rényi introduced what is now called the *probabilistic method* to prove the existence of graphs with certain properties. The idea is to define a probability space over the set of graphs with n vertices, and to show that a random n-vertex graph satisfies the required properties with positive probability, for large enough n. Since then, random graphs have become useful in other areas of mathematics, as well as outside mathematics. Random graphs have been used to model complex networks, and have been useful in computer science, physics, social science, biology and neuroscience (see [19] for a good overview on some applications).

About 15-20 years after Erdős' paper, the study of *random regular graphs* began with the works of Bender and Canfield [7], Bollobás [8] and Wormald [31, 32]. This involves studying random graphs in which the degrees of vertices are restricted. We usually study random regular graphs via another related probability space called the configuration model, which is described in Chapter 2. Interested readers can refer to Wormald [33] for more background.

In graph theory, researchers are interested in the existence of and the number of certain subgraphs. Some of these problems, like determining the existence of a Hamilton cycle, are NP-complete. Thus, the next best plan of attack is to find results that hold probabilistically rather than deterministically. The small subgraph conditioning method gives us a tool to analyse random variables of interest, such as the number of Hamilton cycles or spanning trees in a random regular graph. By conditioning on the number of certain small subgraphs, usually the short cycles of given lengths, the method (mysteriously) gives us the asymptotic distribution of such a random variable.

Hypergraphs are generalisations of graphs where the size of each edge is not restricted to two vertices. An s-uniform hypergraph is one where each edge contains s vertices. It has been shown that for every $s \geq 4$, the problem of deciding the existence of a spanning tree in a given s-uniform hypergraph is NP-complete [4]. The main motivating question of this thesis is: What is number of spanning trees in a random r-regular s-uniform hypergraph? We will attempt to answer the question through the application of the small subgraph conditioning method.

1.1 Structure of thesis

The structure of this thesis is as follows. Chapters 2 and 3 describe the background of random regular uniform hypergraphs and trees. This lays the foundation for Chapters 4, 5 and 6, which contain original results relating to the number of spanning trees in a random r-regular s-uniform hypergraph.

Chapter 2 gives an overview of basic definitions in graph theory, the configuration model and the small subgraph conditioning method. We state some useful results which help with the verification of the hypotheses of the small subgraph conditioning method, and provide a detailed proof that almost all cubic graphs are Hamiltonian. We then generalise the configuration model to hypergraphs.

Chapter 3 describes some results relating to the enumeration of trees, which become useful in subsequent chapters.

Chapter 4 contains results relating to the expected number of spanning trees in the configuration model. We perform analysis on the expression obtained, deriving a threshold result which helps determine the asymptotic existence of spanning trees in a random r-regular s-uniform hypergraph.

Chapter 5 contains a technical proof of an asymptotic expression for the joint moment between the number of spanning trees and the number of short cycles, a crucial component in the verification of the conditions of the small subgraph conditioning method.

Chapter 6 contains the proof of the conjecture of Greenhill, Kwan and Wind about the asymptotic distribution of the number of spanning trees in a random r-regular graph. We conjecture that a generalisation of this result olds true for r-regular s-uniform hypergraphs, and prove it for the case $s = 3$.

1.2 Notation

We use $\mathbb{N}$ to mean the natural numbers, including zero, and $\mathbb{Z}^+$ to mean the positive integers. We use $[n]$ to denote the set $\{1, \ldots, n\}$. For a non-negative integer j and real x, we write $(x)_j$ to denote the falling factorial

$$x(x-1)\ldots(x-j+1).$$

We use the notation Pr for probability and $\mathbb{E}$ for expectation.

For functions f, g, we write $f(n) \sim g(n)$ to mean

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1.$$

We write $f(n) = O(g(n))$ to mean that there exists $M > 0$ such that, for sufficiently large n,

$$|f(n)| \leq M\,|g(n)|\,.$$

Note that $f(n) \sim g(n)$ implies $f(n) = O(g(n))$. We write $f(n) = o(g(n))$ to mean

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0.$$

CHAPTER 2

Random regular graphs and random regular uniform hypergraphs

A journey of a thousand miles begins with a single step.

– 老子 (Laozi)

In this chapter, we begin with some basic definitions and an overview of the configuration model and the small subgraph conditioning method.

2.1 Graphs and random graphs

We review some basic definitions in graph theory and random graph theory. Diestel [13] is a good reference.

Definition 2.1. A (simple) *graph* $G = (V, E)$ consists of a set of *vertices* V and a set E of 2-element subsets of V. Elements of E are called *edges*.

Definition 2.2. A *multigraph* $G = (V, E)$ consists of a set of vertices V, and a multiset E of 2-element multisubsets of V. Elements of E are called *edges*.

In other words, multigraphs are allowed to have loops and multiple edges. We sometimes write an edge $\{a, b\}$ as ab for brevity.

Example 2.3. Figure 2.1(a) is a graph with $V = \{1, 2, 3\}$ and $E = \{12, 23, 13\}$. Figure 2.1(b) is a non-simple multigraph with $V = \{1, 2, 3\}$ and $E = \{13, 13, 23, 22\}$.

(a) A graph. (b) A non-simple multigraph.

Figure 2.1

Definition 2.4. For graphs $G = (V, E)$ and $G' = (V', E')$, if $V' \subseteq V$ and $E' \subseteq E$, then G' is a *subgraph* of G. In addition, if $V' = V$, then we say that G' is a *spanning subgraph* of G.

Definition 2.5. A *path of length k* is a sequence $v_1, e_1, v_2, e_2, \ldots, e_k, v_{k+1}$ where the e_i are distinct edges and the v_i are distinct vertices such that $v_i, v_{i+1} \in e_i$. If there is no ambiguity, we sometimes refer to a path by a sequence of vertices $v_1 v_2 \ldots v_{k+1}$ or sequence of edges $e_1 e_2 \ldots e_k$.

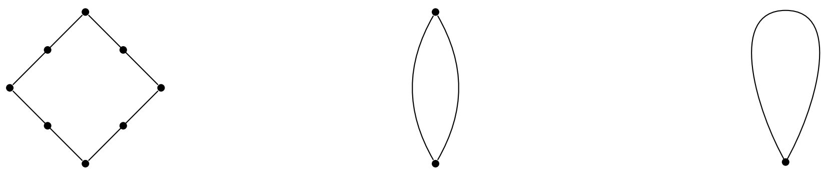
Definition 2.6. A graph is *connected* if there is a path between every pair of vertices.

Definition 2.7. A 1-*cycle* (or *loop*) is a multigraph with one vertex x and one edge xx. A 2-*cycle* is a multigraph of the form $V = \{x_1, x_2\}$, $E = \{x_1x_2, x_1x_2\}$. For $j \geq 3$, a j-*cycle* is a graph $G = (V, E)$ of the form

$$V = \{x_1, \ldots, x_j\}, \qquad E = \{x_1x_2, x_2x_3, \ldots, x_{j-1}x_j, x_jx_1\},$$

where all the x_i are distinct. There are j vertices and j edges in a j-cycle. If there is no ambiguity, we sometimes refer to a j-cycle by a sequence of vertices $x_1x_2 \ldots x_jx_1$.

Example 2.8. Figure 2.2 shows an 8-cycle, a 2 cycle and a 1-cycle.



(a) An 8-cycle. (b) A 2-cycle. (c) A 1-cycle.

Figure 2.2

Definition 2.9. A *Hamilton cycle* of a graph G is a subgraph which is a cycle that contains every vertex of G. A graph is *Hamiltonian* if it contains a Hamilton cycle.

We also define a tree, the main object of study in this thesis.

Definition 2.10. A graph with no cycles is called a *forest.* A connected graph with no cycles is called a *tree.* A *spanning tree* of a graph G is a spanning subgraph of G which is a tree. A vertex of degree 1 in a tree is called a *leaf.*

Example 2.11. Figure 2.3 is an example of a tree on 6 vertices. The leaves of this tree are 3, 5 and 6.

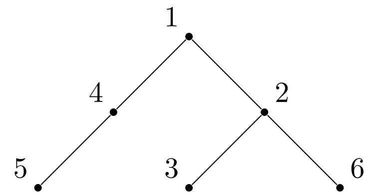


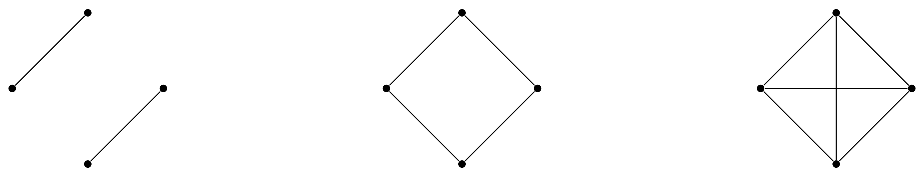
Figure 2.3: A tree with 6 vertices.

We are interested in graphs with some more structure.

Definition 2.12. The *degree* of a vertex v, denoted $\deg(v)$, is the number of edges which contain v.

Definition 2.13. For an integer $r \geq 1$, a r*-regular* (multi)graph is one where all vertices have degree r. A 3-regular graph is called a *cubic graph.*

Example 2.14. Figure 2.4 shows some regular graphs on $n = 4$ vertices.



(a) A 1-regular graph. (b) A 2-regular graph. (c) A 3-regular graph.

Figure 2.4

The first result every graph theory student learns is the Handshaking Lemma.

Lemma 2.15 (Handshaking Lemma)**.** *For a graph $G = (V, E)$, we have*

$$\sum_{v \in V} \deg(v) = 2\,|E|\,.$$

Proof. We have

$$\sum_{v \in V} \deg(v) = \sum_{v \in V} \sum_{e \in E : v \in e} 1 = \sum_{e \in E} \sum_{v \in V : v \in e} 1 = \sum_{e \in E} 2 = 2\,|E|\,,$$

making use of a double counting argument. □

How does one generate graphs randomly? The uniform model and the binomial model are two random graph models.

- **Uniform model.** Let $\mathcal{S}_n$ be the set of all graphs with vertex set $[n]$. Then $|\mathcal{S}_n| = 2^{\binom{n}{2}}$. Define a uniform probability distribution over the set $\mathcal{S}_n$, so $\Pr(G) = 2^{-\binom{n}{2}}$ for each $G \in \mathcal{S}_n$.
- **Binomial model** (Gilbert, 1959)**.** Let $V = [n]$, and fix $p \in [0, 1]$. Independently for each pair of vertices $\{i, j\}$, an edge ij is included with probability p. This random graph model is denoted $G(n, p)$. Note that $G(n, \frac{1}{2})$ is the uniform model.

What if we restrict ourselves to regular graphs? Let $\mathcal{S}_{n,r}$ be the set of all r-regular graphs on the vertex set $[n]$ and let $\mathcal{G}_{n,r}$ denote the uniform probability space over $\mathcal{S}_{n,r}$.

Example 2.16. If $n = 4$ and $r = 2$, then $\mathcal{G}_{4,2}$ is the uniform probability space over $\mathcal{S}_{4,2}$. That is, in $\mathcal{G}_{4,2}$, each of the graphs in Figure 2.5 are assigned a probability of $\frac{1}{3}$.

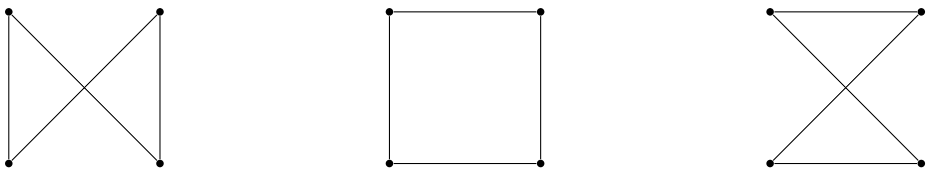


Figure 2.5: Elements of $\mathcal{S}_{4,2}$.

If $n = 4$ and $r = 3$, then $\mathcal{G}_{4,3}$ is the uniform probability space over a single graph K_4 (the complete graph on 4 vertices), as shown in Figure 2.6.

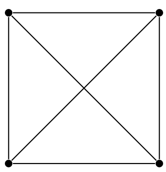


Figure 2.6: The only element in $\mathcal{S}_{4,3}$.

Neither the uniform nor the binomial model is useful in the uniform generation of random regular graphs. This is because, in a random regular graph, the existence of an edge is not independent of the existence of other edges. We therefore work in another model called the *configuration model.* The configuration model is closely related probability space which allows us to perform computations more easily.

2.2 Configuration model

We now describe the *configuration model* or *pairing model* of random regular graphs, given in this form in 1980 by Bollabás in [8].

Suppose $r \geq 1$. Consider a set of rn *points* partitioned into n *cells* $v_1, \dots, v_n$ of r points each, where rn is even. A perfect matching (or partition into 2-element disjoint subsets) of the points into $\frac{1}{2}rn$ *parts* (or *pairs*) is called a *partition* (or *pairing* or *configuration*). A partition P projects to a multigraph $G(P)$ by "collapsing" each cell into a single vertex. That is, a part $\{x, y\}$ in P corresponds to an edge $\{v_i, v_j\}$ of $G(P)$, where $x \in v_i$ and $y \in v_j$. Here, loops and multiple edges may arise.

Each (simple) graph corresponds to precisely $(r!)^n$ partitions. Hence, a regular graph can be chosen uniformly at random by choosing a partition uniformly at random and rejecting the result if it has loops or multiple edges. Non-simple multigraphs are not produced uniformly at random, as the probability depends on the number of loops and edges of each multiplicity.

Denote the set of possible partitions $\Omega_{n,r}$ and the uniform probability space over $\Omega_{n,r}$ by $\mathcal{P}_{n,r}$. A *subpartition* P' is a subset of a partition $P \in \Omega_{n,r}$, and note that a subpartition will project to a subgraph of $G(P)$. We will often refer to properties of a partition P, when we are in fact referring to properties of its projection $G(P)$ as a multigraph.

Example 2.17. Consider the case $n = 6$ and $r = 3$: the cubic graphs on 6 vertices. There are $\frac{18!}{9!2^9} = 34\,459\,425$ partitions in $\Omega_{6,3}$. One such partition is given on the left in Figure 2.7, which projects to the multigraph on the right. In this case, the corresponding multigraph has multiple edges and loops so it is rejected.

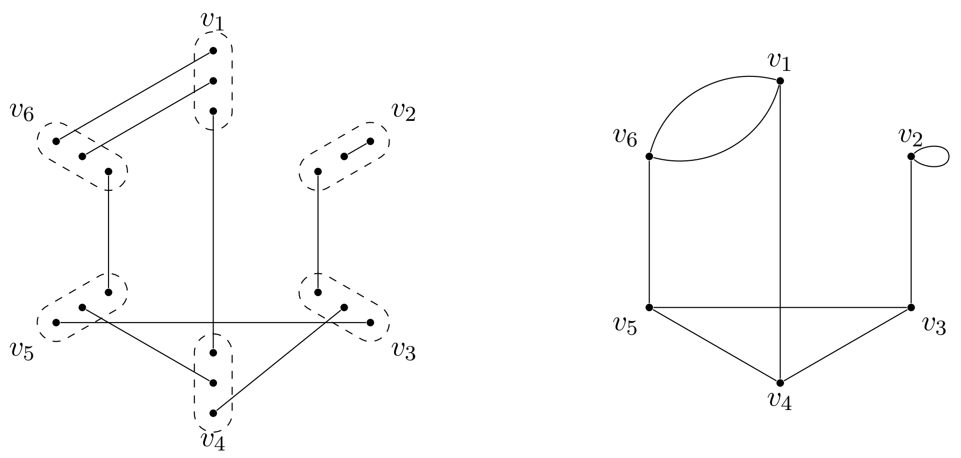


Figure 2.7: Rejected partition for $n = 6$ and $r = 3$.

On the other hand, Figure 2.8 shows a partition that projects to a graph (with no loops and multiple edges).

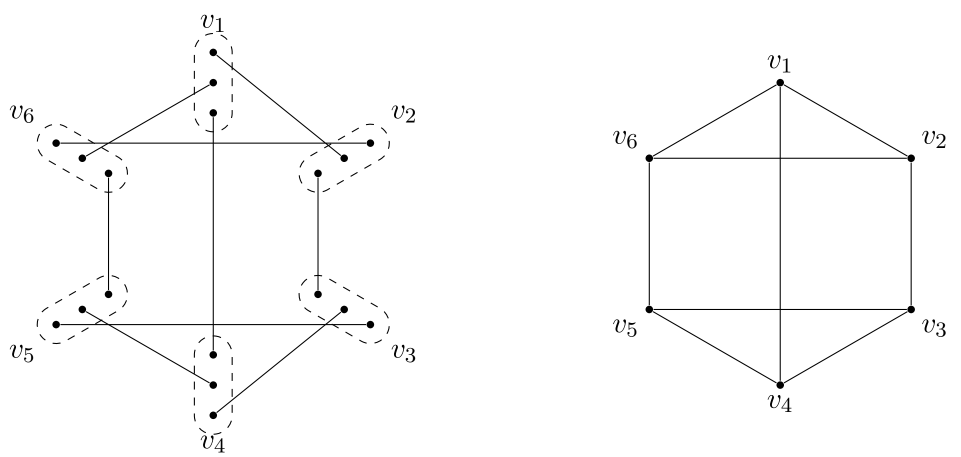


Figure 2.8: Accepted partition for $n = 6$ and $r = 3$.

Remark 2.18. We distinguish between $\mathcal{S}_{n,r}$, the set of r-regular graphs on n vertices, and $\mathcal{G}_{n,r} = (\mathcal{S}_{n,r}, 2^{\mathcal{S}_{n,r}}, \Pr)$, the uniform probability space over $\mathcal{S}_{n,r}$. Likewise, we distinguish between $\Omega_{n,r}$, the set of all partitions, and $\mathcal{P}_{n,r} = (\Omega_{n,r}, 2^{\Omega_{n,r}}, \Pr)$, the uniform probability space over $\Omega_{n,r}$.

	Graphs	Partitions
Probability space	$\mathcal{G}_{n,r}$	$\mathcal{P}_{n,r}$
Set	$\mathcal{S}_{n,r}$	$\Omega_{n,r}$

The configuration model allows us to prove properties of $\mathcal{G}_{n,r}$ by performing computations in $\mathcal{P}_{n,r}$ and conditioning on the event that the corresponding multigraph has no loops or multiple edges. For example, it provides an easy way to enumerate r-regular graphs asymptotically. The number of perfect matchings of t points, where t is even, is

$$f(t) = \frac{t!}{(t/2)!\,2^{t/2}}, \tag{2.1}$$

so the number of partitions in $\Omega_{n,r}$ is

$$|\Omega_{n,r}| = f(rn) = \frac{(rn)!}{(rn/2)!\,2^{rn/2}}. \tag{2.2}$$

Hence, the number of r-regular graphs on n vertices is precisely

$$|\mathcal{S}_{n,r}| = \frac{(rn)!\ \Pr(\text{Simple})}{(rn/2)!\,2^{rn/2}\,(r!)^n},$$

where "Simple" is the event that the corresponding multigraph to a partition has no loops or multiple edges, and we divide by $(r!)^n$ as each graph corresponds to precisely $(r!)^n$ partitions. Thus, an asymptotic formula for $|\mathcal{S}_{n,r}|$ can be found by estimating $\Pr(\text{Simple})$. Bender and Canfield [7] showed that, for fixed r, we have

$$\Pr(\text{Simple}) \sim \exp\left(\frac{1-r^2}{4}\right),$$

so by applying Stirling's approximation (see Corollary 2.26), we have

$$|\mathcal{S}_{n,r}| \sim \sqrt{2}e^{(1-r^2)/4}\left(\frac{r^r n^r}{e^r (r!)^2}\right)^{n/2}.$$

2.3 Small subgraph conditioning method

2.3.1 The theorem

Robinson and Wormald showed in [25] and [26] that almost all r-regular graphs are Hamiltonian, for any fixed $r \geq 3$. The proofs essentially employed the same general method, which is now known as the *small subgraph conditioning method*. Janson [18] stated the small subgraph conditioning method in its current form in 1995. We restate the theorem here (with slightly different notation).

Theorem 2.19 (Janson [18, Theorem 1]). *Let $\lambda_j > 0$ and $\zeta_j \geq -1$, $j = 1, 2, \ldots,$ be constants and suppose that for each n there are random variables $X_{j,n}$, $j = 1, 2, \ldots,$ and Y_n (defined on the same probability space) such that $X_{j,n}$ is a non-negative integer valued and $\mathbb{E}Y_n \neq 0$ (at least for large n), and furthermore the following conditions are satisfied:*

(A1) $X_{j,n} \xrightarrow{d} Z_j$ *as $n \to \infty$ jointly for all j, where $Z_j \sim \mathrm{Po}(\lambda_j)$ are independent Poisson random variables;*

(A2) *For any finite sequence $x_1, \ldots, x_m$ of non-negative integers,*

$$\frac{\mathbb{E}(Y_n | X_{1,n} = x_1, \ldots, X_{m,n} = x_m)}{\mathbb{E}Y_n} \to \prod_{j=1}^{m} (1 + \zeta_j)^{x_j} e^{-\lambda_j \zeta_j} \qquad \text{as } n \to \infty;$$

(A3) $\sum_{j \geq 1} \lambda_j \zeta_j^2 < \infty$;

(A4) $$\frac{\mathbb{E}Y_n^2}{(\mathbb{E}Y_n)^2} \to \exp\left(\sum_{j \geq 1} \lambda_j \zeta_j^2\right) \qquad \text{as } n \to \infty.$$

Then

$$\frac{Y_n}{\mathbb{E}Y_n} \xrightarrow{d} W = \prod_{j=1}^{\infty} (1 + \zeta_j)^{Z_j} e^{-\lambda_j \zeta_j} \qquad \text{as } n \to \infty;$$

moreover, this and the convergence in (A1) *hold jointly. The infinite product defining W converges asymptotically almost surely and in L^2, with*

$$\mathbb{E}W = 1 \qquad \text{and} \qquad \mathbb{E}W^2 = \exp\left(\sum_{j \geq 1} \lambda_j \zeta_j^2\right) = \lim_{n \to \infty} \frac{\mathbb{E}Y_n^2}{(\mathbb{E}Y_n)^2}.$$

In particular, $W > 0$ almost surely if and only if every $\zeta_j > -1$.

Janson remarks in [18] that the index set $\mathbb{Z}^+$ may be replaced by any other countably infinite set, and that 0^0 is defined to be 1.

As mentioned, the small subgraph conditioning method was first introduced by Robinson and Wormald, when they proved that almost all r-regular graphs are Hamiltonian.

Theorem 2.20 (Robinson and Wormald [26])**.** *Let $r \geq 3$ be an integer and let H_n be the number of Hamilton cycles in G from $\mathcal{G}_{n,r}$. Then* $\Pr(H_n > 0) \to 1$ *as $n \to \infty$.*

We will provide a detailed proof of this theorem for the case $r = 3$ in the next section to illustrate the small subgraph conditioning method in a relatively simple setting.

As the small subgraph conditioning method has become more well known, a number of asymptotic results about random regular graphs have been discovered. The number of perfect matchings in a random regular graph was essentially studied by Robinson and Wormald [26], and the result given explicitly by Janson [18]. In addition, the number of 1-factorisations in a random regular graph was studied by Janson [18] and the number of 2-factors in a random regular graph was studied by Robalewska [24].

Tutte conjectured in 1972 that every 4-edge connected graph has a nowhere-zero 3-flow, and Prałat and Wormald [23] used the small subgraph conditioning method to prove this conjecture holds asymptotically. Delcourt and Postle [12] proved that random 4-regular graphs have 3-star decompositions a.a.s., provided that the number of vertices is divisible by 3.

There have also been some results from the small subgraph conditioning method outside of random regular graphs. Coja-Oghlan and Wormald [10] used the method to determine the limiting distribution of the logarithm of the number of satisfying assignments of random regular k-SAT formulas. We refer to the original papers for the relevant definitions.

Of particular interest in this thesis, Greenhill, Kwan and Wind [16] used the small subgraph conditioning method to determine the asymptotic distribution of spanning trees for cubic graphs.

Theorem 2.21 (Greenhill, Kwan and Wind [16, Theorem 1.2])**.** *Define*

$$\lambda_j = \frac{2^j}{2j}, \qquad \zeta_j = -\frac{2^{j+1} - 1}{2^{2j}}.$$

Let $Y_{\mathcal{G}}$ be the number of spanning trees in a random graph G from $\mathcal{G}_{n,3}$. Then the asymptotic distribution of this quantity is

$$\frac{Y_{\mathcal{G}}}{\mathbb{E}Y_{\mathcal{G}}} \xrightarrow{d} \prod_{j=3}^{\infty} (1 + \zeta_j)^{Z_j} \, e^{-\lambda_j \zeta_j},$$

where $Z_j \sim \mathrm{Po}(\lambda_j)$ are independent Poisson random variables.

Greenhill et al. also conjectured in [16, Conjecture 1.3] that an analogous result holds for arbitrary fixed $r \geq 3$. The first original work of this thesis is the proof of this conjecture, which will be shown in Chapter 6.

Theorem 2.22 (Liang). *For fixed $r \geq 3$, define*

$$\lambda_j = \frac{(r-1)^j}{2j}, \qquad \zeta_j = -\frac{2(r-1)^j - 1}{(r-1)^{2j}}.$$

Let $Y_{\mathcal{G}}$ be the number of spanning trees in a random graph G from $\mathcal{G}_{n,r}$. Then the asymptotic distribution of this quantity satisfies

$$\frac{Y_{\mathcal{G}}}{\mathbb{E}Y_{\mathcal{G}}} \xrightarrow{d} \prod_{j=3}^{\infty} (1+\zeta_j)^{Z_j} \, e^{-\lambda_j \zeta_j}$$

where $Z_j = \mathrm{Po}(\lambda_j)$ are independent Poisson random variables.

2.3.2 Verifying the conditions

In essence, Theorem 2.19 is a statement about random variables. For a random variable of interest Y (such as the number of Hamilton cycles or the number of spanning trees in a random partition P from $\mathcal{P}_{n,r}$) and a series of random variables X_j (usually the number of j-cycles in a random partition P from $\mathcal{P}_{n,r}$), if we can verify conditions (A1)–(A4), then we can determine the asymptotic distribution of Y.

In almost all applications of the small subgraph conditioning method, X_j is the number of j-cycles in a random r-regular multigraph. Thus, it is common to cite a result from Bollabás [8] which directly establishes (A1).

Theorem 2.23 (Bollobás [8]). *For fixed r, let $X_{j,n}$ be the number of cycles of length j in the random multigraph coming from a partition in $\mathcal{P}_{n,r}$. For $k \geq 1$, $X_{1,n}, \ldots, X_{k,n}$ are asymptotically independent Poisson random variables with means $\lambda_j = \frac{(r-1)^j}{2j}$.*

Condition (A3) is also relatively easy to establish. Once the values of λ_j and ζ_j are known from (A1) and (A2), we usually use Taylor series to compute the infinite sum. The difficult conditions to establish are (A2) and (A4), both of which usually require lengthy combinatorial arguments, use of generating functions (see Wilf [30] for a good introduction) and asymptotic summation techniques such as Laplace's method.

In order to verify condition (A2), the following lemma is helpful.

Lemma 2.24 (Janson [18, Lemma 1])**.** *Let $\lambda'_j \geq 0$, $j = 1, 2, \ldots$ be constants. Suppose that* (A1) *holds, that $Y_n \geq 0$ and that*

$$\text{(A2}'\text{)} \qquad \frac{\mathbb{E}(Y_n (X_{1,n})_{x_1} \ldots (X_{m,n})_{x_m})}{\mathbb{E}Y_n} \to \prod_{j=1}^{m} (\lambda'_j)^{x_i} \qquad \text{as } n \to \infty,$$

for every finite sequence $x_1, \ldots, x_m$ of non-negative integers. Then condition (A2) *holds with $\lambda'_j = \lambda_j(1 + \zeta_j)$.*

When proving that (A2′) holds, it is standard in the literature to only give full details for $\mathbb{E}(YX_j)$, as the calculations for a general finite sequence $x_1, \ldots, x_m$ follow similarly.

In the next section, we will use the small subgraph conditioning method to prove that almost all cubic graphs are Hamiltonian. Before this, we state some useful results which will help us verify the four conditions.

Lemma 2.25. *For fixed $a, b \in \mathbb{R}$, $a > 0$, we have*

$$(an + b)! = \sqrt{2\pi}\,(an)^{b+\frac{1}{2}} \left(\frac{an}{e}\right)^{an} e^{O(1/n)}$$

as $n \to \infty$, with any restrictions on n to ensure that $an + b$ is a non-negative integer.

Proof. We use Stirling's formula in the form:

$$N! = \sqrt{2\pi N} \left(\frac{N}{e}\right)^{N} e^{O(1/N)}.$$

Thus,

$$(an + b)! = \sqrt{2\pi(an + b)} \left(\frac{an + b}{e}\right)^{an+b} e^{O(1/n)}.$$

Now, we have

$$(an + b)^{an+b} = (an)^{an+b} \left(1 + \frac{b}{an}\right)^{an+b} = (an)^{an+b} e^{b+O(1/n)},$$

as $\left(1 + \frac{b}{an}\right)^{n} \to e^{b/a}$ from below. Therefore, we have

$$(an + b)! = \sqrt{2\pi(an + b)}\, \frac{(an)^{an+b}}{e^{an}}\, e^{O(\frac{1}{n})} = \sqrt{2\pi}\,(an)^{b+\frac{1}{2}} \left(\frac{an}{e}\right)^{an} e^{O(1/n)},$$

as required. □

As an immediate corollary of Lemma 2.25, we have the following.

Corollary 2.26. *For fixed $a, b \in \mathbb{R}$, $a > 0$, we have*

$$(an+b)! \sim \sqrt{2\pi}\,(an)^{b+\frac{1}{2}}\left(\frac{an}{e}\right)^{an}$$

as $n \to \infty$, with restrictions on n to ensure that $an + b$ is a non-negative integer.

It is convenient to leave the expression in this form to isolate the *exponential factor*. This becomes useful especially when applying Laplace's method to perform the asymptotic summation.

Greenhill, Janson and Ruciński [15] proved a version of Laplace's method for asymptotic summation, tailored for the small subgraph conditioning method. A *lattice* $\mathcal{L}$ is a discrete subgroup of $\mathbb{R}^m$ and the determinant of $\mathcal{L}$ is the volume of a "unit cell". We refer to [15] for precise definitions. We restate the following lemma from [15, Lemma 6.3].

Lemma 2.27. *Suppose the following:*

(i) $\mathcal{L} \subset \mathbb{R}^m$ *is a lattice with full rank* m.

(ii) $K \subset \mathbb{R}^m$ *is a compact convex set with non-empty interior.*

(iii) $\phi : K \to \mathbb{R}$ *is a continuous function with a unique maximum at some interior point* $x_0 \in \text{int}(K)$.

(iv) ϕ *is a twice continuously differentiable in a neighbourhood of* x_0 *and the Hessian* $H_0 := D^2\phi(x_0)$ *is strictly negative definite.*

(v) $\psi : K_1 \to \mathbb{R}$ *is a continuous function on some neighbourhood* $K_1 \subset K$ *of* x_0 *with* $\psi(x_0) > 0$.

(vi) For each positive integer n *there is a vector* $\ell_n \in \mathbb{R}^m$.

(vii) For each positive integer n *there is a positive real number* b_n *and a function* $a_n : (\mathcal{L} + \ell_n) \cap nK \to \mathbb{R}$ *such that, as* $n \to \infty$,

$$a_n(\ell) = O(b_n e^{n\phi(\ell/n)+o(n)}), \qquad \ell \in (\mathcal{L} + \ell_n) \cap nK, \tag{2.3}$$

and

$$a_n(\ell) = b_n(\psi(\ell/n) + o(1))e^{n\phi(\ell/n)}, \qquad \ell \in (\mathcal{L} + \ell_n) \cap nK_1, \tag{2.4}$$

uniformly for ℓ *in the indicated sets.*

Then, as $n \to \infty$,

$$\sum_{(\mathcal{L}+\ell_n)\cap nK} a_n(\ell) \sim \frac{(2\pi)^{m/2}\psi(x_0)}{\det(\mathcal{L})\det(-H_0)^{1/2}} b_n n^{m/2} e^{n\phi(x_0)}.$$

The idea behind Laplace's method is that the sum estimated can be thought of as approximating an integral, and as n increases, the significant contributions to the integral come from points in the neighbourhood of the unique global maximum x_0. So we approximate the function with a Gaussian function around x_0, a function we know how to integrate. In Figure 2.9, we have $\phi(x) = \frac{\sin x}{x}$ with a global maximum at $x_0 = 0$, and as n grows, the approximation by a Gaussian function is getting better.

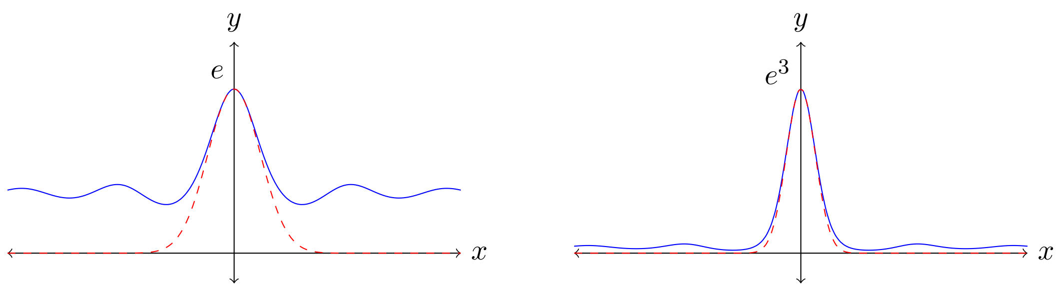


Figure 2.9: $y = e^{n\phi(x)}$ for $n = 1$ and $n = 3$.

2.4 Proof that almost all cubic graphs are Hamiltonian

We are now equipped with the technical machinery to prove that almost all cubic graphs are Hamiltonian. Let H_n be the number of Hamilton cycles in a random graph G from $\mathcal{G}_{n,3}$ and let $Y_n = Y$ be the number of Hamilton cycles in a random partition P from $\mathcal{P}_{n,3}$. We emphasise that these are distinct random variables, but we recover H_n from Y by conditioning on no loops and no multiple edges. Note that n must be even and asymptotics are over $2\mathbb{Z}^+$.

We first find an asymptotic expression for $\mathbb{E}Y$.

Lemma 2.28. *Let Y be the number of Hamilton cycles in a random partition P in $\mathcal{P}_{n,3}$. Then*

$$\mathbb{E}Y \sim \sqrt{\frac{\pi}{2n}} \left(\frac{2}{\sqrt{3}}\right)^n.$$

Proof. By definition,

$$\mathbb{E}Y = \frac{1}{|\Omega_{n,3}|} \sum_{P \in \Omega_{n,3}} Y(P) = \frac{1}{|\Omega_{n,3}|} \sum_{P_H} |\{P \in \Omega_{n,3} : P_H \subseteq P\}|,$$

where P_H sums over all subpartitions such that $H = G(P_H)$ is a Hamilton cycle. Thus, we want to count the number of subpartitions which projects to a Hamilton cycle. First, the number of Hamilton cycles on n vertices is

$$\frac{(n-1)!}{2},$$

because there are $n!$ ways to arrange n vertices in a line, and we divide by $2n$ to adjust for starting vertex and direction. Each vertex corresponds to a cell in the configuration model, and we use 2 out of 3 points in each cell for the Hamilton cycle. So there are $(3 \times 2)^n = 6^n$ subpartitions that project to a given Hamilton cycle. Figure 2.10 shows one such subpartition P_H for $n = 6$, given Hamilton cycle $H = v_1v_2v_3v_4v_5v_6v_1$.

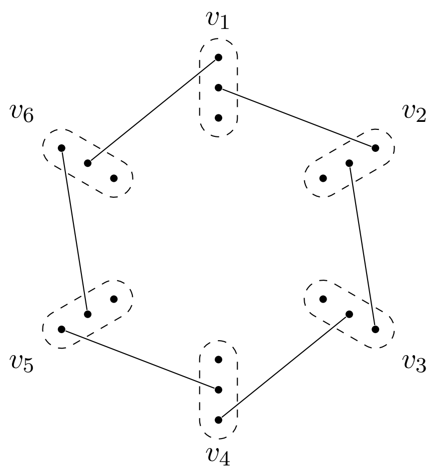


Figure 2.10: For a given Hamilton cycle, there are 3 choices for an “in” points and 2 choices for an “out” point.

Thus the number of subpartitions P_H that project to a Hamilton cycle is

$$\frac{(n-1)!\,6^n}{2}. \tag{2.5}$$

Since $2n$ points have been used, there are n points left, so we partition them arbitrarily in $f(n)$ ways, where f is defined in (2.1). From (2.2), we have $|\Omega_{n,3}| = f(3n)$, so

$$\mathbb{E}Y = \frac{(n-1)!\,6^n}{2}\frac{f(n)}{f(3n)} = \frac{12^n}{2}\frac{(n-1)!\,n!\,(3n/2)!}{(n/2)!\,(3n)!}. \tag{2.6}$$

Applying Stirling's approximation in the form of Corollary 2.26 to (2.6), we have

$$\mathbb{E}Y \sim \frac{12^n}{2}\frac{\sqrt{2\pi}n^{-1/2}\left(\frac{n}{e}\right)^n\sqrt{2\pi}n^{1/2}\left(\frac{n}{e}\right)^n\sqrt{2\pi}(3n/2)^{1/2}\left(\frac{3n}{2e}\right)^{3n/2}}{\sqrt{2\pi}\,(n/2)^{1/2}\left(\frac{n}{2e}\right)^{n/2}\sqrt{2\pi}(3n)^{1/2}\left(\frac{3n}{e}\right)^{3n}} = \sqrt{\frac{\pi}{2n}}\left(\frac{2}{\sqrt{3}}\right)^n,$$

as required. □

We now prove that almost all cubic graphs are Hamiltonian. The following theorem is a restatement of Theorem 2.20 for $r = 3$.

Theorem 2.29 (Robinson and Wormald [25]). *Let H_n be the number of Hamilton cycles in G from $\mathcal{G}_{n,3}$. Then* $\Pr(H_n > 0) \to 1$ *as $n \to \infty$.*

Proof. Let X_j be the number of j-cycles in a random partition P from $\mathcal{P}_{n,3}$. We apply the small subgraph conditioning method by verifying the conditions of Theorem 2.19.

Verifying (A1). Theorem 2.23 verifies (A1).

Verifying (A2′). To verify (A2′), we will show that, for fixed $3 \le j < n$,

$$\frac{\mathbb{E}(YX_j)}{\mathbb{E}Y} \to \lambda_j(1+\zeta_j) = \frac{2^j + (-1)^j - 1}{2j}$$

where

$$\lambda_j = \frac{2^j}{2j} \quad \text{and} \quad \zeta_j = \frac{(-1)^j - 1}{2^j}.$$

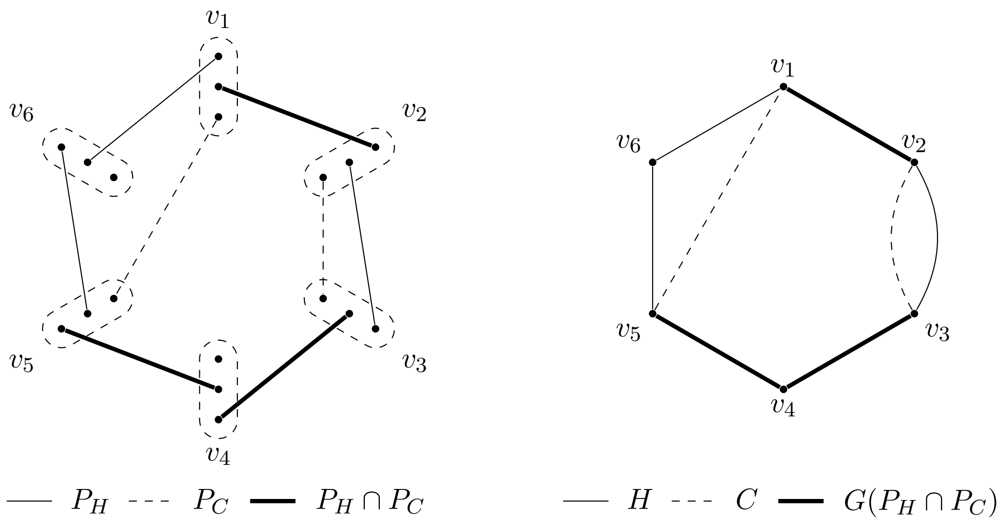
The cases $j = 1$ and $j = 2$ follow similarly.

We may think of the random variable YX_j as counting ordered pairs of subpartitions (P_H, P_C) such that $G(P_H)$ is a Hamilton cycle and $G(P_C)$ is a j-cycle. We can write

$$\begin{aligned}|\Omega_{n,3}|\,\mathbb{E}(YX_j) = \sum_{P\in\Omega_{n,3}} Y(P)X_j(P) &= \sum_{P\in\Omega_{n,3}}\sum_{(P_H,P_C)} \mathbb{1}(P_H \cup P_C \subseteq P)\\ &= \sum_{(P_H,P_C)} |\{P \in \Omega_{n,3} : P_H \cup P_C \subseteq P\}|\end{aligned}$$

where the sum is over all pairs of subpartitions (P_H, P_C) such that $H = G(P_H)$ is a Hamilton cycle and $C = G(P_C)$ is a j-cycle.

To perform this count, we first condition on $P_H \cap P_C$. Given a Hamilton cycle H and a j-cycle C, consider the intersection $G(P_H \cap P_C)$. Each vertex has two edges on the Hamilton cycle, and because $G(P)$ is cubic, each vertex has only one edge not in H. Thus the intersection $G(P_H \cap P_C)$ looks like k vertex-disjoint paths of length at least 1, for some $k \geq 1$. We cannot have length-0 paths or $k = 0$, otherwise a vertex will have more than one edge not in H. The sum of the lengths of the paths must be $j - k$ because there must be k edges not on H. Figure 2.11 shows the intersection between P_C and P_H, where C is a 5-cycle.



(a) $P_H \cap P_C$. (b) $G(P_H \cap P_C)$

Figure 2.11: Here, we have $k = 2$ vertex-disjoint paths. The sum of the lengths of the paths is $j - k = 5 - 2 = 3$.

We evaluate the sum using the following process.

1. Choose a subpartition P_H which projects to a Hamilton cycle H.
2. Choose a sequence of lengths $(\ell_1, \ldots, \ell_k)$, where $\ell_1 + \cdots + \ell_k = j - k$ and $\ell_1, \ldots, \ell_k \geq 1$, recorded with respect to an arbitrary start vertex and direction.
3. Choose where to place the k paths on H.
4. Choose C consistent with the k paths.
5. Choose P_C which projects to C.
6. Partition the remaining points.

Then $|\Omega_{n,3}|\, \mathbb{E}(YX_j)$ is equal to the number of ways to complete the above process, summed over all $k \geq 1$.

By combinatorial arguments, we have:

1. The number of ways to complete Step 1 is $\frac{(n-1)!\,6^n}{2}$ by (2.5).
2. The number of ways to complete Step 2 is $\binom{j-k-1}{k-1}$ by elementary counting methods.
3. Given a sequence of lengths, we have n choices for a starting vertex and 2 choices for a direction. With respect to this starting vertex and direction, the location of the paths in $P_C \cap P_H$ is determined by the number of vertices in between each path. Since the cycle uses up j vertices, we count the number of ways of distributing $n-j$ vertices in k gaps, which is equal to $\binom{n-j+k-1}{k-1}$. However, each particular arrangement is repeated $2k$ times, as shifts of the sequence, such as $(\ell_2, \dots, \ell_k, \ell_1)$, or reflections of the sequence, such as $(\ell_k, \dots, \ell_1)$, give rise to the same set of disjoint paths. Hence, the number of ways of completing Step 3 is

$$2n \times \binom{n-j+k-1}{k-1} \times \frac{1}{2k} = \frac{n}{k}\binom{n-j+k-1}{k-1} \sim \frac{n^k}{k!}. \tag{2.7}$$

4. There are $\frac{(k-1)!}{2}$ ways to cyclically order the k disjoint paths. Since each path has 2 ends, there are 2^k ways to decide the orientation of each path within the cycle. Hence, the number of ways of completing Step 4 is

$$\frac{(k-1)!\,2^k}{2} = (k-1)!\,2^{k-1}. \tag{2.8}$$

5. The number of ways to complete Step 5 is 1, because each cell only has one unused point remaining. Note that we use $2k$ points in this step.
6. There are $3n - 2n - 2k = n - 2k$ points left, so the number of ways to complete Step 6 is $f(n-2k)$.

Putting this together with (2.2), we have

$$f(3n)\,\mathbb{E}(YX_j) \sim \frac{(n-1)!\,6^n}{2}\left(\sum_{k\geq 1} \frac{2^{k-1}\,n^k}{k}\binom{j-k-1}{k-1} f(n-2k)\right).$$

Using Corollary 2.26 we can see that for fixed k, we have

$$\frac{f(n-2k)}{f(n)} = \frac{(n-2k)!\,(\frac{n}{2})!\,2^{n/2}}{\left(\frac{n}{2}-k\right)!\,2^{n/2-k}\,n!} \sim 2^k \frac{\sqrt{2\pi}n^{-2k+1/2}\left(\frac{n}{e}\right)^n}{\sqrt{2\pi}\left(\frac{n}{2}\right)^{-k+1/2}\left(\frac{n}{2e}\right)^{n/2}} \frac{\sqrt{2\pi}\left(\frac{n}{2}\right)^{1/2}\left(\frac{n}{2e}\right)^{n/2}}{\sqrt{2\pi}n^{1/2}\left(\frac{n}{e}\right)^n} = n^{-k},$$

so by dividing by the expression (2.6), we have

$$\frac{\mathbb{E}(YX_j)}{\mathbb{E}Y} \sim \sum_{k\geq 1} \frac{2^k}{2k}\binom{j-k-1}{k-1} = -\frac{1}{2j} + \sum_{k=0}^{\lfloor j/2\rfloor} \frac{2^{k-1}}{j-k}\binom{j-k}{k}.$$

We simplify this sum with the use of a generating function. We have

$$\binom{j-k}{k} = [z^k]\,(1+z)^{j-k} = [z^j]\,(z(1+z))^{j-k},$$

with square brackets denoting coefficient extraction. Thus we have

$$\frac{\mathbb{E}(YX_j)}{\mathbb{E}Y} \sim -\frac{1}{2j} + [z^j]\sum_{k=0}^{\lfloor j/2\rfloor} \frac{2^{j-1}}{j-k}\left(\frac{z(1+z)}{2}\right)^{j-k}.$$

Changing indices with $i = j-k$, we have

$$\begin{aligned}\frac{\mathbb{E}(YX_j)}{\mathbb{E}Y} &\sim -\frac{1}{2j} + 2^{j-1}[z^j]\sum_{i=0}^{\infty}\frac{1}{i}\left(\frac{z(1+z)}{2}\right)^{i} \\ &= -\frac{1}{2j} + 2^{j-1}[z^j]\left(-\log\left(1-\frac{z(1+z)}{2}\right)\right),\end{aligned}$$

making use of the Taylor expansion

$$-\log(1-z) = \sum_{i=1}^{\infty}\frac{z^i}{i}.$$

Further simplifying, we have

$$\begin{aligned}\frac{\mathbb{E}(YX_j)}{\mathbb{E}Y} &\sim -\frac{1}{2j} + 2^{j-1}[z^j]\left(-\log(1-z) - \log\left(1+\frac{z}{2}\right)\right) \\ &= -\frac{1}{2j} + 2^{j-1}\left(\frac{1}{j} + \frac{(-1)^j}{j\,2^j}\right) \\ &= \frac{2^j + (-1)^j - 1}{2j}.\end{aligned}$$

Verifying (A3). Condition (A1) gave us $\lambda_j = \frac{2^j}{2j}$ and condition (A2$'$) gave us $\zeta_j = \frac{(-1)^j - 1}{2^j}$. So

$$\lambda_j \zeta_j^2 = \frac{1-(-1)^j}{j\,2^j} = \begin{cases} 0 & \text{for even } j, \\ \dfrac{2}{j\,2^j} & \text{for odd } j. \end{cases}$$

Using the Taylor expansion

$$\log\left(\frac{1+z}{1-z}\right) = 2\sum_{j \text{ odd}} \frac{z^j}{j},$$

with $z = 1/2$, we see that

$$\sum_{j\geq 1} \lambda_j \zeta_j^2 = \log 3.$$

Verifying (A4). We may think of the random variable Y^2 as counting ordered pairs of subpartitions, where both entries project to a Hamilton cycle. Thus, we can write

$$\mathbb{E}Y^2 = \frac{1}{|\Omega_{n,3}|} \sum_{P \in \Omega_{n,3}} Y(P)^2 = \frac{1}{|\Omega_{n,3}|} \sum_{(P_{H_1}, P_{H_2})} |\{P \in \Omega_{n,3} : P_{H_1} \cup P_{H_2} \subseteq P\}|,$$

where (P_{H_1}, P_{H_2}) sums over pairs of subpartitions such that $H_1 = G(H_1)$ and $H_2 = G(H_2)$ are Hamilton cycles.

We perform this count by conditioning on $P_{H_1} \cap P_{H_2}$. Fix P_{H_1}. If $P_{H_1} = P_{H_2}$, then $P_{H_1} \cap P_{H_2} = P_{H_1}$. In this case, there are $f(n)$ ways to partition the remaining points. Otherwise, as in the short cycle calculations verifying (A2), $P_{H_1} \cap P_{H_2}$ will look like k vertex-disjoint paths, for some $k \geq 1$, with the sum of the lengths of the paths $n-k$. Because the random variable Y is the same as X_n, we can evaluate the sum by counting the number of ways to complete the same process as in the verification of (A2), summed over all $k \geq 1$, with (H, C) replaced with (H_1, H_2). However, the calculations will vary slightly in Step 3 as k is no longer fixed.

By combinatorial arguments, we have:

1. The number of ways to complete Step 1 is $\frac{(n-1)!\,6^n}{2}$ by (2.5).
2. The number of ways to complete Step 2 is $\binom{n-k-1}{k-1}$ by elementary counting methods.
3. Given a sequence of lengths, we have n choices for a starting vertex, and 2 choices for a direction. Because there is just a single edge between successive paths, this is enough to specify an arrangement of k paths on H_1. However, each particular arrangement is repeated $2k$ times, so the number of ways to complete Step 3 is $2n/2k = n/k$. (Alternatively, let $j = n$ in (2.7)).

4. We use the same argument as in the verification of (A2′). By (2.8), the number of ways to complete Step 4 is $(k-1)!\,2^{k-1}$.
5. The number of ways to complete Step 5 is 1, because each cell only has one unused point remaining. Note that we use $2k$ points in this step.
6. There are $3n-2n-2k=n-2k$ points left, so the number of ways to complete Step 6 is $f(n-2k)$.

Putting this together, we have

$$f(3n)\,\mathbb{E}Y^2 = \frac{(n-1)!\,6^n}{2}\left(f(n)+\sum_{k=1}^{n/2}\binom{n-k-1}{k-1}\frac{n(k-1)!\,2^{k-1}}{k}f(n-2k)\right).$$

By (2.6), we have

$$\begin{aligned}\frac{\mathbb{E}Y^2}{\mathbb{E}Y} &= 1+\sum_{k=1}^{n/2}\binom{n-k-1}{k-1}\frac{n(k-1)!\,2^{k-1}}{k}\frac{f(n-2k)}{f(n)}\\ &= 1+\sum_{k=1}^{n/2}\frac{(n-k-1)!}{(k-1)!(n-2k)!}\frac{n(k-1)!\,2^{k-1}}{k}\frac{(n-2k)!}{(\frac{n}{2}-k)!2^{n/2-k}}\frac{(n/2)!2^{n/2}}{n!}\\ &= \sum_{k=0}^{n/2}a_n(k),\end{aligned}$$

where

$$a_n(k)=\begin{cases}1 & \text{for } k=0\\ \dfrac{n\,2^{2k-1}}{k}\dfrac{(n-k-1)!\,(n/2)!}{(n/2-k)!\,n!} & \text{for } 1\le k\le \dfrac{n}{2}.\end{cases}$$

We now wish to apply Laplace's method to compute the asymptotic summation of this expression.

Define

$$b_n=\frac{1}{2n} \tag{2.9}$$

$$\psi(\alpha)=\frac{1}{\alpha\sqrt{(1-\alpha)(1-2\alpha)}}, \tag{2.10}$$

$$\phi(\alpha)=\alpha\log 2+g(1-\alpha)-\frac{1}{2}g(1-2\alpha), \tag{2.11}$$

where $g(x) = x \log x$ for $x > 0$ and $g(0) = 0$. Intuitively, we let $\alpha = k/n$ in $a_n(k)$ and show that

$$\sum_{k=0}^{n/2} a_n(k) \sim \sum_{\alpha \in \{0, \frac{1}{n}, \dots, \frac{n/2}{n}\}} b_n \psi(\alpha)\, e^{n\phi(\alpha)}.$$

This sum can be thought of as a Riemann approximation of an integral, which in turn, is approximated by a Gaussian around the global maximum of ϕ. We will apply Lemma 2.27 to demonstrate its usage (although we remark its usage is very much overkill in the one-dimensional case).

We see that

$$\phi'(\alpha) = \log(2 - 4\alpha) - \log(1 - \alpha),$$
$$\phi''(\alpha) = \frac{1}{1-\alpha} - \frac{2}{1-2\alpha}.$$

The first six conditions of Lemma 2.27 hold:

(i) Let $\mathcal{L} = \mathbb{Z}$, a lattice with rank 1 and determinant 1.

(ii) Let $K = \left[0, \frac{1}{2}\right]$. It is a compact convex set with non-empty interior. Note that $\mathcal{L} \cap nK = \{0, 1, \dots, n/2\}$.

(iii) The function $\phi : K \to \mathbb{R}$ defined in (2.11) is continuous. The only solution to $\phi'(\alpha) = 0$ is $\alpha_0 = \frac{1}{3}$, and $\phi''(\alpha_0) < 0$, so there is a local maximum at $\alpha_0 = \frac{1}{3}$, with $\phi\left(\frac{1}{3}\right) = \log 2 - \frac{1}{2}\log 3$. Checking the boundary, we have $\phi(0) = 0$ and $\phi\left(\frac{1}{2}\right) = 0$. Thus, we conclude that α_0 gives the unique global maximum in K.

(iv) The function ϕ is twice differentiable in the interior of K, with $\phi''(\alpha_0) = -9/2 < 0$.

(v) Let $K_1 = \left(\frac{1}{4}, \frac{5}{12}\right) \subset K$ (any sufficiently small open subset of K containing α_0 will do). Then K_1 is a neighbourhood of α_0, and the function $\psi : K_1 \to \mathbb{R}$, defined in (2.10) is a continuous function on K_1 with $\psi(\alpha_0) = 9/\sqrt{2} > 0$.

(vi) Let $\ell_n = 0$.

It remains to prove that condition (vii) holds. Introduce the scaled variable $\alpha = k/n$.

Observe that (vii) holds when $k = 0$, since we have $\phi(0) = 0$, and hence

$$a_n(0) = 1 = O\left(\frac{1}{2n} e^n\right).$$

By Corollary 2.26, for $k \geq 1$, we have

$$\begin{aligned}a_n(k) &\sim \frac{n2^{2\alpha n-1}}{\alpha n} \frac{\sqrt{2\pi}((1-\alpha)n)^{-1/2}\left(\frac{(1-\alpha)n}{e}\right)^{(1-\alpha)n}(n/2)^{1/2}\left(\frac{n}{2e}\right)^{n/2}}{\sqrt{2\pi}((\frac{1}{2}-\alpha)n)^{1/2}\left(\frac{(1-2\alpha)n}{2e}\right)^{(\frac{1}{2}-\alpha)n}\sqrt{2\pi}n^{1/2}\left(\frac{n}{e}\right)^{n}} \\ &= b_n\,\psi(\alpha)\,e^{n\phi(\alpha)}.\end{aligned}$$

Because $\psi(k/n) = e^{o(n)}$, we have

$$a_n(k) = O(b_n e^{n\phi(k/n)+o(n)}),$$

and thus (2.3) holds.

To show (2.4) holds, applying Lemma 2.25 gives us

$$a_n(k) = b_n\,\psi(k/n)\,e^{n\phi(k/n)}\,e^{O(1/n)}.$$

Note that the deliberate choice of K_1 ensures that $\psi(k/n) = O(1)$, or in other words, $\psi(k/n)$ is bounded above within K_1. Therefore, we have

$$\lim_{n\to\infty} \psi(k/n)(e^{O(1/n)} - 1) = 0,$$

which implies that $\psi(k/n)e^{O(1/n)} = \psi(k/n) + o(1)$. Hence,

$$a_n(k) = b_n\,(\psi(k/n) + o(1))\,e^{n\phi(k/n)}.$$

Thus, all the conditions of Lemma 2.27 are met, so

$$\frac{\mathbb{E}Y^2}{\mathbb{E}Y} = \sum_{\mathcal{L}\cap nK} \sim \frac{\sqrt{2\pi}\,\psi(\alpha_0)}{\det(\mathcal{L})\sqrt{-\phi''(\alpha_0)}} b_n\,n^{1/2}\,e^{n\phi(\alpha_0)} = 3\sqrt{\frac{\pi}{2n}}\left(\frac{2}{\sqrt{3}}\right)^n.$$

Hence dividing by the expression in Lemma 2.28, we have

$$\frac{\mathbb{E}Y^2}{(\mathbb{E}Y)^2} \sim 3.$$

Therefore (A1)–(A4) are verified, and we have

$$\frac{Y}{\mathbb{E}Y} \xrightarrow{d} \prod_{j=1}^{\infty}(1+\zeta_j)^{Z_j} e^{-\lambda_j \zeta_j} = \prod_{\substack{j \geq 1 \\ j \text{ odd}}} \left(1 - \frac{2}{2^j}\right)^{Z_j} e^{1/j},$$

where $Z_j \sim \text{Po}\left(\frac{2^j}{2j}\right)$.

Finally, conditioning on no loops and multiple edges ($X_1 = X_2 = 0$) recovers H_n and we have

$$\frac{H_n}{\mathbb{E}Y} \xrightarrow{d} e \prod_{\substack{j \geq 3 \\ j \text{ odd}}} \left(1 - \frac{2}{2^j}\right)^{Z_j} e^{1/j}.$$

We can use (A2) to see that $\mathbb{E}H_n = \mathbb{E}(Y|X_1 = X_2 = 0) \sim e\,(\mathbb{E}Y)$, so we can obtain the normalised asymptotic distribution of H_n:

$$\frac{H_n}{\mathbb{E}H_n} \xrightarrow{d} \prod_{\substack{j \geq 3 \\ j \text{ odd}}} \left(1 - \frac{2}{2^j}\right)^{Z_j} e^{1/j}.$$

From this, we see that $\Pr(H_n > 0) \to 1$ as the distribution on the right hand side is positive with probability 1. □

2.5 Random regular uniform hypergraphs

Hypergraphs are a generalisation of graphs where an edge does not need to contain exactly two vertices. We generalise the configuration model to a s-uniform hypergraphs and discuss the application of the small subgraph conditioning method. For more background on hypergraphs, see [14].

2.5.1 Definitions

Definition 2.30. A *hypergraph* $G = (V, E)$ consists of a set of vertices V and a multiset E of non-empty multisubsets of V, which we call *edges*.

Definition 2.31. We say that G is *simple* if E is a set of sets: that is, there are no repeated edges and no edge contains a repeated vertex.

Definition 2.32. If we have hypergraphs $G = (V, E)$ and $G' = (V', E')$ with $V' \subseteq V$ and $E' \subseteq E$, then G' is a *subhypergraph* of G. In addition, if $V' = V$, thenG' is a *spanning subhypergraph* of G.

Example 2.33. Figure 2.12 represents the simple hypergraph $G = (V, E)$, where $V = [7]$ and $E = \{123,\ 23,\ 356,\ 4\}$.

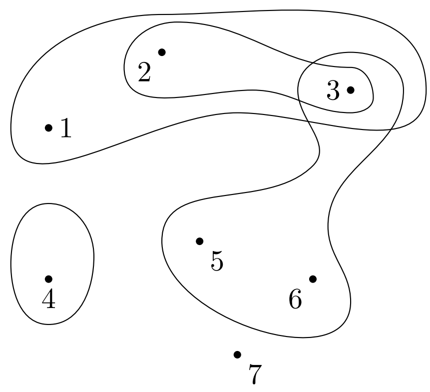


Figure 2.12: A simple hypergraph on 7 vertices.

As with graphs, we have the concepts of a path and a cycle.

Definition 2.34. A (Berge) *path of length k* is a sequence $v_1, e_1, v_2, e_2, \ldots, e_k, v_{k+1}$ where the e_i are distinct edges and the v_i are distinct vertices such that $v_i, v_{i+1} \in e_i$. For brevity, we will sometimes refer to a path by a sequence of edges $e_1 e_2 \ldots e_k$.

Definition 2.35. A hypergraph is *connected* if there is a path between every pair of vertices.

Definition 2.36. A 1-*cycle* (or *loop*) is a hypergraph with one edge which contains a repeated vertex. A 2-*cycle* is a hypergraph with two edges which intersect in at least 2 vertices. For $j \geq 3$, a j-*cycle* is a hypergraph with j edges which can be labelled $e_1, e_2, \ldots, e_j$ such that there exists distinct vertices $v_1, \ldots, v_j$ where $v_i \in e_i \cap e_{i+1}$ for $i = 1, \ldots, j$ (where $e_{j+1} \equiv e_1$).

The definition of a tree will be given in Chapter 3.

As hypergraphs can become very complicated, we will restrict ourselves to certain types of hypergraphs.

Definition 2.37. The *degree* of a vertex v is the number of edges which contain it, denoted $\deg(v)$.

Definition 2.38. For an integer $r \geq 1$, a hypergraph is said to be r-*regular* if every vertex has degree r.

Definition 2.39. For an integer $s \geq 2$, a hypergraph is said to be *s-uniform* if every edge contains s vertices.

Example 2.40. A graph is a 2-uniform simple hypergraph.

Example 2.41. Figure 2.13 shows a 2-regular 3-uniform hypergraph with $V = [6]$ and $E = \{123, 345, 156, 246\}$.

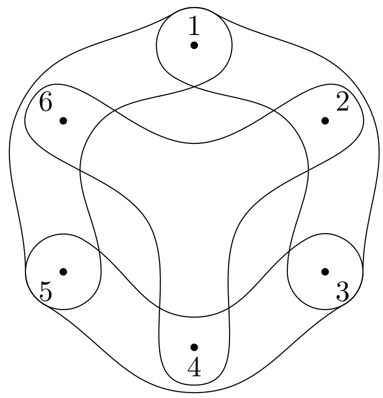


Figure 2.13: A 2-regular 3-uniform hypergraph.

We can define a special type of cycle for s-uniform hypergraphs.

Definition 2.42. An s-uniform 1-cycle is *loose* if the edge contains $s-1$ distinct vertices, and an s-uniform 2-cycle is *loose* if the intersection of the two edges has size 2. For $j \geq 3$, an s-uniform j-cycle is *loose* if

$$|e_k \cap e_\ell| = \begin{cases} 1 & \text{if } k - \ell \equiv \pm 1 \pmod{j}, \\ 0 & \text{otherwise} \end{cases}$$

for $k \neq \ell$. A loose j-cycle C contains $(s-1)j$ vertices.

Definition 2.43. Let v be a vertex in a loose j-cycle C. We say v is *C-external* if v has degree 2 in C. Otherwise, we say v has degree 1 and we say it is *C-internal*.

Example 2.44. Let C_1 be the 5-uniform loose 3-cycle on the left of Figure 2.14. Then u is C_1-external and v is C_1-internal. The cycle C_2 on the right is a 5-uniform 4-cycle which is not loose.

For s-uniform hypergraphs, we also have a version of the Handshaking Lemma.

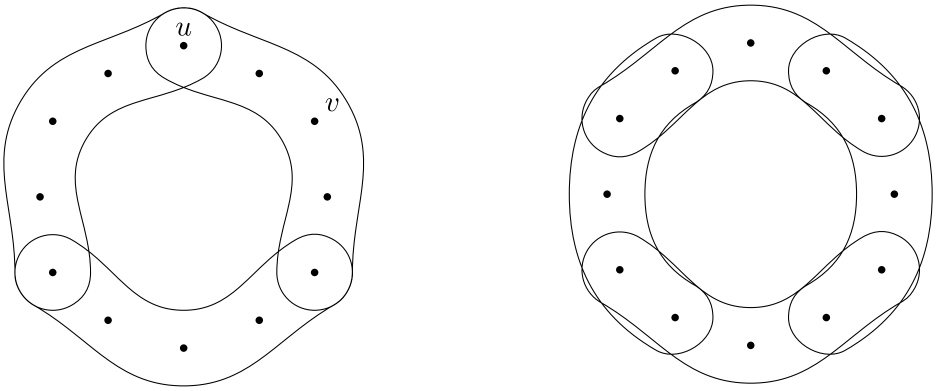


Figure 2.14: Some cycles.

Lemma 2.45 (Handshaking Lemma for s-uniform hypergraphs)**.** *For an s-uniform hypergraph $G = (V, E)$, we have*

$$\sum_{v \in V} \deg(v) = s\,|E|\,.$$

Proof. We have

$$\sum_{v \in V} \deg(v) = \sum_{v \in V} \sum_{e \in E : v \in e} 1 = \sum_{e \in E} \sum_{v \in V : v \in e} 1 = \sum_{e \in E} s = s\,|E|\,,$$

making use of a double counting argument. □

2.5.2 Hypergraph configuration model

We can now introduce a generalised configuration model to work with r-regular s-uniform hypergraphs. Let $\mathcal{S}_{n,r,s}$ be the set of r-regular s-uniform simple hypergraphs on $[n]$. Let $\mathcal{G}_{n,r,s}$ be the uniform probability space over $\mathcal{S}_{n,r,s}$. By the Handshaking Lemma for s-uniform hypergraphs, we need $s \mid rn$.

The configuration model for hypergraphs involves considering a set of rn *points* partitioned in n *cells* $v_1, \ldots, v_n$, each with r points. We partition the points into $\frac{rn}{s}$ subsets of size s each, called a *partition*. Each partition corresponds to a hypergraph $G(P)$ in which the cells are regarded as vertices and the partitions as edges. Each simple hypergraph corresponds to precisely $(r!)^n$ partitions, so an r-regular s-uniform simple hypergraph can be chosen uniformly at random by choosing a partition uniformly at random and rejecting the result if it has loops or multiple edges.

Denote the set of possible partitions by $\Omega_{n,r,s}$ and the uniform probability space over $\Omega_{n,r,s}$ by $\mathcal{P}_{n,r,s}$. Define a *subpartition* P' to be a subset of a partition $P \in \Omega_{n,r,s}$, which will project to a subhypergraph of $G(P)$.

Similarly to the graph case, the configuration model allows us to prove some properties of $\mathcal{G}_{n,r,s}$ by performing computations in $\mathcal{P}_{n,r,s}$. The number of s-subset partitions of the set of t points, where $s \mid t$, is

$$f_s(t) = \frac{t!}{(t/s)!(s!)^{t/s}},$$

so the number of partitions in $\Omega_{n,r,s}$ is

$$|\Omega_{n,r,s}| = \frac{(rn)!}{(rn/s)!\,(s!)^{rn/s}}. \tag{2.12}$$

Therefore, the number of r-regular s-uniform simple hypergraphs on n vertices is precisely

$$|\mathcal{S}_{n,r,s}| = \frac{(rn)!\ \Pr(\text{Simple})}{(rn/s)!\,(s!)^{rn/s}(r!)^n},$$

where "Simple" is the event that the corresponding hypergraph to a partition has no loops or multiple edges. Thus, an asymptotic formula for $|\mathcal{S}_{n,r,s}|$ can be found by estimating $\Pr(\text{Simple})$. Cooper, Frieze, Molloy and Reed showed in [11] that for fixed integers $r \geq 1$ and $s \geq 3$, we have

$$\Pr(\text{Simple}) \sim \exp\left(-\frac{(r-1)(s-1)}{2}\right). \tag{2.13}$$

2.5.3 Small subgraph conditioning with hypergraphs

There have been a few results which have applied the small subgraph conditioning method to the hypergraph configuration model. Cooper et al. [11] proved the existence of a perfect matching in a random r-regular s-uniform hypergraph, for fixed integers $r \geq 2$ and $s \geq 3$. Specifically, in [11, Theorem 1], they showed that the probability of a random r-regular s-uniform hypergraph contains a perfect matching tends to 0 if $s > \sigma_r$ and tends to 1 if $s < \sigma_r$, where

$$\sigma_r = \frac{\log r}{(r-1)\log\left(\frac{r}{r-1}\right)} + 1.$$

Altman, Greenhill, Isaev and Ramadurai [3] investigated loose Hamilton cycles in r-regular s-uniform hypergraphs and established a threshold result: that is, for each $s \geq 3$, there exists a real number $\rho(s)$ such that the probability of a random r-regular s-uniform hypergraph containing a loose Hamilton cycle tends to 0 if $r \leq \rho(s)$ and tends to 1 if $r > \rho(s)$, where

$$\frac{e^{s-1}}{s-1} - \frac{s-2}{2} - \frac{(s^2-s+1)^2}{(s-1)e^{s-1}} < \rho(s) < \frac{e^{s-1}}{s-1} - \frac{s-2}{2}.$$

When using the small subgraph conditioning method with the hypergraph configuration model, it is convenient to let X_1 be the number of 1-cycles in a random partition P from $\mathcal{P}_{n,r,s}$ and for $j \geq 2$, let X_j be the number of *loose* j-cycles in a random partition P from $\mathcal{P}_{n,r,s}$. Cooper et al. proved in [11] that $X_j \to Z_j$ as $n \to \infty$ where Z_j are asymptotically independent Poisson random variables with mean

$$\lambda_j = \frac{(r-1)^j(s-1)^j}{2j}. \tag{2.14}$$

To be more precise, Cooper et al. worked with the random variable X'_j, the number of j-cycles (not necessarily loose), and showed that X'_j has the same asymptotic distribution as X_j, as the contribution to X'_j from non-loose j-cycles forms a negligible fraction of X'_j.

However, there are some shortcomings when applying the small subgraph conditioning method to regular uniform hypergraphs. Recall that, in our proof of Theorem 2.29, we were able to condition on the event "Simple", which is the same as $X_1 = X_2 = 0$, to recover the asymptotic distribution of H_n (the number of Hamilton cycles in a random graph in $\mathcal{G}_{n,r}$) from Y (the number of Hamilton cycles in a random partition in $\mathcal{P}_{n,r}$). We cannot do this in the hypergraph case, because the event $X_1 = X_2 = 0$ is not the same as "Simple". Instead, "Simple" is the event where a random partition has no 1-cycles and no repeated edges.

Fortunately, we can translate some asymptotic properties from the configuration model. For any event $\mathcal{E} \subseteq \Omega_{n,r,s}$, we have

$$\Pr(P \in \mathcal{E} \mid \text{Simple}) \leq \frac{\Pr(P \in \mathcal{E})}{\Pr(\text{Simple})}. \tag{2.15}$$

Altman et al. proved the following lemma in [3]:

Lemma 2.46 ([3, Lemma 2.1])**.** *Fix integers $r, s \geq 2$. For any positive integer n such that $s \mid rn$, let $\widehat{P}$ be a uniformly random partition in $\Omega_{n,r,s}$ with no 1-cycles, let $P_{\mathcal{S}}$ be a uniformly random simple partition in $\Omega_{n,r,s}$ and let P be a uniformly random partition in $\Omega_{n,r,s}$. Let $Y : \Omega_{n,r,s} \to \mathbb{Z}$ be a random variable. Then as $n \to \infty$ along integers such that $s \mid rn$, the following two properties hold.*

(a) *If* $\Pr(Y(P) \in A) = o(1)$, *then* $\Pr(Y(P_{\mathcal{S}}) \in A) = o(1)$ *for any* $A \subseteq \mathbb{Z}$.

(b) $\Pr(Y(\widehat{P}) \in A) - \Pr(Y(P_{\mathcal{S}}) \in A) = o(1)$ *for any* $A \subseteq \mathbb{Z}$.

Property (a) follows from (2.15), while property (b) follows from (2.13) and the fact that the probability that two parts in a random partition P from $\mathcal{P}_{n,r,s}$ give rise to a repeated edge is $o(1)$ (as remarked by Cooper et al. in [11]).

Property (b) essentially tells us that the distribution of Y that arises from conditioning on $X_1 = 0$ is asymptotically equivalent to the distribution Y conditioned on "Simple". This allows us to apply the following corollary from [3]:

Corollary 2.47. *Suppose that Y_n and $X_{j,n}$ satisfy conditions* (A1)–(A4) *of Theorem 2.19. Let $\widehat{Y}_n$ be the random variable obtained from Y_n by conditioning on the event $X_{1,n} = 0$. Then*

$$\frac{\widehat{Y}_n}{\mathbb{E}Y_n} \xrightarrow{d} e^{\lambda_1 \zeta_1} \prod_{j=2}^{\infty} (1+\zeta_j)^{Z_j} e^{\lambda_j \zeta_j} \quad \textit{as} \quad n \to \infty.$$

Moreover, if $\zeta_j > -1$ for all $j \geq 1$, then asymptotically almost surely $\widehat{Y}_n > 0$.

However, convergence in distribution does not guarantee that the expected values converge. So the small subgraph conditioning method for hypergraphs does not give an asymptotic expression for the expected value of a random variable in $\mathcal{G}_{n,r,s}$. Altman et al. were able to derive an asymptotic distribution for the number of loose Hamilton cycles in $\mathcal{G}_{n,r,s}$, but were only able to conjecture an expression for the expected value. Aldosari and Greenhill proved this conjecture using the switching method in [2].

CHAPTER 3

Trees

This one's from the Book!

– Paul Erdős

One of the aims of this thesis is to apply the small subgraph conditioning method to determine the distribution of spanning trees in a random regular uniform hypergraph. Before we discuss this, this chapter gives an overview of some existing results on trees and Prüfer sequences, and generalises these results for hypergraphs. These results will be useful when we make combinatorial arguments to compute the first moment, the joint factorial moment and the second moment in the verification of Theorem 2.19. Section 3.3 contains a generalisation of Prüfer sequences to hypergraphs which are not necessarily uniform. They are not directly used in the small subgraph conditioning method, but are interesting in their own right.

3.1 Trees and Prüfer sequences

We begin by giving an overview on some well-known results on trees in the graph case. Moon [21] is a good reference for this section.

Theorem 3.1 (Cayley's formula)**.** *The number of trees on $n \geq 2$ vertices is n^{n-2}.*

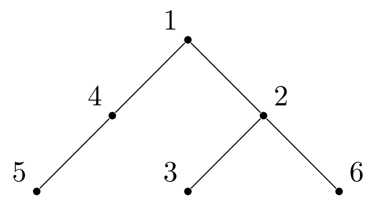
Many proofs of Cayley's formula are known, beautiful enough to be included in "the Book" (see [1]). One such proof uses Kirchoff's matrix tree theorem, which relates the number of trees with a determinant of a matrix. We will prove Theorem 3.1 through the use of *Prüfer sequences*.

Algorithm 3.2 (`Tree-to-Prüfer`)**.** For a tree $T \equiv T_0$ on $V = [n]$, start with the empty vector $\boldsymbol{c}$. Repeat the following process until one edge remains:

1. In T_i, identify the leaf ℓ which has the smallest labelled vertex.
2. Append $\boldsymbol{c}$ with the label of this leaf's neighbour.
3. Set $T_{i+1} := T_i - \ell$.

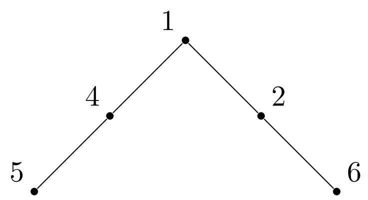
Then $\boldsymbol{c} \in [n]^{n-2}$, and we call it the *Prüfer sequence* (or *Prüfer code*) of T.

Example 3.3. Suppose we have the graph shown in Figure 2.3.



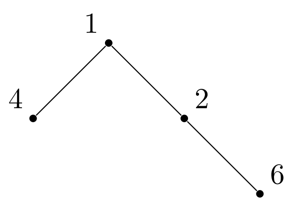
Running through the algorithm, we have:

- Delete 3 and record 2.



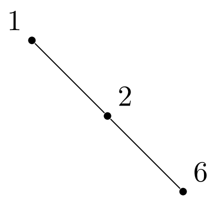
$$\boldsymbol{c} = (2, \cdot, \cdot, \cdot)$$

- Delete 5 and record 4.



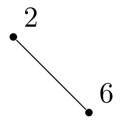
$$\boldsymbol{c} = (2, 4, \cdot, \cdot)$$

- Delete 4 and record 1.



$$\boldsymbol{c} = (2, 4, 1, \cdot)$$

- Delete 1 and record 2.



$\boldsymbol{c} = (2, 4, 1, 2)$

Conversely, we can recover the spanning tree from an element of $[n]^{n-2}$.

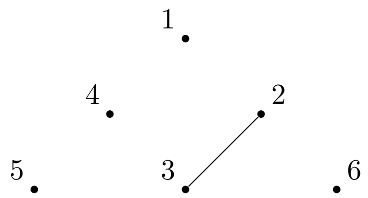
Algorithm 3.4 (`Prüfer-to-Tree`)**.** For $\boldsymbol{c} \in [n]^{n-2}$, start with $P = [n-1]$, vertex set $V = [n]$ and edge set $E = \varnothing$. Repeat the following process until $\boldsymbol{c}$ is empty:

1. Find the smallest label v in P which is not in $\boldsymbol{c}$.
2. Build an edge which joins v and the first vertex in $\boldsymbol{c}$.
3. Delete v from P and delete the first vertex from $\boldsymbol{c}$.

Finally, build the last edge between the last vertex in P and n. Then $T = (V, E)$ is the tree which corresponds to $\boldsymbol{c}$.

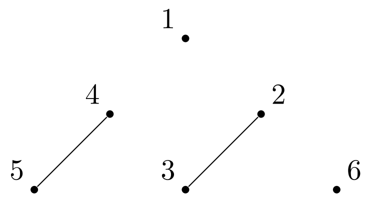
Example 3.5. Suppose $\boldsymbol{c} = (2, 4, 1, 2)$. Start with $P = \{1, 2, 3, 4, 5\}$.

- The smallest vertex in P but not in $\boldsymbol{c}$ is 3. Draw the edge 23.



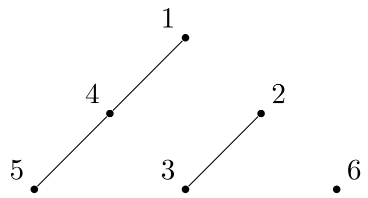
$P = \{1, 2, \cancel{3}, 4, 5\}, \quad \boldsymbol{c} = (\cancel{2}, 4, 1, 2)$

- The smallest vertex in P but not in $\boldsymbol{c}$ is 5. Draw the edge 45.



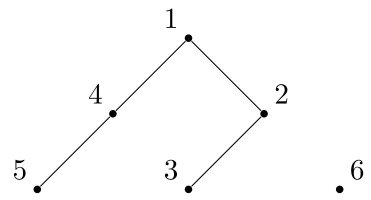
$P = \{1, 2, \cancel{3}, 4, \cancel{5}\}, \quad \boldsymbol{c} = (\cancel{2}, \cancel{4}, 1, 2)$

- The smallest vertex in P but not in $\boldsymbol{c}$ is 4. Draw the edge 14.



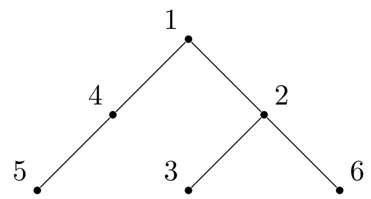
$P = \{1, 2, \cancel{3}, \cancel{4}, \cancel{5}\}, \quad \boldsymbol{c} = (\cancel{2}, \cancel{4}, \cancel{1}, 2)$

- The smallest vertex in P but not in $\boldsymbol{c}$ is 1. Draw the edge 12.



$$P = \{\not{1}, 2, \not{3}, \not{4}, \not{5}\}, \quad \boldsymbol{c} = (\not{2}, \not{4}, \not{1}, \not{2})$$

- Draw the edge 26. Note that we have recovered the tree from Example 3.3.



$$P = \{\not{1}, \not{2}, \not{3}, \not{4}, \not{5}\}, \quad \boldsymbol{c} = (\not{2}, \not{4}, \not{1}, \not{2})$$

Each tree produces a unique Prüfer sequence. Moreover, we have the following proposition.

Proposition 3.6. *There is a bijection between the set of spanning trees on n vertices and $[n]^{n-2}$.*

We will prove a more general version of this proposition in the next section.

Remark 3.7. Cayley's formula (Theorem 3.1) follows directly from this proposition.

In fact, Prüfer sequences allow us to prove a stronger result about the number of trees with a given degree sequence.

Lemma 3.8. *Let T be a tree on vertex set $[n]$. Suppose that $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n)$ is the degree sequence of T, where δ_i is the degree of the vertex i. Then*

$$\delta_1 + \cdots + \delta_n = 2(n-1).$$

Proof. By the Handshaking Lemma (Lemma 2.15), the degree sum is equal to 2 times the number of edges. But every tree has $n-1$ edges, so the result follows. □

Proposition 3.9. *Let $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n) \in \mathbb{N}^n$ such that $\delta_1 + \cdots + \delta_n = 2(n-1)$. Then the number of trees on $[n]$ with degree sequence $\boldsymbol{\delta}$ is*

$$\binom{n-2}{\delta_1 - 1, \dots, \delta_n - 1} = \frac{(n-2)!}{(\delta_1 - 1)! \dots (\delta_n - 1)!}.$$

Proof. Let $\boldsymbol{c}$ be the Prüfer sequence of a spanning tree T with degree sequence δ. By the construction of a Prüfer sequence, the number of times i appears in $\boldsymbol{c}$ is $\delta_i - 1$. Thus, the number of possible sequences $\boldsymbol{c}$ which give a tree with degree sequence $\boldsymbol{\delta}$ is equal to the number of permutations of a sequence with $\delta_1 - 1$ entries equal to 1, $\delta_2 - 1$ entries equal to 2, ..., $\delta_n - 1$ entries equal to n. □

3.2 Prüfer sequences for s-uniform trees

There have been various generalisations of Prüfer sequences to hypertrees, see [27], [20] and [6]. We give a special case of Lavault's algorithm in [20] for s-uniform trees.

Definition 3.10. A *tree* (or *hypertree*) is a simple connected hypergraph with no cycles. In particular, this means that the intersection of two edges is size at most 1. An *s-uniform tree* is a tree which is an s-uniform hypergraph. A *spanning tree* of a hypergraph H is a spanning subhypergraph which is a tree.

Proposition 3.11. *If H is an s-uniform spanning tree with t edges, then the number of vertices in H is $n = (s-1)t + 1$. In particular, $s - 1 \mid n - 1$.*

Proof. We prove this by induction on t. A tree with one edge has $s = (s-1) + 1$ vertices. Adding an edge increases the number of vertices by $s - 1$. □

Definition 3.12. A *leaf* of a hypertree is a set ℓ of vertices of degree 1 belonging to the same edge e, where the size of ℓ is one less than the size of e. The *stem* of a leaf is the unique vertex in $e \setminus \ell$.

Example 3.13. Figure 3.1 shows a 3-uniform spanning tree on $n = 9$ vertices and $t = 4$ edges. The leaves are $\{1,9\}, \{4,5\}$ and $\{2,8\}$ and the corresponding stems are $3, 3$ and 7.

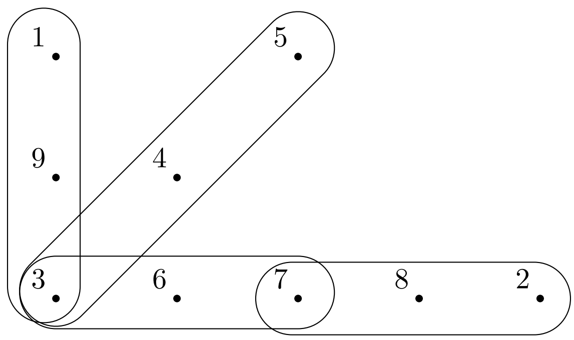


Figure 3.1: A 3-uniform spanning tree on 9 vertices.

Proposition 3.14. *Every s-uniform tree with at least 2 edges has at least 2 leaves.*

Proof. Let T be an s-uniform tree. Consider the longest path $e_1, e_2, \dots, e_k$ in T. We claim that the edges e_1 and e_k must each contain a leaf. Suppose that e_1 does not contain a leaf. Then e_1 has at least two vertices which have degree greater than or equal to 2. One of these vertices v will be in e_2. The other vertex u will not be in e_2, because T has no 2-cycles. Then u is contained in another edge. This edge cannot be in the path we are considering, otherwise there would be a cycle in T. Therefore, we can add this edge to our path to obtain a longer path in T, a contradiction. By the same argument, we conclude that e_k has a leaf. □

We now describe an algorithm which encodes an s-uniform tree.

Algorithm 3.15 (`Hypertree-to-Prüfer`)**.** For a tree $T \equiv T_0$ on $n = (s-1)t+1$ vertices, start with the empty vector $\boldsymbol{c}$ and an empty partition P. Repeat the following process until one edge e remains:

1. In T_i, consider all leaves which do not contain the vertex n (there exists at least one such leaf by Proposition 3.14). Out of these leaves, identify the leaf ℓ which has the smallest-labelled vertex.
2. Add ℓ to the partition P and append $\boldsymbol{c}$ with the label of the stem of ℓ.
3. Set $T_{i+1} := T_i - \ell$.

The final edge e will contain the vertex n. Add $e \setminus n$ to P.

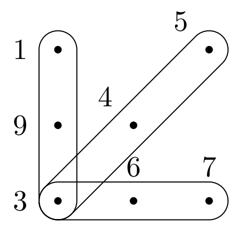
At the conclusion of the algorithm, P will be a partition of the set $[n-1]$ into disjoint $(s-1)$-subsets and $\boldsymbol{c} \in [n]^{t-1}$. We call $(P, \boldsymbol{c})$ the *Prüfer sequence* of T.

Remark 3.16. For the case $s = 2$, this algorithm returns $(P, \boldsymbol{c})$, where P is a partition of the set $[n-1]$ into 1-subsets. Because there is only one such partition, we can omit P. Thus, this algorithm reduces to `Tree-to-Prüfer` in the graph case.

Let us apply this algorithm to the 3-uniform spanning tree in Figure 3.1.

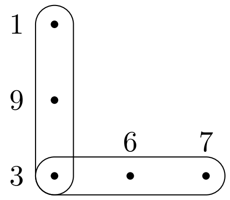
Example 3.17. Here, $n = 9$ and $t = 4$. Start with the empty vector $\boldsymbol{c}$ and empty partition P.

- The leaf with the smallest label which does not contain 9 is $\{2, 8\}$. Add $\{2, 8\}$ to P and record 7. Delete $\{2, 8\}$.



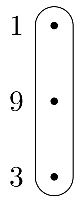
$$P = \{\{2, 8\}\}, \quad \boldsymbol{c} = (7, \cdot, \cdot)$$

- The leaf with the smallest label which does not contain 9 is $\{4,5\}$. Add $\{4,5\}$ to P and record 3. Delete $\{4,5\}$.



$$P = \{\{2,8\},\{4,5\}\}, \quad \boldsymbol{c} = (7,3,\cdot)$$

- The leaf with the smallest label which does not contain 9 is $\{6,7\}$. Add $\{6,7\}$ to P and record 3. Delete $\{6,7\}$.



$$P = \{\{2,8\},\{4,5\},\{6,7\}\}, \quad \boldsymbol{c} = (7,3,3)$$

- Add $\{1,3\}$ to P.

$$P = \{\{2,8\},\{4,5\},\{6,7\},\{1,3\}\}, \quad \boldsymbol{c} = (7,3,3).$$

Let $\mathscr{P}_{n-1,s-1}$ denote the set of all possible partitions of $[n-1]$ into subsets of size $s-1$.

Proposition 3.18. `Hypertree-to-Prüfer` *gives a surjective map from the set of s-uniform trees on $n = (s-1)t+1$ vertices to $\mathscr{P}_{n-1,s-1} \times [n]^{t-1}$.*

Proof. We prove the proposition by induction on the number of edges t. When $t = 1$, the algorithm returns $(P, \boldsymbol{c})$, where $P = \{\{1,\dots,n-1\}\}$ and $\boldsymbol{c}$ is the "empty vector", so the proposition is trivially true.

Suppose the proposition is true for s-uniform trees with $t-1$ edges, and consider an s-uniform tree with t edges and $n = (s-1)t+1$ vertices. We want to show that the algorithm surjectively maps into $\mathscr{P}_{n-1,s-1} \times [n]^{t-1}$.

Take an arbitrary $P \in \mathscr{P}_{n-1,s-1}$ and an arbitrary $\boldsymbol{c} = (c_1,\dots,c_{t-1}) \in [n]^{t-1}$. Find the lexicographically smallest set in P which does not contain any elements of $\boldsymbol{c}$, and call it $\{a_1,\dots,a_{s-1}\}$, where $a_1,\dots,a_{s-1} \in [n-1] \setminus \{c_1,\dots,c_{t-1}\}$. Note that such a set exists because the length of $\boldsymbol{c}$ is $t-1$ and there are t sets to choose from in P.

Consider now the partition $P' = P \setminus \{\{a_1,\dots,a_{s-1}\}\}$ and $\boldsymbol{c}' = (c_2,\dots,c_{t-1})$. By the inductive hypothesis, there exists a tree $T' = (V',E')$ with $V' = [n] \setminus \{a_1,\dots,a_{s-1}\}$ which maps to $(P',\boldsymbol{c}')$ via this algorithm.

Let $T = (V, E)$, where $V = [n]$ and $E = E' \cup \{\{a_1, \ldots, a_{s-1}, c_1\}\}$. Then we can see that T maps to $(P, \boldsymbol{c})$ when we run through the algorithm. This completes the proof.

$\square$

We now describe an algorithm which takes an element of $(P, \boldsymbol{c}) \in \mathscr{P}_{n-1,s-1} \times [n]^{t-1}$ and outputs a unique tree.

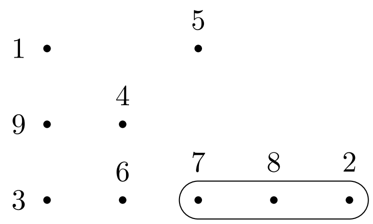
Algorithm 3.19 (`Prüfer-to-Hypertree`)**.** Let $n = (s-1)t + 1$, where $t \in \mathbb{Z}^+$. For $(P, \boldsymbol{c}) \in \mathscr{P}_{n-1,s-1} \times [n]^{t-1}$, start with vertex set $V = [n]$ and edge set $E = \varnothing$. Repeat the following process until $\boldsymbol{c}$ is empty:

1. Out of the sets in P which do not contain any vertices in $\boldsymbol{c}$, find the set $\{a_1, \ldots, a_{s-1}\}$ which contains the smallest label.
2. Build an edge containing the vertices $a_1, \ldots, a_{s-1}$ and the first vertex in $\boldsymbol{c}$.
3. Delete $\{a_1, \ldots, a_{s-1}\}$ from P and delete the first vertex in $\boldsymbol{c}$.

At the conclusion of these steps, one set will remain in P. Build the last edge with the last set in P and the vertex n. Then $T = (V, E)$ is an s-uniform tree corresponding to $(P, \boldsymbol{c})$.

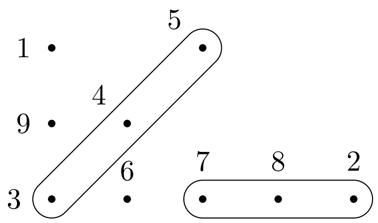
Example 3.20. Suppose $P = \{\{2, 8\}, \{4, 5\}, \{6, 7\}, \{1, 3\}\}$ and $\boldsymbol{c} = (7, 3, 3)$. Start with vertex set $V = [9]$ and edge set $E = \varnothing$.

- The smallest set in P which contains the smallest label and does not contain any vertices in $\boldsymbol{c}$ is $\{2, 8\}$. Build the edge $\{2, 8, 7\}$.



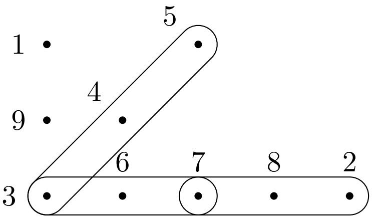
$P = \{\cancel{\{2, 8\}}, \{4, 5\}, \{6, 7\}, \{1, 3\}\}, \quad \boldsymbol{c} = (\cancel{7}, 3, 3)$

- The smallest set in P which contains the smallest label and does not contain any vertices in $\boldsymbol{c}$ is $\{4, 5\}$. Build the edge $\{3, 4, 5\}$.



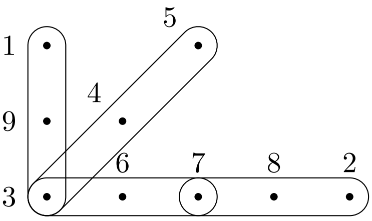
$P = \{\cancel{\{2, 8\}}, \cancel{\{4, 5\}}, \{6, 7\}, \{1, 3\}\}, \quad \boldsymbol{c} = (\cancel{7}, \cancel{3}, 3)$

- The smallest set in P which contains the smallest label and does not contain any vertices in $\boldsymbol{c}$ is $\{6, 7\}$. Build the edge $\{3, 6, 7\}$.



$P = \{\cancel{\{2,8\}}, \cancel{\{4,5\}}, \cancel{\{6,7\}}, \{1,3\}\}, \quad \boldsymbol{c} = (\cancel{7}, \cancel{3}, \cancel{3})$

- Build the edge $\{1, 3, 9\}$.



$P = \{\cancel{\{2,8\}}, \cancel{\{4,5\}}, \cancel{\{6,7\}}, \cancel{\{1,3\}}\}, \quad \boldsymbol{c} = (\cancel{7}, \cancel{3}, \cancel{3})$

Proposition 3.21. `Prüfer-to-Hypertree` *gives a surjective map from* $\mathscr{P}_{n-1,s-1} \times [n]^{t-1}$ *to the set of* s*-uniform spanning trees on* $[n]$*, where* $n = (s-1)t + 1$*.*

Proof. We prove the proposition by induction on t. In the case where $t = 1$, the algorithm outputs the only possible tree on s vertices (an edge containing all s vertices).

Suppose the proposition is true for trees with $t - 1$ edges, and take an arbitrary tree $T = (V, E)$ with $n = (s-1)t + 1$ vertices. Let $\{a_1, \dots, a_{s-1}\}$ be the smallest leaf which does not contain n (such a leaf exists by Proposition 3.14). Suppose the stem of this leaf is c_1.

Consider the tree $T' = (V', E')$, where $V' = V \setminus \{a_1, \dots, a_{s-1}\}$ and $E' = E \setminus \{e\}$, where $e = \{a_1, \dots, a_{s-1}, c_1\}$. By the inductive hypothesis, there exists P', a partition of the set $[n-1] \setminus \{a_1, \dots, a_{s-1}\}$ into disjoint $(s-1)$-subsets, and $\boldsymbol{c}' = (c_2, \dots, c_{t-1})$, such that $(P', \boldsymbol{c}')$ maps to T'.

Let $P = P' \cup \{a_1, \dots, a_{s-1}\}$ and $\boldsymbol{c} = (c_1, c_2, \dots, c_{t-1})$. Then we can see that $(P, \boldsymbol{c})$ maps to T when we run through the algorithm. □

Propositions 3.18 and 3.21 together prove the following theorem.

Theorem 3.22. *Let* $n = (s-1)t + 1$ *for some* $t \in \mathbb{Z}^+$*. Then there is a bijection between the set of* s*-uniform spanning trees on* n *vertices and* $\mathscr{P}_{n-1,s-1} \times [n]^{t-1}$*.*

Remark 3.23. Taking $s = 2$ in Theorem 3.22 proves Proposition 3.6.

Using Theorem 3.22, we can generalise Cayley's formula to s-uniform hypergraphs. This formula was shown in [28, Theorem 2] and [20, Corollary 1].

Corollary 3.24. *The number of s-uniform trees on $n = (s-1)t + 1$ vertices is*

$$\frac{n^{\frac{n-1}{s-1}-1}\,(n-1)!}{\left(\frac{n-1}{s-1}\right)!\,((s-1)!)^{\frac{n-1}{s-1}}} = \frac{n^{t-1}\,(n-1)!}{t!\,((s-1)!)^{t}}.$$

Proof. By Theorem 3.22, it suffices to count the size of $\mathscr{P}_{n-1,s-1} \times [n]^{t-1}$. The number of ways of partitioning $[n-1]$ into disjoint $(s-1)$-subsets is

$$|\mathscr{P}_{n-1,s-1}| = \frac{\binom{n-1}{s-1}\binom{n-1-(s-1)}{s-1}\cdots\binom{s-1}{s-1}}{\left(\frac{n-1}{s-1}\right)!} = \frac{(n-1)!}{\left(\frac{n-1}{s-1}\right)!\,((s-1)!)^{\frac{n-1}{s-1}}}. \tag{3.1}$$

The size of $[n]^{t-1}$ is n^{t-1}. Multiplying these gives us the result. □

We can also count the number of s-uniform trees with a specified degree sequence.

Lemma 3.25. *Let T be a s-uniform spanning tree on $n = (s-1)t + 1$ vertices (where t is the number of edges). Suppose $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)$ is the degree sequence of T (where δ_i is the degree of the vertex i). Then*

$$\delta_1 + \cdots + \delta_n = \frac{s(n-1)}{s-1} = st.$$

Proof. This is follows directly from the Handshaking Lemma for s-uniform hypergraphs (Lemma 2.45). □

This allows us to give the following definition.

Definition 3.26. An *(s-uniform) tree degree sequence* on $n = (s-1)t + 1$ vertices is a sequence $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n) \in \mathbb{N}^n$ such that

$$\delta_1 + \cdots + \delta_n = st \quad \text{and} \quad \delta_i \geq 1 \text{ for } i = 1, \ldots, n.$$

Let $\mathcal{D}_n^{(s)} = \mathcal{D}_n$ be the set of s-uniform tree degree sequences. We will mostly omit the superscript for brevity.

From our analysis of Prüfer sequences, we can prove the following result, a special case of [5, Theorem 1.1].

Corollary 3.27. *The number of trees with tree degree sequence $\boldsymbol{\delta} \in \mathcal{D}_n$ is*

$$\binom{t-1}{\delta_1-1,\ldots,\delta_n-1}\frac{(n-1)!}{t!\,((s-1)!)^t} = \frac{(s-1)(n-2)!}{((s-1)!)^{\frac{n-1}{s-1}}}\prod_{i=1}^{n}\frac{1}{(\delta_i-1)!}.$$

Proof. By the construction of the Prüfer sequence, the number of times i appears in $\boldsymbol{c}$ is $\delta_1 - 1$. Thus, the number of possible $\boldsymbol{c}$ which will give a tree of degree sequence $\boldsymbol{\delta}$ is equal to the number of permutations of a sequence of $\delta_i - 1$ entries equal to 1, $\delta_2 - 1$ entries equal to 2, ..., $\delta_n - 1$ entries equal to n, which is

$$\frac{(t-1)!}{(\delta_1-1)!\ldots(\delta_n-1)!} = \binom{\frac{n-1}{s-1}-1}{\delta_1-1,\ldots,\delta_n-1}.$$

Each of these sequences is then paired with a different partition $P \in \mathscr{P}_{n-1,s-1}$, and by (3.1), the number of choices for P is

$$\frac{(n-1)!}{t!\,((s-1)!)^t}.$$

Multiplying gives us the result. □

3.3 Prüfer sequences for arbitrary hypertrees

The algorithms described in Section 3.2 can be adapted to general hypertrees, as long as each edge has at least two vertices. Let $\mathbb{P}_{n-1,t}$ be the set of all partitions of $[n-1]$ into t disjoint non-empty subsets and let T be a (not necessarily s-uniform) hypertree with n vertices and t edges. Then if we run through `Prüfer-to-Hypertree`, we will instead obtain an element of $\mathbb{P}_{n-1,t} \times [n]^{t-1}$. Similarly, if we take an element $(P, \boldsymbol{c}) \in \mathbb{P}_{n-1,t} \times [n]^{t-1}$ and apply `Hypertree-to-Prüfer`, then we will obtain a hypertree with n vertices and t edges. In fact, we can prove the following proposition.

Proposition 3.28. *There is a bijection between $\mathbb{P}_{n-1,t} \times [n]^{t-1}$ and the set of spanning hypertrees on n vertices with t edges, where each edge has at least two vertices.*

We omit the proof of this, as it is similar to the s-uniform case.

Definition 3.29. The *Stirling numbers of the second kind*, written $\left\{ {n \atop k} \right\}$, count the number of ways to partition the set $[n]$ into k non-empty subsets. It is well known (see Wilf [30]) that

$$\left\{ {n \atop k} \right\} = \frac{1}{k!}\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}j^n.$$

With this definition in mind, Proposition 3.28 gives us the following results.

Corollary 3.30. *The number of hypertrees on n vertices with t edges, where each edge has at least two vertices, is*

$$n^{t-1}\left\{ {n-1 \atop t} \right\}.$$

This reaffirms the following result by Warme from his Ph.D thesis [29, Corollary 3.15.1].

Corollary 3.31. *The number of hypertrees on n vertices, where each edge has at least two vertices, is*

$$\sum_{t=1}^{n-1} n^{t-1}\left\{ {n-1 \atop t} \right\}.$$

CHAPTER 4

Expected number of spanning trees

Expect the unexpected.

– Unknown

This chapter contains some original results about the expected number of spanning trees in a random r-regular s-uniform hypergraph from the configuration model. We establish threshold results which allow us to asymptotically determine the existence of spanning trees as the number of vertices grows.

4.1 First moment calculation

For the remainder of the thesis, unless otherwise stated, we fix integers $r \geq 3$ and $s \geq 2$. We work in the configuration model $\mathcal{P}_{n,r,s}$, where $s \mid rn$ and $n = (s-1)t + 1$, for some $t \in \mathbb{N}$. Let Y be the number of s-uniform spanning trees in a random partition P from $\mathcal{P}_{n,r,s}$.

Theorem 4.1. *For fixed $r \geq 3$, $s \geq 2$, we have*

$$\mathbb{E}Y \sim \frac{(r-1)^{1/2}(s-1)}{n\,(rs-r-s)^{\frac{s+1}{2(s-1)}}} \left(\frac{(s-1)^{r/s}(r-1)^{r-1}}{r^{\frac{rs-r-s}{s}}\,(rs-r-s)^{\frac{rs-r-s}{s(s-1)}}} \right)^n .$$

Proof. Let $\mathcal{T}_n$ be the set of all s-uniform trees on n vertices. For a random partition P from $\mathcal{P}_{n,r,s}$, we can write

$$\mathbb{E}Y = \sum_{P_T : G(P_T) \in \mathcal{T}_n} \Pr(P_T \subseteq P) = \sum_{P_T : G(P_T) \in \mathcal{T}_n} \frac{|\{P \in \Omega_{n,r,s} : P_T \subseteq P\}|}{|\Omega_{n,r,s}|}.$$

Selecting P_T uses up st points, by Lemma 3.25. Thus, for a given P_T, the size of $\{P \in \Omega_{n,r,s} : P_T \subseteq P\}$ is $f_s(rn - st)$, the number of ways to partition the remaining $rn - st$ points after the points of the tree are selected. We can then split the sum according to the tree that P_T projects to:

$$|\Omega_{n,r,s}|\, \mathbb{E}Y = f_s(rn - st) \sum_{P_T : G(P_T) \in \mathcal{T}_n} 1 = f_s(rn - st) \sum_{T \in \mathcal{T}_n} \sum_{P_T : G(P_T) = T} 1.$$

We further split the sum on tree degree sequence. Recall from Definition 3.26 that $\mathcal{D}_n$ is the set of possible tree degree sequences on n vertices. Given $\boldsymbol{\delta} \in \mathcal{D}_n$, define $\mathcal{T}_n(\boldsymbol{\delta})$ to be the set of trees with a degree sequence $\boldsymbol{\delta}$.

Thus, we can write

$$|\Omega_{n,r,s}|\, \mathbb{E}Y = f_s(rn - st) \sum_{\boldsymbol{\delta} \in \mathcal{D}_n} \sum_{T \in \mathcal{T}_n(\boldsymbol{\delta})} \sum_{P_T : G(P_T) = T} 1.$$

To compute the inner sum, we must count the number of subpartitions that project to a given tree $T \in \mathcal{T}_n(\boldsymbol{\delta})$, for a given $\boldsymbol{\delta} \in \mathcal{D}_n$. Exactly δ_i of the points in cell i must contribute to P_T, and there are $(r)_{\delta_j}$ ways to choose and order these points. So there are

$$\prod_{j=1}^{n} (r)_{\delta_j}$$

possible partitions P_T which project to T.

Thus

$$|\Omega_{n,r,s}|\, \mathbb{E}Y = f_s(rn - st) \sum_{\boldsymbol{\delta} \in \mathcal{D}_n} |\mathcal{T}_n(\boldsymbol{\delta})| \prod_{j=1}^{n} (r)_{\delta_j}.$$

By Corollary 3.27, the number of trees with degree sequence $\boldsymbol{\delta}$ is

$$|\mathcal{T}_n(\boldsymbol{\delta})| = \frac{(s-1)(n-2)!}{((s-1)!)^t} \prod_{j=1}^{n} \frac{1}{(\delta_j - 1)!}.$$

Thus,

$$|\Omega_{n,r,s}|\,\mathbb{E}Y = f_s(rn-st)\,\frac{(s-1)(n-2)!}{((s-1)!)^t}\left(\sum_{\boldsymbol{\delta}\in\mathcal{D}_n}\prod_{j=1}^{n}\frac{(r)_{\delta_j}}{(\delta_j-1)!}\right).$$

We use a generating function to compute the sum. Recalling that $\delta_1+\cdots+\delta_n = st$, the number of subpartitions P_T which project to a spanning tree is

$$\begin{aligned}
\frac{(s-1)(n-2)!}{((s-1)!)^t}\sum_{\boldsymbol{\delta}\in\mathcal{D}_n}\prod_{j=1}^{n}\frac{(r)_{\delta_j}}{(\delta_j-1)!} &= \frac{(s-1)(n-2)!}{((s-1)!)^t}[z^{st}]\left(\sum_{j=1}^{\infty}\frac{(r)_j}{(j-1)!}z^j\right)^n\\
&= \frac{(s-1)(n-2)!}{((s-1)!)^t}\,[z^{st}]\left(\sum_{j\geq 1} rz\binom{r-1}{j-1}z^{j-1}\right)^n\\
&= \frac{(s-1)(n-2)!}{((s-1)!)^t}\,[z^{st}]\left(rz(1+z)^{r-1}\right)^n\\
&= \frac{r^n(s-1)(n-2)!}{((s-1)!)^t}\binom{(r-1)n}{t-1}, \qquad (4.1)
\end{aligned}$$

where square brackets denotes coefficient extraction.

By (2.12) and recalling that $t=\frac{n-1}{s-1}$, we have

$$\begin{aligned}
\mathbb{E}Y &= \frac{f_s(rn-st)}{f_s(rn)}\,\frac{r^n(s-1)(n-2)!}{((s-1)!)^t}\binom{(r-1)n}{t-1}\\
&= \frac{\left(\frac{rs-r-s}{s-1}n+\frac{s}{s-1}\right)!\,(rn/s)!\,(s!)^{rn/s}\cdot r^n(s-1)(n-2)!\cdot((r-1)n)!}{\left(\frac{rs-r-s}{s(s-1)}n+\frac{1}{s-1}\right)!\,(s!)^{\frac{rs-r-s}{s(s-1)}n+\frac{1}{s-1}}\,(rn)!\cdot((s-1)!)^{\frac{n-1}{s-1}}\cdot\left(\frac{n-s}{s-1}\right)!\left(\frac{rs-r-s}{s-1}n+\frac{s}{s-1}\right)!}\\
&= r^n s^{\frac{n-1}{s-1}}\cdot\frac{((r-1)n)!}{(rn)!}\cdot\frac{(n-1)!}{\left(\frac{n-1}{s-1}\right)!}\cdot\frac{(rn/s)!}{\left(\frac{rs-r-s}{s(s-1)}n+\frac{1}{s-1}\right)!} \qquad (4.2)
\end{aligned}$$

Applying Corollary 2.26 gives the result. The omitted computational details are similar to the calculations at the end of the the proof of Lemma 2.28. □

Remark 4.2. Letting $s=2$ recovers the result from [16] for graphs:

$$\mathbb{E}Y \sim \frac{(r-1)^{1/2}}{n(r-2)^{3/2}}\left(\frac{(r-1)^{r-1}}{(r^2-2r)^{r/2-1}}\right)^n.$$

4.2 Threshold analysis

Consider the base of the exponential factor in Theorem 4.1:

$$\frac{(s-1)^{r/s}(r-1)^{r-1}}{r^{\frac{rs-r-s}{s}}(rs-r-s)^{\frac{rs-r-s}{s(s-1)}}}.$$

We can ask the interesting question: When is this expression less than or equal to 1 and when is this expression greater than 1? If the expression is less than 1, then $\mathbb{E}Y = o(1)$, and if the expression is greater than 1, then $\mathbb{E}Y \to \infty$.

Define the logarithm of the base of the exponential factor as a function:

$$L(r,s) = L_s(r) = \frac{r}{s}\log(s-1)+(r-1)\log(r-1)-\frac{rs-r-s}{s}\log r-\frac{rs-r-s}{s(s-1)}\log(rs-r-s),$$

treating $r \geq 2$ ($r \geq 3$ when $s = 2$) as a continuous variable and s as a fixed integer. We want to determine when $L_s(r) > 0$ and when $L_s(r) < 0$.

Lemma 4.3. *For $s \in \{2,3,4\}$, $L_s(r) > 0$ for $r \in (2,\infty)$. Furthermore, for $s \in \{3,4\}$, $L_s(2) > 0$.*

Proof. This can be checked directly using elementary calculus. □

Figure 4.1 shows a plot of $L_s(r)$ for $s = 2, 3$ and 4.

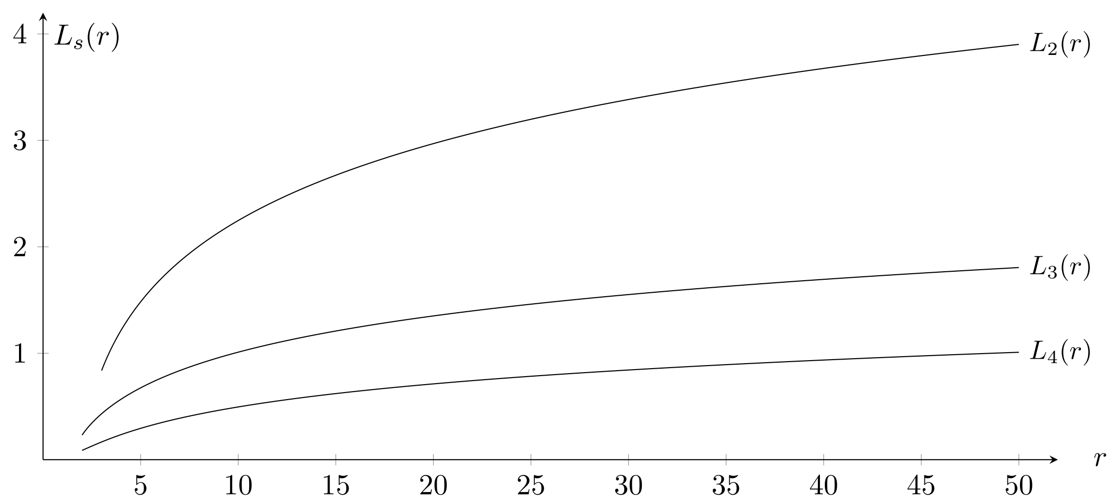


Figure 4.1: Plot of $L_2(r), L_3(r)$ and $L_4(r)$.

This implies that for $s \in \{2, 3, 4\}$, we have $\mathbb{E}Y \to \infty$. However, for $s \geq 5$, the expression $L_s(r)$ may be negative (perhaps surprisingly). Figure 4.2 shows a plot of $L_s(r)$ for $s = 5, 6$ and 7.

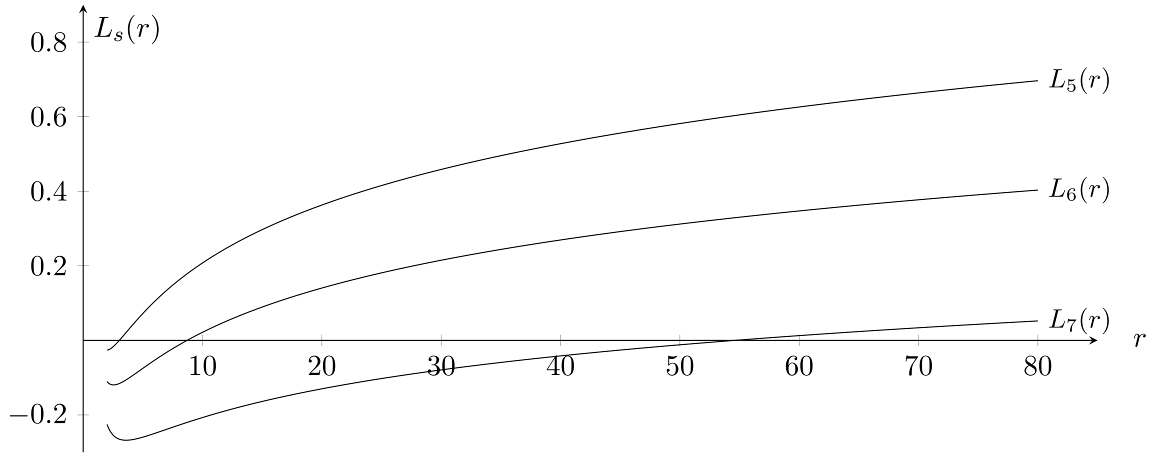


Figure 4.2: Plot of $L_5(r)$, $L_6(r)$ and $L_7(r)$.

Further exploring the nature of the function, Figure 4.3 is a contour plot of $L(r, s)$.

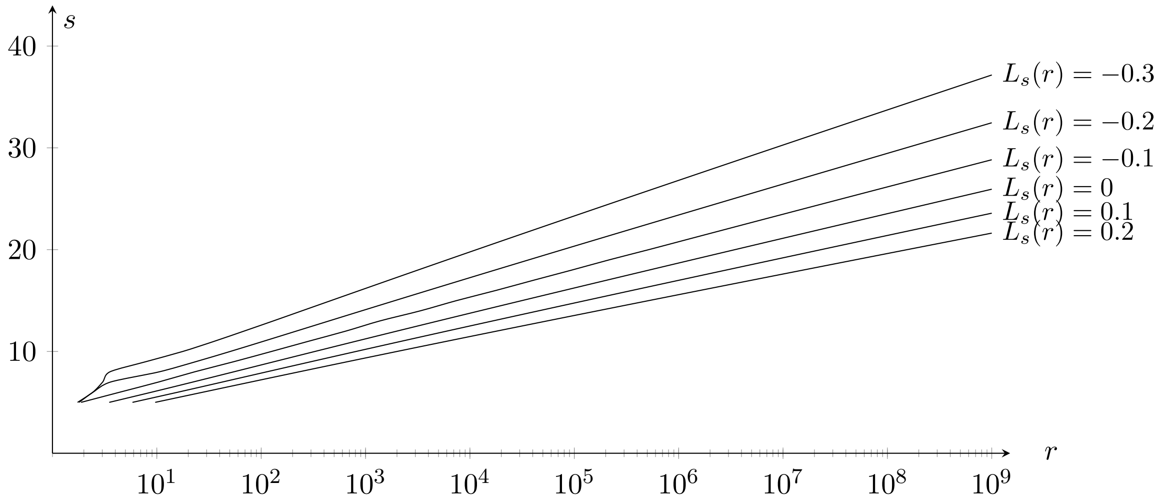


Figure 4.3: Contour plot of $L_s(r)$.

This plot leads us to believe that, for a fixed $s \geq 5$, there exists a unique threshold $r = \rho(s)$, where $L_s(\rho(s)) = 0$.

Indeed, we have the following theorem.

Theorem 4.4. *For fixed integer $s \geq 5$, there exists a unique real number $\rho(s) > 2$ such that $L(\rho(s), s) = 0$,*

$$L(r,s) < 0 \quad \text{for } r \in [2, \rho(s)) \quad \text{and} \quad L(r,s) > 0 \text{ for } r \in (\rho(s), \infty).$$

Furthermore, if $s \geq 6$, then $\rho(s) > s$.

We check this theorem directly for $s = 5$. In order to prove this theorem for $s \geq 6$, we will prove that:

1. The function $L_s(r)$ has precisely one inflexion point at $r = \frac{1}{2}(s + \sqrt{s(s-4)}) < s$, where the second derivative changes from positive to negative.
2. The function $L_s(r)$ is decreasing at $r = 2$ and increasing at $r = s$.
3. The function $L_s(r)$ is negative at $r = 2$ and $r = s$.
4. The function $L_s(r)$ is monotonically increasing for $r \geq s$.
5. The function $L_s(r)$ approaches infinity as $r \to \infty$.

We can see from Figure 4.2 that $L_6(r)$ and $L_7(r)$ are initially decreasing, and then increasing. It turns out that for all $s \geq 6$, the function $L_s(r)$ has this shape.

For reference, we have the following first and second derivatives with respect to r:

$$\begin{aligned} L_s'(r) &= \frac{1}{r} + \log(r-1) - \frac{s-1}{s}\log r - \frac{1}{s}\log\left(r - \frac{s}{s-1}\right), \\ L_s''(r) &= \frac{1}{r^2}\left(\frac{1}{r-1} - \frac{r}{rs-r-s}\right). \end{aligned}$$

The following inequality from [22, p. 435] is useful in the subsequent proofs.

Lemma 4.5 ([22])**.** *For $b \geq 1$ and $|a| < b$, we have*

$$\left(1 + \frac{a}{b}\right)^b \geq e^a \left(1 - \frac{a^2}{b}\right).$$

We first prove that the function $L_s(r)$ has precisely one inflexion point.

Lemma 4.6. *For $s \geq 5$, $L_s''(r) \geq 0$ when $r \leq \frac{1}{2}(s + \sqrt{s(s-4)})$ and $L_s''(r) < 0$ when $r > \frac{1}{2}(s + \sqrt{s(s-4)})$.*

Proof. We can see that $L_s''(r) \geq 0$ if and only if $r^2 - rs + s \leq 0$, which reduces to the result. □

The following lemma shows us that $L_s(r)$ is decreasing at $r = 2$ and increasing at $r = s$.

Lemma 4.7. *For $s \geq 6$, $L_s'(2) < 0$ and $L_s'(s) > 0$.*

Proof. We have $L_s'(2) < 0$ if and only if $\sqrt{e} < 2^{\frac{s-1}{s}}(2 - \frac{s}{s-1})^{1/s}$. For $s \geq 6$, we have

$$2^{\frac{s-1}{s}}\left(2 - \frac{s}{s-1}\right)^{1/s} > 2^{5/6}\left(2 - \frac{6}{5}\right)^{1/s} = 2^{5/6}\left(\frac{4}{5}\right)^{1/6} > \sqrt{e}.$$

For the second inequality, we have $L_s'(s) > 0$ if and only if $e\left(1 - \frac{1}{s}\right)^s > \frac{s-2}{s-1}$. Making use of Lemma 4.5 with $a = -1$, we have

$$e\left(1 - \frac{1}{s}\right)^s > \frac{s-1}{s} > \frac{s-2}{s-1},$$

completing the proof. □

The following lemma shows that the value of $L_s(r)$ at $r = 2$ and $r = s$ are negative.

Lemma 4.8. *For $s \geq 6$, $L_s(2) < 0$ and $L_s(s) < 0$.*

Proof. We have

$$L_s(2) = \frac{1}{s(s-1)} \log\left(\frac{(s-1)^{2(s-1)}}{2^{(s-1)(s-2)}(s-2)^{s-2}}\right),$$

and this is negative if and only if

$$(s-1)^s < \left(\frac{2^{s-1}(s-2)}{s-1}\right)^{s-2}.$$

It suffices to show that $(s-1)^s < 2^{(s-2)^2}$ because $\frac{1}{2} < \frac{s-2}{s-1}$. We proceed by induction. As $5^6 < 2^{16}$, the base case is true. Suppose the claim is true for some $s \geq 6$. Then, using the fact that $s < 2^{s-3}$ when $s \geq 6$, we have

$$s^{s+1} = s\left(\frac{s}{s-1}\right)^s (s-1)^s < 2^{s-3}\, 2^s\, 2^{(s-2)^2} = 2^{(s-1)^2}.$$

For the second part, we have

$$L_s(s) = \log\left(\frac{(s-1)^s}{s^{\frac{s(s-2)}{s-1}}(s-2)^{\frac{s-2}{s-1}}}\right).$$

For $s = 6$, we verify directly: $L_6(6) = 6\log(5) - 4\log(6) - \frac{4\log(24)}{5} < 0$. From now on, assume $s \geq 7$. We have

$$\frac{(s-1)^s}{s^{\frac{s(s-2)}{s-1}}(s-2)^{\frac{s-2}{s-1}}} = \left(\frac{s-1}{s}\right)^s \left(\frac{s^s}{(s-2)^{s-2}}\right)^{\frac{1}{s-1}} < e^{-1}\left(\frac{s^s}{(s-2)^{s-2}}\right)^{\frac{1}{s-1}}.$$

and it can be shown that $\left(\frac{s^s}{(s-2)^{s-2}}\right)^{\frac{1}{s-1}}$ is a decreasing function of s. Indeed, $\left(\frac{s^s}{(s-2)^{s-2}}\right)^{\frac{1}{s-1}} = \frac{s}{s-2}(s(s-2))^{\frac{1}{s-1}}$, and both $\frac{s}{s-2}$ and $(s(s-2))^{\frac{1}{s-1}}$ are positive decreasing functions.

So we have, for $s \geq 7$,

$$L_s(s) = \log\left(\frac{(s-1)^s}{s^{\frac{s(s-2)}{s-1}}(s-2)^{\frac{s-2}{s-1}}}\right) < \log\left(e^{-1}\left(\frac{7^7}{5^5}\right)^{1/6}\right) < 0.$$

□

The following lemma tells us that the function $L_s(r)$ approaches infinity as $r \to \infty$.

Lemma 4.9. *For fixed $s \geq 2$,*

$$L_s(r) \to \infty \qquad \text{as } r \to \infty.$$

Proof. We can rewrite the base of the exponential factor as

$$(s-1)^{\frac{1}{s-1}} \frac{(r-1)^{r-1}}{r^{\frac{rs-r-s}{s}}\left(r - \frac{s}{s-1}\right)^{\frac{rs-r-s}{s(s-1)}}} = ((r-1)(s-1))^{\frac{1}{s-1}} \left(\frac{r-1}{r}\right)^{\frac{rs-r-s}{s}} \left(\frac{r-1}{r-\frac{s}{s-1}}\right)^{\frac{rs-r-s}{s(s-1)}}.$$

Now, we have

$$\left(\frac{r-1}{r}\right)^{\frac{rs-r-s}{s}} \to \exp\left(-\frac{s-1}{s}\right) \quad \text{and} \quad \left(\frac{r-1}{r-\frac{s}{s-1}}\right)^{\frac{rs-r-s}{s(s-1)}} \to \exp\left(\frac{1}{s(s-1)}\right)$$

as $r \to \infty$, so

$$L_s(r) \sim \frac{1}{s-1}\log(r-1) + \frac{1}{s-1}\log(s-1) - \frac{s-2}{s-1},$$

which approaches ∞ as $r \to \infty$. □

We can now prove Theorem 4.4.

Proof of Theorem 4.4. We can verify this directly when $s = 5$. So suppose $s \geq 6$.

First, we claim that $L(r, s)$ is negative when $r \leq s$. Lemmas 4.6 and 4.7 together imply there is one unique stationary point of $L_s(r)$ when $r \in [2, s]$, which is a local minimum. Thus, the maximum value of $L_s(r)$ in the interval $[2, s]$ is either $L_s(2)$ or $L_s(r)$. Lemma 4.8 tells us these values are both negative, so the claim is proven.

Next, we show that $L_s(r)$ is monotonically increasing when $r \geq s$. Using Lemma 4.5 with $a = -1$, and the fact that $r \geq s$, we have

$$e^s \left(1 - \frac{1}{r}\right)^{rs} \geq \left(1 - \frac{1}{r}\right)^{s} \geq \left(1 - \frac{1}{r}\right)^{r} > \left(1 - \frac{s}{r(s-1)}\right)^{r}.$$

This inequality is equivalent to $L_s'(r) > 0$.

Since $L_s(s)$ is negative and $L_s(r)$ tends to infinity as $r \to \infty$, combining with the above tells us that $L_s(r)$ has precisely one root $\rho(s) > s$ in $[2, \infty)$ for all $s \geq 6$. □

Now that we know there exists a threshold $\rho(s)$, we can conclude a result about the existence of spanning trees in $\mathcal{G}_{n,r,s}$.

Corollary 4.10. *Suppose $r \geq 3$ and $s \geq 5$ are fixed positive integers such that $r \leq \rho(s)$. Let $Y_{\mathcal{G}}$ be the number of spanning trees in a random graph G from $\mathcal{G}_{n,r,s}$. Then, as $n \to \infty$ (with any obvious restrictions on n), we have*

$$\Pr(Y_{\mathcal{G}} = 0) \to 1.$$

Proof. When $r < \rho(s)$, the base of exponential factor in Theorem 4.1 is less than 1, so $\mathbb{E}Y \to 0$. Since Y is a non-negative random variable, we have

$$\Pr(Y \geq 1) = \sum_{k=1}^{\infty} \Pr(Y = k) \leq \sum_{k=0}^{\infty} k \Pr(Y = k) = \mathbb{E}Y \to 0.$$

By Lemma 2.46, we conclude that $\Pr(Y_{\mathcal{G}} \geq 1) \to 0$, which implies the result. □

It would be useful to know approximately the value of $\rho(s)$. We have the following conjecture.

Conjecture 4.11. *For fixed integer $s \geq 5$, let $\rho(s)$ be the unique real number such that $L(\rho(s), s) = 0$. Then*

$$\frac{e^{s-2}}{s-1} - \frac{s-1}{2} < \rho(s) < \frac{e^{s-2}}{s-1} - \frac{s-3}{2}.$$

We have numerical evidence to support this conjecture. From the proof of Theorem 4.4, we know that $L_s(r)$ is strictly increasing for $r \geq s$. Thus, it suffices to show that

$$L_s\left(\frac{e^{s-2}}{s-1} - \frac{s-1}{2}\right) < 0 \qquad \text{and} \qquad L_s\left(\frac{e^{s-2}}{s-1} - \frac{s-3}{2}\right) > 0.$$

We can compute $L_s\left(\frac{e^{s-2}}{s-1} - \frac{s-1}{2}\right)$ and $L_s\left(\frac{e^{s-2}}{s-1} - \frac{s-3}{2}\right)$ for various values of s numerically:

s	5	50	500	5 000
$L_s\left(\frac{e^{s-2}}{s-1} - \frac{s-1}{2}\right)$	-2.902×10^{-4}	-7.271×10^{22}	-2.637×10^{-217}	-1.245×10^{-2171}
$L_s\left(\frac{e^{s-2}}{s-1} - \frac{s-3}{2}\right)$	3.698×10^{-2}	6.980×10^{-22}	2.627×10^{-217}	1.245×10^{-2171}

The conjecture can be possibly proven through a clever transformation of the variables, similarly to [3, Lemma 6.1].

CHAPTER 5

Effect of short cycles

A generating function is a clothesline on which we hang up a sequence of numbers for display.

– Herbert Wilf, *generatingfunctionology*

In this chapter, we determine an asymptotic expression for $\mathbb{E}(YX_j)$ to verify condition (A2) of Theorem 2.19. In addition, $\mathbb{E}(YX_j)$ gives us the value of ζ_j, allowing us to verifying condition (A3) of Theorem 2.19. The computations are highly technical in nature. We closely follow the structure of the argument for the graph case in [16, Section 3] and show how to adapt the calculations to hypergraphs.

5.1 Joint moment calculation

Let X_1 be the number of 1-cycles in a random partition P from $\mathcal{P}_{n,r,s}$ and for $j \geq 2$, let X_j be the number of loose j-cycles in a random partition P from $\mathcal{P}_{n,r,s}$.

Fix $r \geq 3$ and $s \geq 2$. We perform the calculation for $j \geq 3$. The calculations for $j = 1$ and $j = 2$ follow similarly. For a random partition P from $\mathcal{P}_{n,r,s}$, we can write

$$\mathbb{E}(YX_j) = \sum_{(P_T,P_C)} \Pr(P_T \cup P_C \subseteq P) = \sum_{(P_T,P_C)} \frac{|\{P \in \Omega_{n,r,s} : P_T \cup P_C \subseteq P\}|}{|\Omega_{n,r,s}|},$$

where the sum is over all pairs (P_T, P_C) such that $G(P_T) = T$ for some spanning tree T and $G(P_C) = C$ for some loose j-cycle C.

To perform this count, we condition on the intersection between P_T and P_C. We can construct a map ι, which takes $P_T \cap P_C$ to an element of $\mathcal{I}_j := \{0,1\}^j \setminus \{(1,\ldots,1)\}$. Let $\mathcal{C}_{n,j}$ be the set of s-uniform loose j-cycles on n vertices. Given a fixed $C \in \mathcal{C}_{n,j}$ with a specified starting edge and direction: if the kth edge of the cycle is included in the intersection, then the kth element in the corresponding sequence is one; otherwise it is zero. All sequences in $\{0,1\}^j$ represent possible intersections, except for $(1,\ldots,1)$ because a tree contains no cycles. Thus, we can rewrite the sum as

$$|\Omega_{n,r,s}|\,\mathbb{E}(YX_j) = \sum_{I\in\mathcal{I}_j}\sum_{C\in\mathcal{C}_{n,j}}\sum_{\substack{(P_T,P_C):\\ \iota(P_T\cap P_C)=I,\\ G(P_C)=C\\ G(P_T)\in\mathcal{T}_n}} \left|\{P\in\Omega_{n,r,s} : P_T\cup P_C\subseteq P\}\right|.$$

We evaluate the sum, for a given I, using the following process:

1. Choose some loose j-cycle C on n vertices.
2. Choose a partition P_C which projects to C.
3. Extend this to $P_T \cup P_C$ consistent with I.
4. Partition remaining points arbitrarily.

Then $|\Omega_{n,r,s}|\,\mathbb{E}(YX_j)$ is equal to the number of ways to complete the above process, summed over all $I \in \mathcal{I}_j$. We make combinatorial arguments the count the number of ways of completing each step.

For Step 1, we can choose a sequence of $(s-1)j$ vertices in $(n)_{(s-1)j}$ ways. Recall that, in a loose j-cycle, C-external vertices are those of degree 2, and C-internal vertices are those of degree 1. We divide by $2j$ to adjust for direction and starting point (where a starting point is a C-external vertex) and also divide by another $((s-2)!)^j$ to adjust for the order of the $s-2$ C-internal vertices in each edge. Hence, the number of ways to complete Step 1 is

$$s_1 = \frac{(n)_{(s-1)j}}{2j((s-2)!)^j}.$$

For Step 2, we choose two points for each C-external vertex and one point for each C-internal vertex in the configuration model. So the number of completing Step 2 is

$$s_2 = (r(r-1))^j r^{(s-2)j}.$$

Thus, we have

$$s_1 s_2 \sim \frac{n^{(s-1)j}}{2j}\left(\frac{r^{s-1}(r-1)}{(s-2)!}\right)^j.$$

For Step 3, we construct an irregular configuration model $\mathcal{P}_{n',I}$ from the points that are so far unused. The C-external vertices have $r-2$ points unused and the C-internal vertices have $r-1$ points unused. Recall that I describes disjoint paths on the cycle C. For each path with at least one edge, collect all the unused points and combine them together into an *irregular cell.* If such a path has k edges, it consists of $k+1$ vertices with $r-2$ points unused and $(s-2)k$ vertices with $r-1$ points unused. Thus, the resulting irregular cell has $(k+1)(r-2)+(s-2)k(r-1)=(rs-r-s)k+(r-2)$ points.

For each C-external cell not in $P_T \cap P_C$, we can also form irregular cells with $r-2$ points. Let $u=u(I)$ be the number of paths in $P_T \cap P_C$ plus the number of C-external cells not in T. We shall call these cells the *external irregular cells.* It can be shown that u is equal to the number of zeros in the binary sequence I. For each C-internal vertex not in $P_T \cap P_C$, we can form an *internal irregular cell* with $r-1$ points. There are $(s-2)u$ such cells. To summarise the properties of our irregular configuration model:

- The total number of cells is $n' := n-(s-1)j+(s-1)u$.
- There are $n-(s-1)j$ regular cells with r points each.
- There are $(s-2)u$ internal irregular cells with $r-1$ points each.
- There are u external irregular cells with $(rs-r-s)k+(r-2)$ points, if the cell was collapsed from $k \geq 0$ edges.

The number of ways to complete Step 3 equals to the number of ways of choosing a partition P' that projects to a tree in this irregular configuration model. The projection $T'=G(P')$ of this partition corresponds exactly to a tree T with some subpaths contracted to single vertices.

We perform this count by conditioning on the degree of each cell of T'. Put an arbitrary ordering on the u external irregular cells and the $(s-2)u$ internal irregular cells, and let d_i be the number of points in the ith external irregular cell. For a degree sequence $\boldsymbol{\delta}$, let $|\boldsymbol{\delta}|$ be its degree sum. Define $\mathcal{D}_{\text{irreg}}$ to be the set of possible degree sequences for the irregular cells, and $\mathcal{D}_{\text{reg}}^{(\boldsymbol{\delta'},\boldsymbol{\delta''})}$ be the set of possible degree sequences for the regular cells, given $(\boldsymbol{\delta'},\boldsymbol{\delta''}) \in \mathcal{D}_{\text{irreg}}$:

$$\mathcal{D}_{\text{irreg}} = \{(\boldsymbol{\delta'},\boldsymbol{\delta''}) \in \mathbb{N}^u \times \mathbb{N}^{(s-2)u} : 1 \leq \delta'_i \leq d_i, \quad 1 \leq \delta''_i \leq r-1\},$$

$$\mathcal{D}_{\text{reg}}^{(\boldsymbol{\delta'},\boldsymbol{\delta''})} = \left\{ \boldsymbol{\delta} \in \mathbb{N}^{n-(s-1)j} : \sum_{i=1}^{n-(s-1)j} \delta_i = \frac{s(n'-1)}{s-1} - |\boldsymbol{\delta'}| - |\boldsymbol{\delta''}|, \quad \delta_i \geq 1 \right\}.$$

Let $\mathcal{T}_{n'}(\boldsymbol{\delta}, \boldsymbol{\delta'}, \boldsymbol{\delta''})$ be the trees on n' vertices which have degree sequence consistent with $(\boldsymbol{\delta}, \boldsymbol{\delta'}, \boldsymbol{\delta''})$. Then, the number of ways to complete Step 3 is

$$s_3 = \sum_{(\boldsymbol{\delta'},\boldsymbol{\delta''})\in\mathcal{D}_{\text{irreg}}} \sum_{\boldsymbol{\delta}\in\mathcal{D}_{\text{reg}}^{(\boldsymbol{\delta'},\boldsymbol{\delta''})}} \sum_{T'\in\mathcal{T}_{n'}(\boldsymbol{\delta},\boldsymbol{\delta'},\boldsymbol{\delta''})} \sum_{P'_{T'}:G(P'_{T'})=T'} \left|\{P' \in \mathcal{P}_{n',I} : P'_{T'} \subseteq P'\}\right|.$$

To simplify this, for a given $(T', \boldsymbol{\delta}, \boldsymbol{\delta'}, \boldsymbol{\delta''})$, the number of partitions that project to T' with degree sequence $(\boldsymbol{\delta}, \boldsymbol{\delta'}, \boldsymbol{\delta''})$ is

$$\left(\prod_{i=1}^{u}(d_i)_{\delta'_i}\right)\left(\prod_{i=1}^{(s-2)u}(r-1)_{\delta''_i}\right)\left(\prod_{i=1}^{n-(s-1)j}(r)_{\delta_i}\right).$$

For a given $(\boldsymbol{\delta}, \boldsymbol{\delta'}, \boldsymbol{\delta''})$, by Corollary 3.27, the number of trees in $\mathcal{T}_{n'}(\boldsymbol{\delta}, \boldsymbol{\delta'}, \boldsymbol{\delta''})$ is

$$\frac{s-1}{\left(\prod_{i=1}^{u}(\delta'_i-1)!\right)\left(\prod_{i=1}^{(s-2)u}(\delta''_i-1)!\right)\left(\prod_{i=1}^{n-(s-1)j}(\delta_i-1)!\right)} \times \frac{(n'-2)!}{((s-1)!)^{\frac{n'-1}{s-1}}}.$$

So

$$s_3 = \sum_{(\boldsymbol{\delta'},\boldsymbol{\delta''})\in\mathcal{D}_{\text{irreg}}} A_{(\boldsymbol{\delta'},\boldsymbol{\delta''})}\left(\prod_{i=1}^{u}\frac{(d_i)_{\delta'_i}}{(\delta'_i-1)!}\right)\left(\prod_{i=1}^{(s-2)u}\frac{(r-1)_{\delta''_i}}{(\delta''_i-1)!}\right),$$

where, using Corollary 2.26,

$$\begin{aligned}
A_{(\boldsymbol{\delta'},\boldsymbol{\delta''})} &= \frac{(s-1)(n'-2)!}{((s-1)!)^{\frac{n'-1}{s-1}}} \sum_{\delta\in\mathcal{D}_{\text{reg}}^{(\boldsymbol{\delta'},\boldsymbol{\delta''})}} \prod_{i=1}^{n-(s-1)j}\frac{(r)_{\delta_i}}{(\delta_i-1)!} \\
&= \frac{(s-1)(n'-2)!}{((s-1)!)^{\frac{n'-1}{s-1}}}\left[z^{\frac{s(n'-1)}{s-1}-|\boldsymbol{\delta'}|-|\boldsymbol{\delta''}|}\right]\left(rz(1+z)^{r-1}\right)^{n-(s-1)j} \\
&= \frac{(s-1)(n'-2)!\,r^{n-(s-1)j}}{((s-1)!)^{\frac{n'-1}{s-1}}}\binom{(r-1)(n-(s-1)j)}{\frac{s(n'-1)}{s-1}-|\boldsymbol{\delta'}|-|\boldsymbol{\delta''}|-(n-(s-1)j)} \\
&\sim \frac{(r-1)^{1/2}(s-1)^2(rs-r-s)^{su-\frac{s}{s-1}-|\boldsymbol{\delta'}|-|\boldsymbol{\delta''}|-\frac{1}{2}}}{((s-1)!)^{\frac{(s-1)u-1}{s-1}}n^{(s-1)j-(s-1)u+2}}\left(\frac{(rs-r-s)^{rs-r-s}(s-1)!}{((r-1)(s-1))^{(r-1)(s-1)}r^{s-1}}\right)^j \\
&\quad\times\left(\frac{nr(r-1)^{r-1}(s-1)^{r-1}}{e((s-1)!)^{\frac{1}{s-1}}(rs-r-s)^{\frac{rs-r-s}{s-1}}}\right)^n.
\end{aligned}$$

Lastly, for Step 4, out of the rn points in the original configuration model, $2j$ of them have been used for the C-external vertices, $(s-2)j$ for the C-internal vertices, and $\frac{s(n'-1)}{s-1}$

for the tree. So there are $rn - 2j - (s-2)j - \frac{s(n'-1)}{s-1} = \frac{(rs-r-s)n}{s-1} - s\left(u - \frac{1}{s-1}\right)$ points remaining. Hence, the number of ways to complete Step 4 is, using Corollary 2.26,

$$
\begin{aligned}
s_4 &= f_s\left(\frac{(rs-r-s)n}{s-1} - s\left(u - \frac{1}{s-1}\right)\right) \\
&\sim \frac{\sqrt{s}(s-1)^{(s-2)u-1}((s-1)!)^{u-\frac{1}{s-1}}}{(rs-r-s)^{(s-1)u-1}n^{(s-1)u-1}} \left(\frac{n(rs-r-s)}{e(s-1)((s-1)!)^{\frac{1}{s-1}}}\right)^{\frac{(rs-r-s)}{s}n}.
\end{aligned}
$$

By (2.12), Theorem 4.1 and Corollary 2.26, we also have

$$
|\Omega_{n,r,s}|\,\mathbb{E}Y \sim \frac{\sqrt{s}\sqrt{r-1}(s-1)}{n(rs-r-s)^{\frac{s+1}{2(s-1)}}} \left(\frac{r(r-1)^{r-1}s^{\frac{r(s-1)}{s}}}{(rs-r-s)^{\frac{rs-r-s}{s(s-1)}}((s-2)!)^{\frac{r}{s}}e^{\frac{r(s-1)}{s}}}\right)^n.
$$

Using the fact that

$$
(s-1)^{(s-1)u}(rs-r-s)^{u-|\boldsymbol{\delta'}|-|\boldsymbol{\delta''}|} = \prod_{i=1}^{u} \frac{s-1}{(rs-r-s)^{\delta'_i-1}} \prod_{i=1}^{(s-2)u} \frac{s-1}{(rs-r-s)^{\delta''_i}},
$$

we can put everything together to see that

$$
\begin{aligned}
\frac{\mathbb{E}(YX_j)}{\mathbb{E}Y} &\sim \sum_{I\in\mathcal{I}_j} \frac{s_1s_2s_3s_4}{|\Omega_{n,r,s}|\,\mathbb{E}Y} \\
&\sim \sum_{I\in\mathcal{I}_j} \sum_{(\boldsymbol{\delta'},\boldsymbol{\delta''})\in\mathcal{D}_{\text{irreg}}} \frac{(s-1)^{(s-1)u}(rs-r-s)^{u-|\boldsymbol{\delta'}|-|\boldsymbol{\delta''}|}}{2j} \left(\frac{rs-r-s}{(r-1)(s-1)}\right)^{(rs-r-s)j} \\
&\quad \times \left(\prod_{i=1}^{u} \frac{(d_i)_{\delta'_i}}{(\delta'_i-1)!}\right) \left(\prod_{i=1}^{(s-2)u} \frac{(r-1)_{\delta''_i}}{(\delta''_i-1)!}\right) \\
&= \frac{1}{2j}\left(\frac{rs-r-s}{(r-1)(s-1)}\right)^{(rs-r-s)j} \sum_{I\in\mathcal{I}_j}\sum_{\boldsymbol{\delta'}} \left(\prod_{i=1}^{u} \frac{(d_i)_{\delta'_i}(s-1)}{(\delta'_i-1)!(rs-r-s)^{\delta'_i-1}}\right) \\
&\quad \times \sum_{\boldsymbol{\delta''}} \left(\prod_{i=1}^{(s-2)u} \frac{(r-1)_{\delta''_i}(s-1)}{(\delta''_i-1)!(rs-r-s)^{\delta''_i}}\right) \qquad (5.1)
\end{aligned}
$$

where $\boldsymbol{\delta'}$ sums over the set $\{\boldsymbol{\delta'} \in \mathbb{N}^u : 1 \le \delta'_i \le d_i\}$ and $\boldsymbol{\delta''}$ sums over the set $\{\boldsymbol{\delta''} \in \mathbb{N}^{(s-2)u} : 1 \le \delta''_i \le r-1\}$.

To simplify this, for $k \ge 1$, let $q[k]$ be the number of paths of length k in the intersection encoded by I, and $q[0]$ be the number of C-external vertices not in T (which we may think of as "length-0 paths"). Note that $\sum_{k=0}^{j-1} q[k] = u$, and by adding up all the vertices in the cycle, we have

$$(s-2)u + \sum_{k=0}^{j-1}((s-1)k+1)q[k] = (s-1)j,$$

so combining this, we have

$$\sum_{k=0}^{j-1} k\, q[k] = j - u.$$

Recall that if the ith irregular cell was collapsed from a path of length k, then $d_i = (rs-r-s)k + (r-2)$. Then, we have

$$\begin{aligned}
&\sum_{\boldsymbol{\delta'}} \left(\prod_{i=1}^{u} \frac{(d_i)_{\delta'_i}(s-1)}{(\delta'_i-1)!(rs-r-s)^{\delta'_i-1}} \right) \\
&= \prod_{k=0}^{j-1} \left(\sum_{\ell=1}^{(rs-r-s)k+(r-2)} \frac{((rs-r-s)k+r-2)_\ell(s-1)}{(\ell-1)!(rs-r-s)^{\ell-1}} \right)^{q[k]} \\
&= \prod_{k=0}^{j-1} \left(((rs-r-s)k+r-2)(s-1) \sum_{\ell=1}^{(rs-r-s)k+(r-2)} \frac{\binom{(rs-r-s)k+r-3}{\ell-1}}{(rs-r-s)^{\ell-1}} \right)^{q[k]} \\
&= \prod_{k=0}^{j-1} \left(\left(k + \frac{r-2}{rs-r-s}\right)(rs-r-s)(s-1)\left(1+\frac{1}{rs-r-s}\right)^{(rs-r-s)k+r-3} \right)^{q[k]} \\
&= \left(\frac{(rs-r-s)^2}{r-1}\right)^{u} \left(\frac{(r-1)(s-1)}{rs-r-s}\right)^{(rs-r-s)j-(r-1)(s-2)u} \prod_{k=0}^{j-1} \left(k+\frac{r-2}{rs-r-s}\right)^{q[k]}
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\sum_{\boldsymbol{\delta''}} \left(\prod_{i=1}^{(s-2)u} \frac{(r-1)_{\delta''_i}(s-1)}{(\delta''_i-1)!(rs-r-s)^{\delta''_i}} \right) &= \left(\sum_{\ell=1}^{r-1} \frac{(r-1)_\ell(s-1)}{(\ell-1)!(rs-r-s)^\ell} \right)^{(s-2)u} \\
&= \left(\frac{(r-1)(s-1)}{rs-r-s} \sum_{\ell=1}^{r-1} \binom{r-2}{\ell-1} \frac{1}{(rs-r-s)^{\ell-1}} \right)^{(s-2)u} \\
&= \left(\frac{(r-1)(s-1)}{rs-r-s} \left(1+\frac{1}{rs-r-s}\right)^{r-2} \right)^{(s-2)u} \\
&= \left(\frac{(r-1)(s-1)}{rs-r-s} \right)^{(r-1)(s-2)u}.
\end{aligned}$$

Thus from (5.1)

$$\frac{\mathbb{E}(YX_j)}{\mathbb{E}Y} \sim \frac{1}{2j} \sum_{I \in \mathcal{I}_j} \left(\frac{(rs-r-s)^2}{r-1} \right)^u \prod_{k=0}^{j-1} \left(k + \frac{r-2}{rs-r-s} \right)^{q[k]}.$$

We will compute this sum with the help of a generating function. Because $(1, \ldots, 1) \notin \mathcal{I}_j$, we may identify a particular element in the sequence to be zero. By symmetry, we arbitrarily choose the last. Define the coefficients

$$\Lambda_{j,t} = \sum_{\substack{I \in \mathcal{I}_j: \\ u(I)=t \\ I_j=0}} \mu^t \prod_{k=0}^{j-1} (k+\beta)^{q[k]},$$

where we let

$$\mu = \frac{(rs-r-s)^2}{r-1} \quad \text{and} \quad \beta = \frac{r-2}{rs-r-s}$$

for convenience. Now, $\Lambda_{j,t}$ fixes the number of zeros in I to be t, and assumes $I_j = 0$. There are actually j positions to place a zero, but there are t zeros in I. Hence,

$$\frac{\mathbb{E}(YX_j)}{\mathbb{E}Y} \sim \frac{1}{2j} \sum_{t=1}^{j} \frac{j\Lambda_{j,t}}{t} = \frac{1}{2} \sum_{t=1}^{j} \frac{\Lambda_{j,t}}{t}. \tag{5.2}$$

We now evaluate $\Lambda_{j,t}$. We have $\Lambda_{j,1} = \mu\,(j-1+\beta)$, because the only sequence with $u = 1$ (recalling that u represents the number of zeros in the sequence I) and the last element $I_j = 0$ is $(1, \ldots, 1, 0)$. For $t \geq 2$, the sequence start with k ones followed by a zero, where $0 \leq k \leq j-2$. Ranging over these possibilities, we have the following recurrence relation:

$$\Lambda_{j,t} = \mu \sum_{k=0}^{j-2} (k+\beta)\,\Lambda_{j-k-1,t-1}.$$

To solve this, define the generating function

$$\Lambda(x,y) = \sum_{j\geq 1} \sum_{t \geq 1} \Lambda_{j,t}\, x^j y^t.$$

By changing the order of summation and re-indexing, we have

$$\Lambda(x,y) - \sum_{j\geq 1} \Lambda_{j,1} x^j y = \mu \sum_{j \geq 1} \sum_{t \geq 2} \sum_{k=0}^{j-2} (k+\beta)\,\Lambda_{j-k-1,t-1}\, x^j y^t$$

$$= \mu \sum_{k\geq 0} (k+\beta)\, x^{k+1} y \sum_{j\geq k+2} \sum_{t'\geq 1} \Lambda_{j-k-1,t'}\, x^{j-k-1} y^{t'}$$
$$= \mu \sum_{k\geq 0} (k+\beta)\, x^{k+1} y\, \Lambda(x,y).$$

Thus, recalling that $\Lambda_{j,1} = \mu(j-1+\beta)$, we have

$$\Lambda(x,y) = \mu \left[\sum_{j\geq 1} (j-1+\beta)\, x^j y + \sum_{k\geq 0} (k+\beta)\, x^{k+1} y\, \Lambda(x,y) \right]$$
$$= \mu\, (\Lambda(x,y)+1) \sum_{k\geq 0} (k+\beta)\, x^{k+1} y.$$

Recall that by differentiating both sides of $(1-x)^{-1} = \sum_{k\geq 0} x^k$, we have

$$\sum_{k\geq 0} k x^{k+1} = \frac{x^2}{(1-x)^2}.$$

Hence if we define

$$g(x) = \mu \sum_{k\geq 0} (k+\beta)\, x^{k+1} = \mu \left(\frac{x^2}{(1-x)^2} + \frac{\beta x}{1-x} \right),$$

we have $\Lambda(x,y) = y\, g(x)(\Lambda(x,y)+1)$ and thus

$$\Lambda(x,y) = \frac{g(x)y}{1-g(x)y}.$$

Now, going back to (5.2), we have

$$\frac{\mathbb{E}(YX_j)}{\mathbb{E}Y} \sim \frac{1}{2} \sum_{t=1}^{j} \frac{\Lambda_{j,t}}{t} = \frac{1}{2}[x^j] \sum_{t=1}^{j} \frac{1}{t} [y^{t-1}] \frac{g(x)}{1-g(x)y}.$$

Applying the Taylor expansion of $(1-z)^{-1}$ and $\log(1-z)$, we have

$$\frac{\mathbb{E}(YX_j)}{\mathbb{E}Y} \sim \frac{1}{2}[x^j] \sum_{t=1}^{j} \frac{1}{t} [y^{t-1}] \left(g(x) \sum_{k=0}^{\infty} (g(x)y)^k \right) = \frac{1}{2}[x^j] \sum_{t=1}^{j} \frac{g(x)^t}{t} = -\frac{1}{2}[x^j] \log(1-g(x)).$$

Now,

$$1 - g(x) = \frac{\left(1 - \left(\frac{r}{r-1} - s + 1\right) x\right)(1 - (r-1)(s-1)x)}{(1-x)^2},$$

so

$$\begin{aligned}
&\frac{\mathbb{E}(YX_j)}{\mathbb{E}Y} \\
&\sim \frac{1}{2}[x^j] \left(2\log(1-x) - \log\left(1 - \left(\frac{r}{r-1} - s + 1\right) x\right) - \log\left(1 - (r-1)(s-1)x\right)\right) \\
&= \frac{1}{2}[x^j] \sum_{k=1}^{\infty} \frac{-2x^k + \left(\left(\frac{r}{r-1} - s + 1\right) x\right)^k + ((r-1)(s-1)x)^k}{k} \\
&= \frac{\left(\frac{r}{r-1} - s + 1\right)^j + (r-1)^j(s-1)^j - 2}{2j}.
\end{aligned}$$

We have proven the following lemma.

Lemma 5.1. *Let $r \geq 3$ and $s \geq 2$ be fixed integers. For any fixed integer $j \geq 1$,*

$$\frac{\mathbb{E}(YX_j)}{\mathbb{E}Y} \sim \lambda_j(1 + \zeta_j),$$

as $n \to \infty$ (with any necessary restrictions), where

$$\lambda_j = \frac{(r-1)^j(s-1)^j}{2j} \quad \textit{and} \quad \zeta_j = \frac{\left(\frac{r}{r-1} - s + 1\right)^j - 2}{(r-1)^j(s-1)^j}.$$

Similar calculations may be performed for a general m-tuple to prove that

$$\frac{\mathbb{E}[Y(X_1)_{x_1} \ldots (X_m)_{x_m}]}{\mathbb{E}Y} \to \prod_{j=1}^{m} (\lambda_j(1 + \zeta_j))^{x_i}.$$

Roughly speaking, the most asymptotically significant contribution to this value is when the cycles are disjoint, and then the contribution from each cycle is essentially independent and the same as when just considering one cycle. So we conclude that condition (A2′) (and hence (A2)) holds.

Now that we know the value of ζ_j, we can show that condition (A3) holds.

Lemma 5.2. *Let $r \geq 3$ and $s \geq 2$ be fixed integers where $r > \rho(s)$. Then we have*

$$\exp\left(\sum_{j=1}^{\infty} \lambda_j \zeta_j^2\right) = \frac{r^2\sqrt{s-1}}{\sqrt{(r^2 - rs + r + s - 1)\,(rs - r - s)(r-1)}} < \infty.$$

Proof. Using the Taylor expansion of $\log(1-z)$,

$$\begin{aligned}
\sum_{j=1}^{\infty} \lambda_j \zeta_j^2 &= \frac{1}{2}\sum_{j=1}^{\infty} \frac{1}{j}\left(\left(\frac{\left(\frac{r}{r-1} - s + 1\right)^2}{(r-1)(s-1)}\right)^j - 4\left(\frac{\frac{r}{r-1} - s + 1}{(r-1)(s-1)}\right)^j + 4\left(\frac{1}{(r-1)(s-1)}\right)^j\right) \\
&= -2\log\left(1 - \frac{1}{(r-1)(s-1)}\right) + 2\log\left(1 - \frac{\frac{r}{r-1} - s + 1}{(r-1)(s-1)}\right) \\
&\quad - \frac{1}{2}\log\left(1 - \frac{\left(\frac{r}{r-1} - s + 1\right)^2}{(r-1)(s-1)}\right) \\
&= -2\log\left(\frac{rs - r - s}{(r-1)(s-1)}\right) + 2\log\left(\frac{r(rs - r - s)}{(r-1)^2(s-1)}\right) \\
&\quad - \frac{1}{2}\log\left(\frac{(rs - r - s)(r^2 - rs + r + s - 1)}{(r-1)^3(s-1)}\right)
\end{aligned}$$

Taking the exponential of both sides establishes the result. □

CHAPTER 6

Second moment

All know the way, few actually walk it.

– Buddha

So far, conditions (A1)–(A3) of Theorem 2.19 have been verified. In this chapter, we derive an expression for the second moment in the graph case, proving the conjecture of Greenhill et al. in [16]. We also make progress towards finding the expression for the second moment in the r-regular s-uniform hypergraph case. As with the joint moment, the calculations in this chapter are fairly technical.

6.1 Second moment in the graph case

For this section, we return to the graph case ($s = 2$). As the case $r = 3$ has been studied sufficiently in [16], we assume that $r \geq 4$ throughout this section.

In [16], Greenhill et al. concluded that, for $r \geq 4$,

$$|\Omega_{n,r}| \, \mathbb{E}Y^2 = \frac{n! \, r^n \, ((r-2)n)!}{2^{(r/2-1)n+2}} \sum_{b=1}^{n} \frac{2^b}{b! \, ((r/2-1)n - b + 2)!} \sum_{\substack{\nu_1+\cdots+\nu_b=n \\ \nu_i \geq 1}} \prod_{j=1}^{b} \binom{(r-1)\nu_i}{\nu_i}.$$

Note that for the $r = 3$ case, Greenhill et al. were able to evaluate a similar sum using Cauchy's coefficient formula and complex integration. Here, we take a different approach and evaluate the sum for any fixed $r \geq 4$.

Using the substitution $k_i = \nu_i - 1$, we can rewrite the inner sum as

$$\sum_{\substack{k_1+\cdots+k_b=n-b \\ k_i \ge 0}} \prod_{j=1}^{b} \binom{(r-1)k_i + r - 1}{k_i + 1} = (r-1)^b \sum_{\substack{k_1+\cdots+k_b=n-b \\ k_i \ge 0}} \prod_{j=1}^{b} \binom{(r-1)k_i + r - 2}{k_i}.$$

This allows us to apply a generalised form of Jensen's identity [9], given below.

Lemma 6.1 (Generalised Jensen's identity)**.** *For $b \ge 2$, we have*

$$\sum_{\substack{k_1+\cdots+k_b=m \\ k_i \ge 0}} \prod_{i=1}^{b} \binom{x_i + k_i z}{k_i} = \sum_{k=0}^{m} \binom{k+b-2}{k} \binom{x_1 + \cdots + x_b + mz - k}{m-k} z^k.$$

We apply this with $x_i = r-2$, $z = r-1$ and $m = n-b$, writing the $b = 1$ term separately:

$$\begin{aligned}|\Omega_{n,r}|\,\mathbb{E}Y^2 = \frac{n!\,r^n((r-2)n)!}{2^{(r/2-1)n+2}} \Bigg(&\frac{2}{((r/2-1)n+1)!} \binom{(r-1)n}{n} \\ &+ \sum_{b=2}^{n} \frac{2^b(r-1)^b}{b!\,((r/2-1)n-b+2)!} \sum_{k=0}^{n-b} \binom{k+b-2}{k} \binom{(r-1)n-k-b}{n-k-b} (r-1)^k \Bigg)\end{aligned}$$

By taking $s = 2$ in Theorem 4.1, we have

$$|\Omega_{n,r}|\,\mathbb{E}Y = \frac{r^n((r-1)n)!}{((r/2-1)n+1)!\,2^{(r/2-1)n+1}}.$$

Hence, dividing by this expression, we can write

$$\begin{aligned}\frac{\mathbb{E}Y^2}{\mathbb{E}Y} &= 1 + \sum_{b=2}^{n} \sum_{k=0}^{n-b} \frac{2^{b-1}(r-1)^{b+k}\,n!\,((r/2-1)n+1)!\,(k+b-2)!\,((r-1)n-k-b)!}{((r-1)n)!\,b!\,((r/2-1)n-b+2)!\,k!\,(b-2)!\,(n-k-b)!} \\ &= 1 + \sum_{\substack{(k,b)\in\mathbb{N}^2 \\ k+b \le n}} a_n(k,b),\end{aligned}$$

where

$$a_n(k,b) = \begin{cases} 0 & \text{for } b = 0, 1, \\ \dfrac{2^{b-1}(r-1)^{b+k}\,n!\,((\frac{r}{2}-1)n+1)!\,(k+b-2)!\,((r-1)n-k-b)!}{((r-1)n)!\,b!\,((\frac{r}{2}-1)n-b+2)!\,k!\,(b-2)!\,(n-k-b)!} & \text{for } 2 \le b \le n. \end{cases}$$

Now, we apply Laplace's method via Lemma 2.27 to sum this expression asymptotically.

Define the scaled domain with scaled variables $\alpha = k/n$ and $\beta = b/n$:

$$K = \{(\alpha, \beta) \in \mathbb{R}^2 : \alpha, \beta \geq 0, \quad \alpha + \beta \leq 1\}. \tag{6.1}$$

Note that $\mathbb{Z}^2 \cap nK = \{(k, b) \in \mathbb{N}^2 : k + b \leq n\}$. Define the function $\phi : K \to \mathbb{R}$ by

$$\begin{aligned}\phi(\alpha, \beta) = (\alpha + \beta) \log(r-1) + g(\alpha + \beta) + g(r - 1 - \alpha - \beta) - 2g(\beta) \\ - \frac{1}{2} g(r - 2 - 2\beta) - g(\alpha) - g(1 - \alpha - \beta),\end{aligned} \tag{6.2}$$

where $g(x) = x \log x$ for $x > 0$ and $g(0) = 0$. We will need to find the unique global maximum of ϕ in the domain K. We state the following theorem, adapted from [17, Theorem 2.14].

Theorem 6.2. *Let $K \subset \mathbb{R}^n$ be a compact convex set. Let $f : K \to \mathbb{R}$ be a function with negative definite Hessian on $\operatorname{int}(K)$. If $\boldsymbol{x_0} \in \operatorname{int}(K)$ is a critical point, then the unique global maximum of f over K is at $\boldsymbol{x_0}$.*

The original statement of this theorem only concerned functions on open sets, but continuity of the function also ensures that the global maximum cannot lie on the boundary.

We can now prove the following lemma.

Lemma 6.3. *Suppose that $r \geq 4$. The function ϕ has a unique global maximum over the domain K at the point (α_0, β_0), where*

$$\alpha_0 = \frac{1}{r}, \qquad \beta_0 = \frac{r-2}{r}.$$

The maximum value of ϕ over this domain equals

$$\phi(\alpha_0, \beta_0) = 2(r-1)\log(r-1) - (r-2)\log(r-2) - \frac{r-2}{2}\log r.$$

Let H_0 be the Hessian of ϕ evaluated at the point (α_0, β_0). Then H_0 is strictly negative definite and

$$\det(-H_0) = \frac{r^3(r^2 - r + 1)}{(r-2)(r-1)^2}.$$

Proof. Observe that K is a compact, convex set in $[0, 1]^2$ and ϕ is a continuous function on K. Therefore, ϕ attains its maximum value at least once.

We compute the first partial derivatives:

$$\begin{aligned}\frac{\partial\phi}{\partial\alpha} &= \log(\alpha+\beta)+\log(1-\alpha-\beta)-\log(r-1-\alpha-\beta)+\log(r-1)-\log\alpha,\\ \frac{\partial\phi}{\partial\beta} &= \log(\alpha+\beta)+\log(1-\alpha-\beta)-\log(r-1-\alpha-\beta)+\log(r-1)\\ &\qquad -2\log(\beta)+\log(r-2-2\beta).\end{aligned}$$

We can verify by substitution that there is a critical point in the interior of K at

$$\alpha_0=\frac{1}{r}\qquad \beta_0=\frac{r-2}{r}.$$

Now let $H(\alpha,\beta)$ be the Hessian evaluated at the point $(\alpha,\beta)\in \operatorname{int}(K)$. We find that

$$\det(H(\alpha,\beta))=\frac{(r-2)\left(-\alpha^2\beta-2\alpha\beta^2-2\beta^2-\beta^3+\alpha^2(2r-4)+\alpha\beta(2r-6)+\beta(r-1)\right)}{\alpha\beta(1-\alpha-\beta)(\alpha+\beta)(r-2-2\beta)(r-1-\alpha-\beta)}.$$

which is positive if and only if

$$\alpha^2(2r-4-\beta)+\beta(r-1-2\beta-\beta^2)+2\alpha\beta(r-3-\beta)>0.$$

This is true in $\operatorname{int}(K)$, as each term is positive since $0<\alpha,\beta<1$ and $r\geq 4$. Furthermore, we have

$$\operatorname{tr}(H(\alpha,\beta))=-\left(\frac{2}{1-\alpha-\beta}+\frac{2\alpha}{\beta(\alpha+\beta)}+\frac{1}{\alpha}+\frac{2(1-\alpha+\beta)}{(r-2-2\beta)(r-1-\alpha-\beta)}\right)<0.$$

Thus, the eigenvalues of $H(\alpha,\beta)$ are strictly negative, so ϕ has a negative definite Hessian on $\operatorname{int}(K)$, and thus by Theorem 6.2, the unique global maximum over K must be at (α_0,β_0). The values of $\phi(\alpha_0,\beta_0)$ and $\det(-H_0)$ follow by direct substitution. □

With this information, we can establish the following asymptotic expression for the second moment of Y.

Theorem 6.4. *Suppose $r\geq 3$ is a fixed integer. Then as $n\to\infty$, we have*

$$\frac{\mathbb{E}Y^2}{(\mathbb{E}Y)^2}\sim\frac{r^2}{\sqrt{(r-1)(r-2)(r^2-r+1)}}.$$

Proof. We apply Lemma 2.27 to compute the sum of $a_n(k,b)$ over $\mathbb{Z}^2\cap nK$. The first six conditions of Lemma 2.27 hold:

(i) Let $\mathcal{L} = \mathbb{Z}^2$, a lattice with rank $m = 2$ and $\det(\mathcal{L}) = 1$.

(ii) The domain K defined in (6.1) is compact, convex and has a non-empty interior.

(iii) The function $\phi : K \to \mathbb{R}$ defined in (6.2) is continuous with a unique maximum at an interior point $\left(\frac{1}{r}, \frac{r-2}{r}\right)$ by Lemma 6.3.

(iv) The function $\phi : K \to \mathbb{R}$ is twice differentiable in the interior of K, with a strictly negative definite Hessian, by Lemma 6.3.

(v) Let K_1 be an open ball centred at $\left(\frac{1}{r}, \frac{r-2}{r}\right)$ of sufficiently small radius (say $1/2r$), ensuring that $K_1 \subset K$. The function $\psi : K_1 \to \mathbb{R}$ defined by

$$\psi(\alpha, \beta) = \frac{(r-1-\alpha-\beta)^{1/2}\beta}{(\alpha+\beta)^{3/2}(r-2-2\beta)^{5/2}\alpha^{1/2}(1-\alpha-\beta)^{1/2}}.$$

is a continuous function on K_1 with $\psi(\alpha_0, \beta_0) = \frac{r^{7/2}}{(r-2)^4\sqrt{r-1}} > 0$.

(vi) Let ℓ_n be the zero vector for each n.

It remains to prove that condition (vii) holds. Define

$$b_n = \frac{(r-2)^{3/2}}{2\pi n^2 (r-1)^{1/2}} \left(\frac{(r-2)^{r/2-1}}{(r-1)^{r-1}} \right)^n,$$

and introduce the scaled variables

$$\alpha = k/n, \qquad \beta = b/n.$$

Observe that (vii) trivially holds when $b = 0, 1$. For $b \geq 2$, we use Lemma 2.25 and Corollary 2.26 (omitting the details) to see that

$$a_n(k, b) = O(b_n e^{n\phi(\alpha,\beta)+o(n)}) \quad \text{and} \quad a_n(\ell) = b_n(\psi(\ell/n) + o(1))e^{n\phi(\ell/n)}.$$

Hence, we can apply Lemma 2.27 to see that

$$\frac{\mathbb{E}Y^2}{\mathbb{E}Y} \sim \frac{2\pi\psi(\alpha_0, \beta_0)}{\det(\mathcal{L})\det(-H)^{1/2}} b_n\, n\, e^{n\phi(\alpha_0,\beta_0)} = \frac{r^2}{n(r-2)^2\sqrt{r^2-r+1}} \left(\frac{(r-1)^{r-1}}{(r^2-2r)^{r/2-1}} \right)^n.$$

Thus, dividing by the expression for $\mathbb{E}Y$ in Remark 4.2 gives us our result. □

Proving this theorem verifies the final condition of the small subgraph conditioning method, and hence we have verified the conjecture of Greenhill et al., that is, Theorem 2.22 holds true.

6.2 Second moment in the hypergraph case

Assume now that $r \geq 3$ and $s \geq 3$ (although we remark the discussion in the section holds for $r \geq 4$ and $s = 2$). We can write

$$|\Omega_{n,r,s}| \, \mathbb{E}Y^2 = \sum_{(P_{T_1}, P_{T_2})} |\{P \in \Omega_{n,r,s} : P_{T_1} \cup P_{T_2} \subseteq P\}|$$

where the sum is over all pairs (P_{T_1}, P_{T_2}) such that $G(P_{T_1}) = T_2$ and $G(P_{T_2}) = T_2$, for some spanning trees T_1, T_2.

We perform this count by conditioning on the intersection between P_{T_1} and P_{T_2}, which will correspond to a union of disjoint trees. Let $b \in \left\{1 + (s-1)t : t = 0, 1, \ldots, \frac{n-1}{s-1}\right\}$ be the number of connected components in this intersection (we can show b must be of this form by adding up the number of vertices in each connected component). We break up the process into the following steps:

1. Choose a partition $\nu = (\nu_1, \ldots, \nu_b)$ of n, where $\nu_i > 0$, $s - 1 \mid \nu_i - 1$ and $\sum_{i=1}^{b} \nu_i = n$. Here, ν_i represents the number of vertices in the ith connected component. (Later, we will divide by $b!$ to account for the assumption that the connected components are labelled).
2. Choose a partition of the n vertices into b groups, where the size of the ith group is ν_i.
3. In each group, choose a spanning tree on that group and a subpartition that projects to that tree.

We then collapse the unused points in each group to an irregular cell. The ith irregular cell will have $r\nu_i - \frac{s(\nu_i - 1)}{s-1} = \frac{rs-r-s}{s-1}\nu_i + \frac{s}{s-1}$ points. In this irregular configuration model, we wish to partition two part-disjoint spanning trees T_1' and T_2', which will extend to T_1 and T_2.

4. Choose $\boldsymbol{\delta}^{(1)}, \boldsymbol{\delta}^{(2)} \in \mathbb{N}^b$, the degree sequence of T_1' and T_2' respectively, such that, for all i,

$$\delta_i^{(1)}, \delta_i^{(2)} \geq 1, \quad \sum_{i=1}^{b} \delta_i^{(1)} = \sum_{i=1}^{b} \delta_i^{(2)} = \frac{s(b-1)}{s-1}, \quad \text{and} \quad \delta_i^{(1)} + \delta_i^{(2)} \leq \frac{rs - r - s}{s-1}\nu_i + \frac{s}{s-1}.$$

5. Choose trees T_1', T_2' consistent with $\boldsymbol{\delta}^{(1)}$ and $\boldsymbol{\delta}^{(2)}$.
6. Choose P_{T_1} and P_{T_2} such that there are no parts in common.
7. Partition remaining points.

Then $|\Omega_{n,r,s}|\,\mathbb{E}Y^2$ is equal to the number of ways to complete the above process, summed over all b.

The number of ways to complete Step 2 is

$$s_2 = \binom{n}{\nu_1,\ldots,\nu_b} = n!\prod_{i=1}^{b}\frac{1}{\nu_i!}.$$

By (4.1), the number of ways to complete Step 3 is

$$s_3 = \frac{(s-1)^b\, r^n}{((s-1)!)^{\frac{n-b}{s-1}}}\prod_{i=1}^{b}(\nu_i-2)!\binom{(r-1)\nu_i}{\frac{\nu_i-s}{s-1}} = \frac{(s-1)^b\, r^n}{((s-1)!)^{\frac{n-b}{s-1}}}\prod_{i=1}^{b}\frac{(\nu_i-2)!((r-1)\nu_i)!}{\left(\frac{\nu_i-s}{s-1}\right)!\left(\frac{rs-r-s}{s-1}\nu_i+\frac{s}{s-1}\right)!}.$$

By Corollary 3.27, the number of ways to complete Step 5 is

$$s_5 = \binom{\frac{b-1}{s-1}-1}{\delta_1^{(1)}-1,\ldots,\delta_b^{(1)}-1}\binom{\frac{b-1}{s-1}-1}{\delta_1^{(2)}-1,\ldots,\delta_b^{(2)}-1}\left(\frac{(b-1)!}{\left(\frac{b-1}{s-1}\right)!((s-1)!)^{\frac{b-1}{s-1}}}\right)^2.$$

As each cell in this irregular configuration model has $\frac{rs-r-s}{s-1}\nu_i+\frac{s}{s-1}$ points, the number of ways to complete Step 6 is

$$s_6 = \prod_{i=1}^{b}\left(\frac{rs-r-s}{s-1}\nu_i+\frac{s}{s-1}\right)_{\delta_i^{(1)}+\delta_i^{(2)}} = \prod_{i=1}^{b}\frac{\left(\frac{rs-r-s}{s-1}\nu_i+\frac{s}{s-1}\right)!}{\left(\frac{rs-r-s}{s-1}\nu_i+\frac{s}{s-1}-\delta_i^{(1)}-\delta_i^{(2)}\right)!}.$$

There are

$$\sum_{i=1}^{b} r\nu_i - \frac{s(\nu_i-1)}{s-1} - \delta_i^{(1)} - \delta_i^{(2)} = \frac{(rs-r-s)n}{s-1} - \frac{s(b-2)}{s-1}$$

points remaining, so the number of ways to complete Step 7 is

$$s_7 = f_s\left(\frac{(rs-r-s)n}{s-1} - \frac{s(b-2)}{s-1}\right) = \frac{\left(\frac{(rs-r-s)n}{s-1}-\frac{s(b-2)}{s-1}\right)!}{\left(\frac{rs-r-s}{s(s-1)}n-\frac{b-2}{s-1}\right)!(s!)^{\frac{rs-r-s}{s(s-1)}n-\frac{b-2}{s-1}}}.$$

It becomes convenient to work with non-negative variables so we let

$$\eta_i^{(1)} = \delta_i^{(1)}-1,\quad \eta_i^{(2)} = \delta_i^{(2)}-1,\quad \eta_i^{(3)} = \frac{rs-r-s}{s-1}\nu_i - \frac{s-2}{s-1} - \eta_i^{(1)} - \eta_i^{(2)},$$

for $i = 1, \ldots, b$. Let

$$\mathcal{S}_1(b) = \left\{ \nu \in \left\{ 1 + (s-1)t : t = 0, 1, \ldots, \frac{n-1}{s-1} \right\}^b : \sum_{i=1}^{b} \nu_i = n \right\}$$

be the set of possible sequences ν from Step 1, and let

$$\begin{aligned}\mathcal{S}_4(\nu) = \Bigg\{ (\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\eta}^{(3)}) \in (\mathbb{N}^b)^3 : \quad & \eta_i^{(1)} + \eta_i^{(2)} + \eta_i^{(3)} = \frac{rs - r - s}{s-1}\nu_i - \frac{s-2}{s-1}, \\ & \sum_{i=1}^{b} \eta_i^{(1)} = \sum_{i=1}^{b} \eta_i^{(2)} = \frac{b-s}{s-1} \Bigg\}\end{aligned}$$

be the set of sequences arising from Step 4. Combining everything, and dividing by $b!$ as promised earlier, we have

$$\begin{aligned}|\Omega_{n,r,s}|\, \mathbb{E}Y^2 &= \sum_{\substack{b=1 \\ s-1|b-1}}^{n} \frac{1}{b!} \sum_{\nu \in \mathcal{S}_1(b)} s_2 s_3 \sum_{(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\eta}^{(3)}) \in \mathcal{S}_4(\nu)} s_5 s_6 s_7 \\ &= \sum_{\substack{b=1 \\ s-1|b-1}}^{n} \frac{n!\, r^n (s-1)^b ((b-1)!)^2}{b!((s-1)!)^{\frac{n+b-2}{s-1}} \left(\frac{rs-r-s}{s(s-1)}n - \frac{b-2}{s-1}\right)! \left(\left(\frac{b-1}{s-1}\right)!\right)^2 (s!)^{\frac{rs-r-s}{s(s-1)}n - \frac{b-2}{s-1}}} \\ &\quad \times \sum_{\nu \in \mathcal{S}_1(b)} \left(\prod_{i=1}^{b} \frac{((r-1)\nu_i)!}{\nu_i(\nu_i - 1)\left(\frac{\nu_i - s}{s-1}\right)!} \right) \\ &\quad \times \sum_{(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\eta}^{(3)}) \in \mathcal{S}_4(\nu)} \binom{\frac{b-1}{s-1} - 1}{\eta_1^{(1)}, \ldots, \eta_b^{(1)}} \binom{\frac{b-1}{s-1} - 1}{\eta_1^{(2)}, \ldots, \eta_b^{(2)}} \binom{\frac{(rs-r-s)n}{s-1} - \frac{s(b-2)}{s-1}}{\eta_1^{(3)}, \ldots, \eta_b^{(3)}}.\end{aligned}$$

Now we compute the sum over $\mathcal{S}_4(\nu)$ through the use of generating functions:

$$\begin{aligned}&\sum_{(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\eta}^{(3)}) \in \mathcal{S}_4(\nu)} \binom{\frac{b-1}{s-1} - 1}{\eta_1^{(1)}, \ldots, \eta_b^{(1)}} \binom{\frac{b-1}{s-1} - 1}{\eta_1^{(2)}, \ldots, \eta_b^{(2)}} \binom{\frac{(rs-r-s)n}{s-1} - \frac{s(b-2)}{s-1}}{\eta_1^{(3)}, \ldots, \eta_b^{(3)}} \\ &= \sum_{(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\eta}^{(3)}) \in \mathcal{S}_4(\nu)} [z_1^{\eta_1^{(1)}} \cdots z_b^{\eta_b^{(1)}}] \left(\sum_{i=1}^{b} z_i \right)^{\frac{b-s}{s-1}} [z_1^{\eta_1^{(2)}} \cdots z_b^{\eta_b^{(2)}}] \left(\sum_{i=1}^{b} z_i \right)^{\frac{b-s}{s-1}} \\ &\qquad [z_1^{\eta_1^{(3)}} \cdots z_b^{\eta_b^{(3)}}] \left(\sum_{i=1}^{b} z_i \right)^{\frac{(rs-r-s)n}{s-1} - \frac{s(b-2)}{s-1}}\end{aligned}$$

$$= [z_1^{\frac{rs-r-s}{s-1}\nu_1 - \frac{s-2}{s-1}} \cdots z_b^{\frac{rs-r-s}{s-1}\nu_b - \frac{s-2}{s-1}}] \left(\sum_{i=1}^{b} z_i\right)^{\frac{(rs-r-s)n}{s-1} - \frac{b(s-2)}{s-1}}$$
$$= \binom{\frac{(rs-r-s)n}{s-1} - \frac{b(s-2)}{s-1}}{\frac{(rs-r-s)\nu_1}{s-1} - \frac{s-2}{s-1}, \ldots, \frac{(rs-r-s)\nu_b}{s-1} - \frac{s-2}{s-1}}.$$

So

$$|\Omega_{n,r,s}|\,\mathbb{E}Y^2 = \sum_{\substack{b=1 \\ s-1|b-1}}^{n} \frac{n!\, r^n (r-1)^b ((b-1)!)^2 \left(\frac{(rs-r-s)n}{s-1} - \frac{b(s-2)}{s-1}\right)!}{b!((s-1)!)^{\frac{n+b-2}{s-1}} \left(\frac{rs-r-s}{s(s-1)}n - \frac{b-2}{s-1}\right)! \left(\left(\frac{b-1}{s-1}\right)!\right)^2 (s!)^{\frac{rs-r-s}{s(s-1)}n - \frac{b-2}{s-1}}}$$
$$\times \sum_{\nu \in \mathcal{S}_1(b)} \prod_{i=1}^{b} \binom{(r-1)\nu_i - 1}{\frac{\nu_i - 1}{s-1}}$$

Note that the summand is equal to $|\Omega_{n,r,s}|\,\mathbb{E}Y$ when $b = 1$. For $b \geq 2$, we let $k_i = \frac{\nu_i - 1}{s-1}$ and use Lemma 6.1, to see that

$$\sum_{\nu \in \mathcal{S}_1(b)} \prod_{i=1}^{b} \binom{(r-1)\nu_i - 1}{\frac{\nu_i - 1}{s-1}} = \sum_{\substack{k_1 + \cdots + k_b = \frac{n-b}{s-1} \\ k_i \geq 0}} \prod_{i=1}^{b} \binom{(r-1)(s-1)k_i + (r-2)}{k_i}$$
$$= \sum_{k=0}^{\frac{n-b}{s-1}} \binom{k+b-2}{k} \binom{(r-1)n - b - k}{\frac{n-b}{s-1} - k} (r-1)^k.$$

Define

$$K = \{(\alpha, \beta) : \alpha, \beta \geq 0, \quad (s-1)\alpha + \beta \leq 1\}, \tag{6.3}$$
$$\mathcal{L} = \mathbb{Z} \times (s-1)\mathbb{Z}. \tag{6.4}$$

Thus, dividing through by the expression in (2.12), we have

$$\mathbb{E}Y^2 = \mathbb{E}Y + \sum_{(k,b) \in (\mathcal{L} + (0,1)) \cap nK} a_n(k, b),$$

where

$$a_n(k,b) = \begin{cases} 0 & \text{for } b \le 1, \\ \dfrac{r^n (b-1)(r-1)^{k+b}\, s^{\frac{n+b-2}{s-1}}\, (k+b-2)!\,((r-1)n-k-b)!\,(rn/s)!\,n!}{b\,k!\,\left(\left(\frac{b-1}{s-1}\right)!\right)^2 \left(\frac{rs-r-s}{s(s-1)}n - \frac{b-2}{s-1}\right)!\,\left(\frac{n-(s-1)k-b}{s-1}\right)!\,(rn)!} & \text{otherwise.} \end{cases}$$

We now wish to apply Laplace's method to compute the asymptotic summation of this expression.

Define

$$\begin{aligned} b_n &= \frac{(s-1)^2}{2\pi n^3} \left(\frac{(s-1)^{r/s}}{r^{\frac{rs-r-s}{s}}} \right)^n \\ \psi(\alpha,\beta) &= \frac{(r-1-\alpha-\beta)^{1/2}}{(\alpha+\beta)^{3/2}(rs-r-s(1+\beta))^{\frac{1}{2}+\frac{2}{s-1}}\beta^{1-\frac{2}{s-1}}\alpha^{1/2}(1-\beta-(s-1)\alpha)^{1/2}} \\ \phi(\alpha,\beta) &= (\alpha+\beta)\log(r-1) + g(\alpha+\beta) + g(r-1-\alpha-\beta) - \frac{2}{s-1}g(\beta) - g(\alpha) \\ &\qquad - \frac{1}{s(s-1)}g(rs-r-s-s\beta) - \frac{1}{s-1}g(1-(s-1)\alpha-\beta), \end{aligned}$$

where $g(x) = x\log x$ for $x > 0$ and $g(0) = 0$. We have the following conjecture about the function ϕ.

Conjecture 6.5. *For fixed integers $r \ge 3$, $s \ge 3$, where $r > \rho(s)$, the unique global maximum of ϕ over K is*

$$\alpha_0 = \frac{1}{r(s-1)}, \quad \beta_0 = \frac{rs-r-s}{r(s-1)}.$$

The maximum value of ϕ over this domain equals

$$\phi(\alpha_0,\beta_0) = 2(r-1)\log(r-1) - \frac{2(rs-r-s)}{s(s-1)}\log(rs-r-s) + \frac{r}{s}\log(s-1) - \frac{rs-r-s}{s}\log r.$$

Let H_0 be the Hessian of ϕ evaluated at the point (α_0,β_0). Then H_0 is strictly negative definite and

$$\det(-H_0) = \frac{r^3(s-1)^2\left(r^2-rs+r+s-1\right)}{(r-1)^2(rs-r-s)}.$$

We can verify that (α_0,β_0) is at least a local maximum as it satisfies $\frac{\partial\phi}{\partial\alpha} = \frac{\partial\phi}{\partial\beta} = 0$ with a negative definite Hessian H_0, but showing uniqueness is harder. For the cases $s = 2$ and $s = 3$, we are able to solve $\frac{\partial\phi}{\partial\alpha} = \frac{\partial\phi}{\partial\beta} = 0$, but for $s \ge 4$, it is more difficult. The following lemma covers the case $s = 3$.

Lemma 6.6. *For fixed integers $r \geq 3$ and $s = 3$, the unique global maximum of ϕ over K is*

$$\alpha_0 = \frac{1}{2r}, \qquad \beta_0 = \frac{2r-3}{2r}.$$

The maximum value of ϕ over this domain equals

$$\phi(\alpha_0, \beta_0) = 2(r-1)\log(r-1) - \frac{2r-3}{3}\log(2r-3) + \frac{r}{3}\log 2 - \frac{2r-3}{3}\log r.$$

Let H_0 be the Hessian of ϕ evaluated at the point (α_0, β_0). Then H_0 is strictly negative definite and

$$\det(-H_0) = \frac{4r^3\left(r^2 - 2r + 2\right)}{(r-1)^2(2r-3)}.$$

Proof. By substitution into $\frac{\partial \phi}{\partial \alpha}$ and $\frac{\partial \phi}{\partial \beta}$, we can see that we have a critical point of ϕ in the interior of K at

$$\alpha_0 = \frac{1}{2r}, \qquad \beta_0 = \frac{2r-3}{2r}.$$

Let $H(\alpha, \beta)$ be the Hessian of ϕ at an interior point (α, β). We find that

$$\det(H(\alpha,\beta)) = \frac{-(2r-3)\left(-\alpha^2\beta - 2\alpha\beta^2 + \beta^2\left(\frac{r^2-3r+3}{2r-3}\right) + \alpha^2(2r-3) + \alpha\beta(2r-4) - \beta^3\right)}{\alpha\beta(\alpha+\beta)(1-2\alpha-\beta)(2r-3-3\beta)(r-1-\alpha-\beta)},$$

which is positive if and only if

$$\beta^2\left(\frac{r^2-3r+3}{2r-3} - \beta\right) + \alpha^2(2r-3) + \alpha\beta(2r-4-\alpha-2\beta) > 0.$$

This is true as each term is positive in the interior of K, as $r \geq 3$ and $\alpha, \beta > 0$ and $2\alpha + \beta < 1$. Furthermore, we have

$$\operatorname{tr}(H(\alpha,\beta)) = -\left(\frac{\beta^2}{\alpha\beta(\alpha+\beta)} + \frac{3\alpha - \beta + 5r - 9}{2(2\alpha+\beta-1)(\alpha+\beta-r+1)} + \frac{3}{2(2r-3-3\beta)}\right) < 0.$$

Hence the eigenvalues of $H(\alpha, \beta)$ are strictly negative, so ϕ has a negative definite Hessian on $\operatorname{int}(K)$, and thus by Theorem 6.2, the unique global maximum over K must be at (α_0, β_0). □

If we assume Conjecture 6.5 is true, then we have the following lemma.

Lemma 6.7. *Suppose $r \geq 3$, $s \geq 2$ are fixed integers. Then as $n \to \infty$, we have*

$$\frac{\mathbb{E}Y^2}{(\mathbb{E}Y)^2} \sim \frac{r^2\sqrt{s-1}}{\sqrt{(r^2-rs+r+s-1)\,(rs-r-s)(r-1)}}.$$

Proof. We apply Lemma 2.27 to compute the sum. The first six conditions of the lemma hold: Now the conditions of Lemma 2.27 hold:

(i) We defined $\mathcal{L} = \mathbb{Z} \times (s-1)\mathbb{Z}$, a lattice with rank $m = 2$ and $\det(\mathcal{L}) = s-1$.

(ii) The domain K, defined in (6.3), is compact and convex with a non-empty interior.

(iii) The function $\phi : K \to \mathbb{R}$ is a continuous function with a unique global maximum (α_0, β_0), since we assume Conjecture 6.5 is true.

(iv) The function $\phi : K \to \mathbb{R}$ is twice differentiable in the interior of K, with a strictly negative definite Hessian, since we assume Conjecture 6.5 is true.

(v) Let K_1 be the open ball around (α_0, β_0) of sufficiently small radius, ensuring that $K_1 \subset K$. The function $\psi : K_1 \to \mathbb{R}$ is a continuous function with

$$\psi(\alpha_0, \beta_0) = \frac{r^{7/2}(s-1)^{5/2}}{\sqrt{r-1}(rs-r-s)^{\frac{2s}{s-1}}} > 0.$$

(vi) Let ℓ_n be $(0,1)$ for each n.

(vii) Lemma 2.25 and Corollary 2.26 verify this condition. We omit the details.

Thus, we can apply Lemma 2.27 to see that

$$\begin{aligned}\mathbb{E}Y^2 &\sim \frac{2\pi\,\psi(\alpha_0,\beta_0)}{\det(\mathcal{L})\det(-H_0)^{1/2}}\, b_n\, n\, e^{n\phi(\alpha_0,\beta_0)} \\ &= \frac{r^2\sqrt{r-1}(s-1)^{5/2}(rs-r-s)^{\frac{1}{2}-\frac{2s}{s-1}}}{n^2\sqrt{r^2-rs+r+s-1}}\left(\frac{(s-1)^{r/s}(r-1)^{r-1}}{r^{\frac{rs-r-s}{s}}(rs-r-s)^{\frac{rs-r-s}{s(s-1)}}}\right)^{2n}.\end{aligned}$$

Dividing by the expression for $\mathbb{E}Y$ twice from Theorem 4.1 gives us the result. □

To recap, (A1) is proven in [11], (A2) is essentially proven in Lemma 5.1 and (A3) is proven in Lemma 5.2. If Conjecture 6.5 is proven, then Lemma 6.7 establishes condition (A4) of Theorem 2.19, and thus, we will have verified all the conditions of Theorem 2.19. Applying Lemma 2.46 and Corollary 2.47, this will give us the asymptotic distribution of Y.

Conjecture 6.8. *For fixed $r \geq 3$ and $s \geq 3$, define*

$$\lambda_j = \frac{(r-1)^j(s-1)^j}{2j} \quad \text{and} \quad \zeta_j = \frac{\left(\frac{r}{r-1} - s + 1\right)^j - 2}{(r-1)^j(s-1)^j}.$$

Let $Y_{\mathcal{G}}$ be the number of spanning trees in a random hypergraph G from $\mathcal{G}_{n,r,s}$, and let Y be the number of spanning trees in a random partition P from $\mathcal{P}_{n,r,s}$. Then the asymptotic distribution of $Y_{\mathcal{G}}$ satisfies

$$\frac{Y_{\mathcal{G}}}{\mathbb{E}Y} \xrightarrow{d} e^{-\lambda_1\zeta_1} \prod_{j=2}^{\infty} (1+\zeta_j)^{Z_j}\, e^{-\lambda_j\zeta_j}$$

where $Z_j = \mathrm{Po}(\lambda_j)$ are independent Poisson random variables.

CHAPTER 7

Conclusion

The more I learn, the more I realise how much I don't know.
– Albert Einstein

In this thesis, we gained insight into the usefulness of the configuration model and the small subgraph conditioning method in proving asymptotic results for random regular graphs and random regular uniform hypergraphs. We proved a conjecture of Greenhill et al. [16] about the asymptotic distribution of the number of spanning trees in a random r-regular graph. We also managed to derive the asymptotic distribution of the number of spanning trees in a random r-regular 3-uniform hypergraph.

Naturally, further work needs to be done to prove Conjecture 6.5, which in turn, will prove Lemma 6.7 and Conjecture 6.8.

In addition, the small subgraph conditioning method for hypergraphs does not give the value of $\mathbb{E}Y_{\mathcal{G}}$, the expected number of spanning trees in a random graph G from $\mathcal{G}_{n,r,s}$. This is because convergence in distribution does not guarantee that expected values converge. It is natural to conjecture that

$$\mathbb{E}Y_{\mathcal{G}} \sim e^{-\lambda_1\zeta_1}\,\mathbb{E}Y = \exp\left(\frac{s}{2} - \frac{1}{2(r-1)}\right)\mathbb{E}Y.$$

It is likely that the application of the results in [2] will be useful in proving this conjecture, as it was for loose Hamilton cycles.

References

[1] M. Aigner and G. M. Ziegler, *Proofs from the Book*, Springer, Berlin, 2010.

[2] H. S. Aldosari and C. Greenhill, Enumerating sparse uniform hypergraphs with given degree sequence and forbidden edges, arXiv:1805.04991.

[3] D. Altman, C. Greenhill, M. Isaev, and R. Ramadurai, A threshold result for loose Hamiltonicity in random regular uniform hypergraphs, arXiv:1611.09423.

[4] L. D. Andersen and H. Fleischner, The NP-completeness of finding A-trails in Eulerian graphs and of finding spanning trees in hypergraphs, *Discrete Applied Mathematics*, **59(3)** (1995), 203–214.

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

[6] R. Bacher, Prüfer codes for hypertrees, arXiv:1301.4955.

[7] 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)** (1978), 296–307.

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

[9] W. Chu, On an extension of a partition identity and its Abel-analog, *Journal of Mathematical Research and Exposition*, **6(4)** (1986), 37–39.

[10] A. Coja-Oghlan and N. C. Wormald, The number of satisfying assignments of random regular k-SAT formulas, *Combinatorics, Probability and Computing*, **27(4)** (2018), 496–530.

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

[12] M. Delcourt and L. Postle, Random 4-regular graphs have 3-star decompositions asymptotically almost surely, *European Journal of Combinatorics*, **72** (2018), 97–111.

[13] R. Diestel, *Graph Theory*, Springer, New York, 2016.

[14] P. Duchet, Hypergraphs, in *Handbook of Combinatorics* **1**, MIT Press, Cambridge, Massachusetts, pp. 381–432, 1995.

[15] C. Greenhill, S. Janson, and A. Ruciński, On the number of perfect matchings in random lifts, *Combinatorics, Probability and Computing* **19** (2010), 791–817.

[16] C. Greenhill, M. Kwan, and D. Wind, On the number of spanning trees in random regular graphs, *The Electronic Journal of Combinatorics* **21(1)** (2014), 1–45.

[17] O. Güler. *Foundations of Optimization*, Springer Science & Business Media, New York, 2010.

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

[19] M. Kang and Z. Petrášek, Random graphs: Theory and applications from nature to society to the brain, *Internationale mathematische Nachrichten* **227** (2014), 1–24.

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

[21] J. W. Moon, *Counting Labelled Trees*, Canadian Mathematical Monographs, Vol. 1, Canadian Mathematical Congress, Montreal, 1970.

[22] R. Motwani and P. Raghavan, *Randomized Algorithms*, Cambridge University Press, Cambridge, 1995.

[23] P. Prałat and N. C. Wormald, Almost all 5-regular graphs have a 3-flow, In *Extended Abstracts Summer 2015*, pages 89–94. Springer, 2017.

[24] H. Robalewska, 2-factors in random regular graphs, *Journal of Graph Theory*, **23** (1996), 215–224.

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

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

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

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

[29] D. M. Warme, *Spanning trees in hypergraphs with applications to Steiner trees*, Ph.D Thesis, University of Virginia, 1998.

[30] H. S. Wilf, *generatingfunctionology*, Academic Press, Cambridge, MA, 1994.

[31] N. C. Wormald, The asymptotic connectivity of labelled regular graphs, *Journal of Combinatorial Theory, Series B* **31(2)**, (1981), 156–167.

[32] N. C. Wormald, The asymptotic distribution of short cycles in random regular graphs. *Journal of Combinatorial Theory, Series B* **31(2)** (1981), 168–182.

[33] N. C. Wormald, Models of random regular graphs, in *Surveys in Combinatorics, 1999* (J.D. Lamb and D.A. Preece, eds), London Mathematical Society Lecture Note Series **267**, Cambridge University Press, Cambridge, pp. 239–298, 1999.