

1-1-2013

Applications of the Lopsided Lovász Local Lemma Regarding Hypergraphs

Austin Tyler Mohr
University of South Carolina

Follow this and additional works at: <http://scholarcommons.sc.edu/etd>

Recommended Citation

Mohr, A. T.(2013). *Applications of the Lopsided Lovász Local Lemma Regarding Hypergraphs*. (Doctoral dissertation). Retrieved from <http://scholarcommons.sc.edu/etd/1606>

This Open Access Dissertation is brought to you for free and open access by Scholar Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact SCHOLARC@mailbox.sc.edu.

APPLICATIONS OF THE LOPSIDED LOVÁSZ LOCAL LEMMA
REGARDING HYPERGRAPHS

by

Austin Mohr

Bachelor of Science
Southern Illinois University at Carbondale, 2007

Master of Science
Southern Illinois University at Carbondale, 2008

Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in

Mathematics

College of Arts and Sciences

University of South Carolina

2013

Accepted by:

László Székely, Major Professor

Joshua Cooper, Committee Member

Linyuan Lu, Committee Member

George McNulty, Committee Member

Peter Nyikos, Committee Member

Edsel Peña, External Examiner

Lacy Ford, Vice Provost and Dean of Graduate Studies

DEDICATION

Dedicated to the memory of Professor Thomas D. Porter,
whose love of graphs lives on in the hearts of his students.

ACKNOWLEDGMENTS

I give my sincere thanks to the Department of Mathematics and the Graduate School of the University of South Carolina at Columbia, whose combined generosity made this endeavor possible.

I am grateful to Dr. Linyuan Lu, whose expertise has touched many of the results contained herein.

I am endlessly indebted to my advisor, Dr. László Székely. He has guided me through this thicket with patience and skill. In so doing, he has enabled me to begin exploring on my own.

Finally, I offer much love to my wife, Mary, who gives me courage when I have none. All my accomplishments are hers also.

ABSTRACT

The Lovász local lemma is a powerful and well-studied probabilistic technique useful in establishing the possibility of simultaneously avoiding every event in some collection. A principle limitation of the lemma's application is that it requires most events to be independent of one another. The lopsided local lemma relaxes the requirement of independence to negative dependence, which is more general but also more difficult to identify. We will examine general classes of negative dependent events involving maximal matchings of uniform hypergraphs, partitions of sets, and spanning trees of complete graphs. The results on hypergraph matchings (together with the configuration model of Bollobás) yield asymptotically the number of regular, uniform hypergraphs avoiding small cycles. Finally, we work toward a characterization of hypergraphs for which the matching paradigm is guaranteed generate negative dependent events.

TABLE OF CONTENTS

DEDICATION	ii
ACKNOWLEDGMENTS	iii
ABSTRACT	iv
LIST OF FIGURES	vii
CHAPTER 1 THE LOPSIDED LOCAL LEMMA	1
1.1 Dependency Graphs and the Local Lemma	1
1.2 2-Coloring Hypergraphs	3
1.3 Negative Dependency Graphs and the Lopsided Local Lemma	4
1.4 Counting Derangements	7
1.5 Asymptotic Enumeration with the Lopsided Local Lemma	10
CHAPTER 2 LOPSIDED LOCAL LEMMA FOR HYPERGRAPH MATCHINGS	14
2.1 Introduction	14
2.2 Example Conflict Graph	15
2.3 Negative Dependency Graph	16
2.4 Positive Dependency Graph	16
2.5 Asymptotics for Avoiding Matchings	17
CHAPTER 3 NEGATIVE DEPENDENCY GRAPHS FOR SET PARTITIONS	19
3.1 Introduction	19
3.2 Example Conflict Graph	20
3.3 A Class of Counterexamples	20

3.4	Results	22
3.5	Failed Attempt via Injection	27
3.6	Further Research	31
CHAPTER 4 NEGATIVE DEPENDENCY GRAPHS FOR SPANNING TREES . . .		33
4.1	Introduction	33
4.2	Example Conflict Graph	34
4.3	Results	36
CHAPTER 5 ENUMERATION OF REGULAR UNIFORM HYPERGRAPHS		42
5.1	Configuration Model for Hypergraphs	42
5.2	Cycles in Hypergraphs	43
5.3	Applying the Lopsided Local Lemma	44
5.4	Further Research	46
CHAPTER 6 PERFECT MATCHING HOSTS		47
6.1	General Results	47
6.2	2-Uniform Perfect Matching Hosts via Random Matchability	52
6.3	2-Uniform Perfect Matching Hosts via Corollary 6.3	54
6.4	Further Research	57
BIBLIOGRAPHY		59
APPENDIX A DETAILS FOR HYPERGRAPH MATCHINGS		62
A.1	Preliminaries	62
A.2	Proofs of Theorems 2.2 and 2.3	64
APPENDIX B USEFUL FACTS ABOUT BELL NUMBERS		72
APPENDIX C PROOF OF THEOREM 5.1		75

LIST OF FIGURES

Figure 2.1	Canonical event for single-edge matching in K_6	15
Figure 2.2	Matchings K , L , and M , respectively.	16
Figure 4.1	Forest in K_4 and its canonical event.	35
Figure 4.2	Forests D , E , and F , respectively.	35
Figure 5.1	Configuration projecting to 3-regular, 2-uniform multihyper- graph on four vertices.	43
Figure 6.1	Perfect matching host (no F as described in theorem).	54
Figure 6.2	Not a perfect matching host, as evidenced by red edges.	54
Figure 6.3	3-uniform, randomly matchable perfect matching host.	58

CHAPTER 1

THE LOPSIDED LOCAL LEMMA

1.1 DEPENDENCY GRAPHS AND THE LOCAL LEMMA

A collection of events is **avoidable** whenever the probability that no event in the collection occurs is nonzero. In the language of probability, a collection of events A_1, \dots, A_n is avoidable precisely when $\Pr\left(\bigwedge_{i=1}^n \overline{A_i}\right) > 0$.

Any finite collection of mutually independent events is avoidable (provided, of course, no event occurs with probability 1); the probability of avoiding the collection is $\prod_{i=1}^n \Pr\left(\overline{A_i}\right)$, which is greater than zero.

The requirement of mutual independence is quite stringent. We might expect the collection can still be avoided as long as the events do not depend strongly on one another. The Lovász local lemma makes this intuition precise by providing restrictions on the interdependence of events sufficient to guarantee the possibility of avoiding the collection. Erdős and Lovász [13] first introduced this idea to establish the existence of a certain hypergraph coloring. Subsequent generalizations culminated in the customary version appearing, for example, in Alon and Spencer [2].

A key component in the lemma is the **dependency graph**, which is a simple graph $([n], E)$ whose edges are situated such that the each event A_i is independent of the event algebra generated by the collection $\{A_j \mid ij \notin E\}$.

The dependency graph is a convenient way to organize information about possible dependencies among the events. For example, suppose we want to know the relationship between the event A_1 and some other events in our collection. If there is an

edge between A_1 and another event, the graph contains no information about their relationship. If we consider a collection of non-neighbors of A_1 , however, the graph tells us that A_1 is independent of the event algebra generated by the non-neighboring events.

Trivially, a complete graph is always a dependency graph. Such a graph is useless, however, since it contains no information about the events. At the other extreme, a dependency graph with no edges tells us we are dealing with a collection of mutually independent events. In practice, therefore, we hope to produce a dependency graph that is as sparse as possible, since this tells us the collection in question is very much like a collection of mutually independent events.

The Lovász local lemma asserts a collection is avoidable whenever there is a corresponding dependency graph together with an intricate upper bound on the probability of each individual event. In fact, it provides an explicit lower bound on the probability of avoiding the collection. A *symmetric* version (that is, one in which the probability of every event is given the same upper bound) of the lemma was first introduced by Erdős and Lovász [13] to address a question about hypergraph 2-colorability, which we will discuss later. Subsequently, Spencer [26, 27] presented various generalizations of the lemma in his work on Ramsey numbers. Presented below is a slight weakening of the customary version appearing in Alon and Spencer [2]. (In that text, the local lemma and the lopsided version to come are both written in terms of digraphs rather than graphs. The topics herein will not require this additional generality.)

Lemma 1.1 (Lovász Local Lemma). *Let A_1, \dots, A_n be events with dependency graph $([n], E)$. If there are numbers $x_1, \dots, x_n \in [0, 1)$ such that*

$$\Pr(A_i) \leq x_i \prod_{ij \in E} (1 - x_j)$$

for all i , then

$$\Pr\left(\bigwedge_{i=1}^n \overline{A_i}\right) \geq \prod_{i=1}^n (1 - x_i) > 0.$$

1.2 2-COLORING HYPERGRAPHS

Erdős and Lovász [13] provide an upper bound on the maximum degree of a properly 2-colorable hypergraph can contain. Their idea was to color the vertices uniformly at random with two colors and impose conditions under which the random process would produce a proper coloring with nonzero probability. They developed the symmetric version of the local lemma as a tool to aid in the analysis of the resulting probability space.

Lemma 1.2 (Lovász Local Lemma, Symmetric Version). *Let A_1, \dots, A_n be events with dependency graph of maximum degree d . If*

- $\Pr(A_i) \leq p$ for all i and
- $ep(d+1) \leq 1$,

then

$$\Pr\left(\bigwedge_{i=1}^n \overline{A_i}\right) > 0.$$

(The symmetric version follows from the Lemma 1.1 by setting each $x_i = \frac{1}{d+1}$ and using the fact that $\left(1 - \frac{1}{d+1}\right)^d > \frac{1}{e}$.)

Let H be a hypergraph in which every edge contains at least k vertices and color the vertices uniformly at random with two colors. Our ambient probability space will therefore contain all possible 2-colorings (both proper and improper) of the vertices of H weighted uniformly. For each edge $f \in E(H)$, define the event A_f to be the collection of all 2-colorings in which the edge f is monochromatic. The event $\bigwedge_{f \in E(H)} \overline{A_f}$ thus contains all 2-colorings in which *no* edge is monochromatic. That is, it contains all proper 2-colorings. We are therefore interested in determining when $\Pr\left(\bigwedge_{f \in E(H)} \overline{A_f}\right) > 0$, which can be approached via the local lemma.

First, notice the events in the collection $\{A_f \mid f \in F\}$ are mutually independent whenever F is a collection of disjoint edges of H . The graph G with $V(G) = E(H)$

and

$$E(G) = \{fg \mid f \text{ and } g \text{ share a vertex in } H\}$$

is therefore a dependency graph whose maximum degree is the same as the maximum degree of H .

By hypothesis, every edge of H has at least k vertices, so $\Pr(A_f) \leq \frac{1}{2^{k-1}}$, which we take as our value for p in the lemma. It remains to maximize d under the constraint $ep(d+1) < 1$, which works out to $d = \lfloor \frac{2^{k-1}}{e} - 1 \rfloor$.

Theorem 1.3. *Let H be a hypergraph in which every edge contains at least k vertices. If H has maximum degree at most $\frac{2^{k-1}}{e} - 1$, then H is 2-colorable.*

Before leaving this problem behind, notice we could only apply the local lemma by ensuring that “most” events were independent. We will see later that the lopsided local lemma can be applied in spaces where there is *no* independence among the events.

1.3 NEGATIVE DEPENDENCY GRAPHS AND THE LOPSIDED LOCAL LEMMA

The Lovász local lemma allows us to detect avoidability if there is only *some* independence among the events (provided we can discover a dependency graph and suitable numbers x_i). Erdős and Spencer [14] analyzed the proof and determined that one can still detect avoidability even if there is *no* independence among the events.

In the following definition, let $N(v)$ denote the set of neighbors of the vertex v together with v itself. A **negative dependency graph** for a collection of events A_1, \dots, A_n is a simple graph $([n], E)$ whose edges are situated such that the inequality

$$\Pr\left(A_i \mid \bigwedge_{j \in S} \overline{A_j}\right) \leq \Pr(A_i) \tag{1.1}$$

holds for each $i \in [n]$ and every subset S of $\overline{N(i)}$ (excluding those S for which the event $\bigwedge_{j \in S} \overline{A_j}$ has probability zero, in which case the conditional probability is

undefined). Stated in this way the inequality might be crudely summarized to say that the probability of an event falls when some of its non-neighbors do not occur.

Alternative formulations of Inequality 1.1 arise from straightforward algebraic manipulation. As before, we assume $i \in [n]$ and $S \subseteq \overline{N(i)}$ are arbitrary, except that we do not consider collections S for which the conditioning event (if any) has probability zero. The first two are the conditional probabilities

$$\Pr \left(\bigwedge_{j \in S} \overline{A_j} \mid A_i \right) \leq \Pr \left(\bigwedge_{j \in S} \overline{A_j} \right)$$

and

$$\Pr(\overline{A_i}) \leq \Pr \left(\bigwedge_{j \in S \cup \{i\}} \overline{A_j} \mid \bigwedge_{j \in S} \overline{A_j} \right).$$

Our final formulation takes the form of the correlation inequality

$$\Pr(A_i) \Pr \left(\bigvee_{j \in S} A_j \right) \leq \Pr \left(A_i \wedge \bigvee_{j \in S} A_j \right).$$

A collection of events may satisfy the inequalities above even though no two events are independent, as we describe in the next section.

The lopsided local lemma differs from the previous version only by replacing “dependency graph” with “negative dependency graph”. Since every dependency graph is a negative dependency graph (but not vice versa), the lopsided version is strictly more general. It was first introduced by Erdős and Spencer [14] with regard to Latin transversals and independently by Albert, Frieze, and Reed [1] in their work on Hamiltonian cycles. The lemma as it appears below is due to Ku [17] and appears as a remark in Alon and Spencer [2].

Lemma 1.4 (Lopsided Local Lemma). *Let A_1, \dots, A_n be events with negative dependency graph $([n], E)$. If there are numbers $x_1, \dots, x_n \in [0, 1)$ such that*

$$\Pr(A_i) \leq x_i \prod_{ij \in E} (1 - x_j)$$

for all i , then

$$\Pr \left(\bigwedge_{i=1}^n \overline{A_i} \right) \geq \prod_{i=1}^n (1 - x_i) > 0.$$

Proof. Take as granted for a moment that

$$\Pr \left(A_i \mid \bigwedge_{j \in S} \overline{A_j} \right) \leq x_i \quad (1.2)$$

for any strict subset S of $[n]$ and any $i \notin S$.

The conclusion of the lopsided local lemma follows from this claim by observing

$$\begin{aligned} \Pr \left(\bigwedge_{i=1}^n \overline{A_i} \right) &= \Pr \left(\overline{A_1} \right) \cdot \Pr \left(\overline{A_2} \mid \overline{A_1} \right) \cdots \Pr \left(\overline{A_n} \mid \bigwedge_{j=1}^{n-1} \overline{A_j} \right) \\ &\geq \prod_{i=1}^n (1 - x_i) \\ &> 0. \end{aligned}$$

It remains to establish Inequality 1.2, which we accomplish by induction on $|S|$. When $|S| = 0$, the claimed inequality reduces to $\Pr(A_i) \leq x_i$, which is provided by the hypotheses of the lopsided local lemma. For $|S| > 0$, set $S_1 = \{j \in S \mid ij \in E\}$ and $S_2 = S \setminus S_1$. Now,

$$\Pr \left(A_i \mid \bigwedge_{j \in S} \overline{A_j} \right) = \frac{\Pr \left(A_i \wedge \bigwedge_{j \in S_1} \overline{A_j} \mid \bigwedge_{k \in S_2} \overline{A_k} \right)}{\Pr \left(\bigwedge_{j \in S_1} \overline{A_j} \mid \bigwedge_{k \in S_2} \overline{A_k} \right)}.$$

We will bound the numerator and denominator separately.

For the numerator, we have

$$\begin{aligned} \Pr \left(A_i \wedge \bigwedge_{j \in S_1} \overline{A_j} \mid \bigwedge_{k \in S_2} \overline{A_k} \right) &\leq \Pr \left(A_i \mid \bigwedge_{k \in S_2} \overline{A_k} \right) \\ &\leq \Pr(A_i) \\ &\leq x_i \prod_{j \in S_1} (1 - x_j), \end{aligned}$$

where the second inequality comes from the fact that A_i is negative dependent of the collection $\{A_k \mid k \in S_2\}$.

For the denominator, write $S_1 = \{j_1, \dots, j_r\}$ (if it is empty, the denominator is

equal to 1). Now,

$$\begin{aligned}
& \Pr \left(\bigwedge_{\ell=1}^r \overline{A_{j_\ell}} \mid \bigwedge_{k \in S_2} \overline{A_k} \right) \\
&= \Pr \left(\overline{A_{j_1}} \mid \bigwedge_{k \in S_2} \overline{A_k} \right) \cdot \Pr \left(\overline{A_{j_2}} \mid \overline{A_{j_1}} \wedge \bigwedge_{k \in S_2} \overline{A_k} \right) \cdots \Pr \left(\overline{A_{j_r}} \mid \bigwedge_{\ell=1}^{r-1} \overline{A_{j_\ell}} \wedge \bigwedge_{k \in S_2} \overline{A_k} \right) \\
&\geq \prod_{\ell=1}^r (1 - x_{j_\ell}),
\end{aligned}$$

where the inequality holds by the induction hypothesis (in each factor, we condition on an intersection of fewer than $|S|$ events).

Combining the two bounds,

$$\begin{aligned}
\Pr \left(A_i \mid \bigwedge_{j \in S} \overline{A_j} \right) &= \frac{\Pr \left(A_i \wedge \bigwedge_{j \in S_1} \overline{A_j} \mid \bigwedge_{k \in S_2} \overline{A_k} \right)}{\Pr \left(\bigwedge_{j \in S_1} \overline{A_j} \mid \bigwedge_{k \in S_2} \overline{A_k} \right)} \\
&\leq \frac{x_i \prod_{j \in S_1} (1 - x_j)}{\prod_{j \in S_1} (1 - x_j)} \\
&= x_i,
\end{aligned}$$

which proves the claim. □

1.4 COUNTING DERANGEMENTS

A **derangement** is a permutation having no fixed point. It is well known that the number of derangements on the set $[N]$ is the integer nearest to $\frac{N!}{e}$ [15]. The lopsided local lemma gives this value as an asymptotic lower bound. (Using the forthcoming machinery of positive dependency graphs, Lu and Székely [21] obtained this as an asymptotic *upper* bound, as well.)

In the uniform probability space containing all permutations on the set $[N]$, let A_i denote the collection of all such permutations having i as a fixed point. The event $\bigwedge_{i=1}^N \overline{A_i}$ contains precisely those permutations having *no* fixed point (that is, the derangements on $[N]$).

No pair of distinct events A_i and A_j are independent, since

$$\Pr(A_i \wedge A_j) = \frac{(N-2)!}{N!} = \frac{1}{N^2 - N},$$

while

$$\Pr(A_i) \Pr(A_j) = \frac{(N-1)!}{N!} \cdot \frac{(N-1)!}{N!} = \frac{1}{N^2}.$$

For this reason, the local lemma fails in the worst possible way. Remarkably, the lop-sided local lemma succeeds in the *best* possible way, allowing for an *edgeless* negative dependency graph.

Theorem 1.5. *In the uniform probability space containing all permutations on the set $[N]$, let A_i denote the collection of all such permutations having i as a fixed point. The graph with vertex set $[N]$ and no edges is a negative dependency graph for the events $\{A_1, \dots, A_N\}$.*

Lu and Székely [20] prove a more general statement about random injections, of which the theorem above is a simple case. Before presenting an alternative proof, let us take a moment to see why we might expect it to hold for just two events. Asking whether $\Pr(A_1 \mid \overline{A_2})$ is less than or equal to $\Pr(A_1)$ can be phrased as follows: Does the knowledge that the element 2 is *not* a fixed point reduce the likelihood that the element 1 *is* a fixed point? The fact that 2 is not a fixed point means it is slightly more likely than usual that it is mapped to 1, so it is slightly *less* likely than usual that 1 is a fixed point.

Proof of Theorem 1.5. Without loss of generality, we will establish the inequality

$$\Pr\left(\bigwedge_{j=1}^k \overline{A_j} \mid A_N\right) \leq \Pr\left(\bigwedge_{j=1}^k \overline{A_j}\right)$$

for any $k \in [N-1]$, which is defined to be

$$\frac{|A_N \wedge \bigwedge_{j=1}^k \overline{A_j}|}{|A_N|} \leq \frac{|\bigwedge_{j=1}^k \overline{A_j}|}{N!}.$$

Now, A_N is the collection of permutations on the set $[N]$ having N as a fixed point, so $|A_N| = (N - 1)!$. Clearing denominators, we are left with establishing the inequality

$$N \left| A_N \wedge \bigwedge_{j=1}^k \overline{A_j} \right| \leq \left| \bigwedge_{j=1}^k \overline{A_j} \right|,$$

which we achieve by the following combinatorial argument.

Since all permutations belonging to $A_N \wedge \bigwedge_{j=1}^k \overline{A_j}$ have N as a fixed point, we may view the event as the collection of all permutations on the set $[N - 1]$ having no fixed points in the set $[k]$. The event $\bigwedge_{j=1}^k \overline{A_j}$ is the collection of all permutations on the set $[N]$ also having no fixed points in the set $[k]$. Denote these collections by \mathcal{A}_{N-1} and \mathcal{A}_N , respectively.

For any $\sigma \in \mathcal{A}_{N-1}$, define σ_i for each $i \in [N]$ via

$$\sigma_i(j) = \begin{cases} N & \text{if } j = i \\ \sigma(i) & \text{if } j = N \\ \sigma(j) & \text{otherwise.} \end{cases}$$

Each σ_i is distinct, since $\sigma_i(N) \neq \sigma_j(N)$ whenever $i \neq j$. Moreover, distinct permutations σ and τ belonging to \mathcal{A}_{N-1} must differ in at least two coordinates, so $\sigma_i \neq \tau_j$ for any i and j . Finally, since σ has no fixed points in $[k]$, neither does σ_i for any i (recall that $N \notin [k]$), which means each σ_i belongs to \mathcal{A}_N . Taken together, we conclude $N|\mathcal{A}_{N-1}| \leq |\mathcal{A}_N|$, as desired. \square

With an edgeless negative dependency graph in hand, we can take each $x_i = \frac{1}{N}$ in the lopsided local lemma, since

$$\Pr(A_i) = \frac{1}{N} = x_i = x_i \prod_{ij \in \emptyset} (1 - x_j).$$

The lopsided local lemma concludes

$$\Pr\left(\bigwedge_{i=1}^N \overline{A_i}\right) \geq \prod_{i=1}^N \left(1 - \frac{1}{N}\right) = \left(1 - \frac{1}{N}\right)^N,$$

which converges to $\frac{1}{e}$. Therefore, $\frac{N!}{e}$ is an asymptotic lower bound for the *number* of derangements on the set $[N]$.

1.5 ASYMPTOTIC ENUMERATION WITH THE LOPSIDED LOCAL LEMMA

For asymptotic analysis, we will be interested in sequences of problems in which the size of the events and/or the ambient probability space depends on some parameter tending toward infinity. For example, the events “at least one head” or “at least \sqrt{N} heads” in the probability space of all possible outcomes of N coin flips induce a sequence of problems when N grows without bound. We will denote such a growing probability space by Ω_N to emphasize that its size depends on N . Lu and Székely [21] provide conditions under which an asymptotic lower bound for $\Pr\left(\bigwedge_{i=1}^n \overline{A}_i\right)$ can be obtained from the lopsided local lemma. Notice $\Pr(A_i)$ will depend on N even if the event A_i does not explicitly reference N , since the size of the ambient probability space Ω_N grows with N .

Theorem 1.6 (Lu, Székely 2011). *Let A_1, \dots, A_n be events in a probability space Ω_N with negative dependency graph $([n], E)$ and set $\mu = \sum_{i=1}^n \Pr(A_i)$. If there is ϵ (depending on N) such that*

- $\Pr(A_i) < \epsilon$ for all i ,
- $\sum_{j:i_j \in E} \Pr(A_j) + 2\Pr^2(A_j) < \epsilon$ for all i , and
- $\epsilon\mu$ tends to zero as N tends to infinity,

then

$$\Pr\left(\bigwedge_{i=1}^n \overline{A}_i\right) \geq (1 - o(1))e^{-\mu}.$$

Negative dependency graphs are more general but also more difficult to identify than dependency graphs. Lu, Székely, and the author [19] explored **conflict graphs** in several disparate classes of combinatorial objects, which has been a successful avenue in discovering negative dependency graphs. For the sake of concreteness, we will define the conflict graph separately in each of the sections on hypergraph

matchings, set partitions, and spanning trees, which are the classes of combinatorial objects with which we will be concerned.

To obtain an asymptotic *upper* bound for $\Pr\left(\bigwedge_{i=1}^n \overline{A_i}\right)$, Lu and Székely [21] introduced the ϵ -near **positive dependency graph**. In the event that the lower and upper bounds match in the limit, one obtains an asymptotic expression for the probability of interest.

For events A_1, \dots, A_n and $\epsilon \in (0, 1)$, an ϵ -near positive dependency graph $([n], E)$ is one in which

- $\Pr(A_i \wedge A_j) = 0$ whenever $ij \in E$ and
- the inequality

$$\Pr\left(A_i \mid \bigwedge_{j \in S} \overline{A_j}\right) \geq (1 - \epsilon) \Pr(A_i)$$

holds for each i and any subset S of $\overline{N(i)}$ (excluding those S for which the event $\bigwedge_{j \in S} \overline{A_j}$ has probability zero, in which case the conditional probability is undefined).

Notice the reversal in the direction of the inequality (as compared with the negative dependency graph) results in an *upper bound* on $\Pr\left(\bigwedge_{i=1}^n \overline{A_i}\right)$.

Theorem 1.7 (Lu, Székely 2011). *If A_1, \dots, A_n are events with an ϵ -near positive dependency graph, then*

$$\Pr\left(\bigwedge_{i=1}^n \overline{A_i}\right) \leq \prod_{i=1}^n [1 - (1 - \epsilon) \Pr(A_i)].$$

With some extra restrictions, this upper bound meets the lower bound in 1.6 asymptotically.

Corollary 1.8. *Let A_1, \dots, A_n be events with an ϵ -near positive dependency graph in a probability space growing with N and set $\mu = \sum_{i=1}^n \Pr(A_i)$. If both $\epsilon\mu$ and*

$\sum_{i=1}^n \Pr^2(A_i)$ tend to zero as N tends to infinity, then

$$\Pr\left(\bigwedge_{i=1}^n \overline{A_i}\right) \leq (1 + o(1))e^{-\mu}.$$

Proof. Theorem 1.7 gives

$$\begin{aligned} \Pr\left(\bigwedge_{i=1}^n \overline{A_i}\right) &\leq \prod_{i=1}^n [1 - (1 - \epsilon) \Pr(A_i)] \\ &= \exp\left(\sum_{i=1}^n \log [1 - (1 - \epsilon) \Pr(A_i)]\right). \end{aligned}$$

Using the fact that

$$\log(1 - x) = -\sum_{k=1}^{\infty} \frac{x^k}{k}$$

for $|x| < 1$, we write

$$\log [1 - (1 - \epsilon) \Pr(A_i)] = -\sum_{k=1}^{\infty} \frac{(1 - \epsilon)^k \Pr^k(A_i)}{k}$$

for each i . Now,

$$\begin{aligned} -\sum_{i=1}^n \sum_{k=1}^{\infty} \frac{(1 - \epsilon)^k \Pr^k(A_i)}{k} &= -\sum_{i=1}^n (1 - \epsilon) \Pr(A_i) - \sum_{i=1}^n \sum_{k=2}^{\infty} \frac{(1 - \epsilon)^k \Pr^k(A_i)}{k} \\ &= -\sum_{i=1}^n (1 - \epsilon) \Pr(A_i) - \sum_{i=1}^n O\left(\Pr^2(A_i)\right) \\ &= -\sum_{i=1}^n (1 - \epsilon) \Pr(A_i) - O\left(\sum_{i=1}^n \Pr^2(A_i)\right). \end{aligned}$$

Substituting into the exponential, we have

$$\begin{aligned} \Pr\left(\bigwedge_{i=1}^n \overline{A_i}\right) &\leq \exp\left(\sum_{i=1}^n \log [1 - (1 - \epsilon) \Pr(A_i)]\right) \\ &= \exp\left(-\sum_{i=1}^n (1 - \epsilon) \Pr(A_i) - O\left(\sum_{i=1}^n \Pr^2(A_i)\right)\right) \\ &= \exp(-\mu) \exp\left(\epsilon\mu + O\left(\sum_{i=1}^n \Pr^2(A_i)\right)\right) \\ &= \exp(-\mu) (1 + o(1)). \end{aligned}$$

□

In a sequence of problems satisfying the conditions of both Theorem 1.6 and Corollary 1.8, we can conclude $\Pr\left(\bigwedge_{i=1}^n \overline{A_i}\right)$ is asymptotic to $e^{-\mu}$. If we further assume that the ambient probability space is equipped with the counting measure, then multiplying by the size of the space gives an asymptotic expression for the number of outcomes avoiding the events A_1, \dots, A_n .

Corollary 1.9. *Let A_1, \dots, A_n be events in a uniform probability space Ω_N equipped with the counting measure and set $\mu = \sum_{i=1}^n \Pr(A_i)$. If the conditions of both Theorem 1.6 and Corollary 1.8 are satisfied, then*

$$\left| \bigcap_{i=1}^n \overline{A_i} \right| = (1 + o(1)) |\Omega_N| e^{-\mu}.$$

CHAPTER 2

LOPSIDED LOCAL LEMMA FOR HYPERGRAPH MATCHINGS

2.1 INTRODUCTION

An **s -matching** (or simply **matching**) in an s -uniform hypergraph is a collection of vertex-disjoint edges (each containing s vertices) and is **maximal** provided no strictly larger matching contains it. Let Ω denote the uniform probability space consisting of all maximal matchings of some underlying s -uniform hypergraph H . Our primary objective in this section will be to define a conflict graph for events in Ω (analogous to the one defined in Chapter 3 for set partitions) and present some conditions under which it is a negative dependency graph.

For a particular matching M , define A_M to be the collection of all maximal matchings extending M . More precisely,

$$A_M = \{L \in \Omega \mid M \subseteq L\}.$$

We call the collection A_M the **canonical event** for the matching M to emphasize its interpretation as an event in the probability space Ω . Two matchings **conflict** whenever their union is not again a matching, and two canonical events conflict when the matchings used to define them conflict.

Finally, let \mathcal{M} be any collection of s -matchings. The **conflict graph** for the collection $\{A_M \mid M \in \mathcal{M}\}$ of canonical events is a simple graph whose vertex set is \mathcal{M} . Two matchings are adjacent in this graph if and only if they conflict.

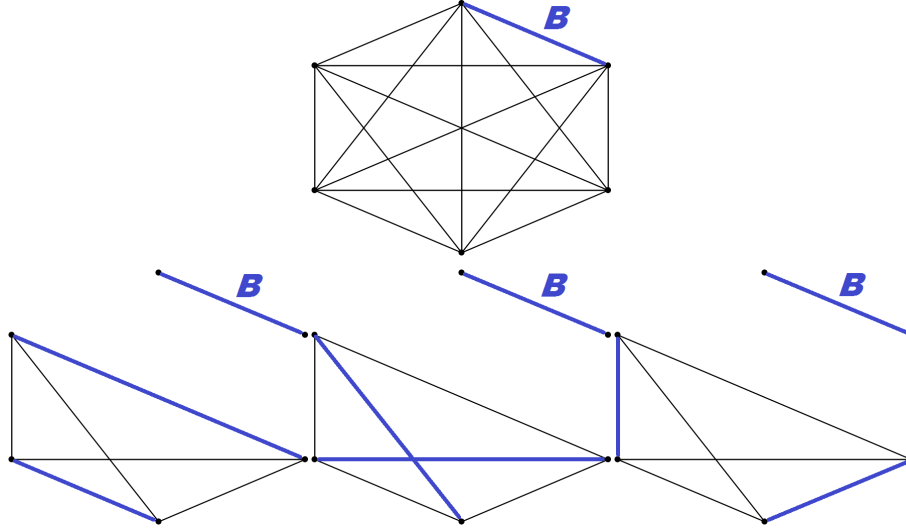


Figure 2.1 Canonical event for single-edge matching in K_6 .

2.2 EXAMPLE CONFLICT GRAPH

Take the complete graph on six vertices to be the underlying graph (a graph is a 2-uniform hypergraph). The single-edge matching B is pictured in Figure 2.2 together with its canonical event A_B , which consists of the three maximal (indeed, perfect) matchings containing the edge B . The edges incident to a vertex of B are not pictured to emphasize the fact that they cannot possibly be including in any matching containing B . For this reason, there is a natural bijection between the outcomes of the canonical event A_B and the collection of perfect matchings of the complete graph on four vertices.

Unrelated to the previous example, consider the three matchings K , L , and M in the complete graph on six vertices pictured in Figure 2.2. The matchings K and M conflict, because their union is not again a matching. The canonical events (not pictured) A_K and A_M are disjoint, since no perfect matching extends both the matchings K and M simultaneously. Similarly, the matchings L and M conflict. The matchings K and L do not conflict, since their union is itself a matching (indeed, a perfect matching). The conflict graph for the these events is therefore the graph with

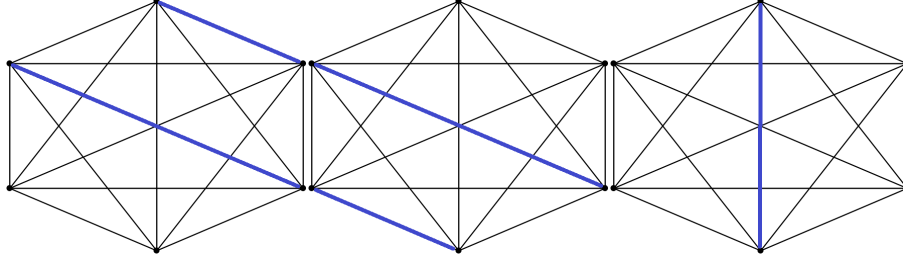


Figure 2.2 Matchings K , L , and M , respectively.

vertex set $\{K, L, M\}$ and edge set $\{KM, LM\}$.

2.3 NEGATIVE DEPENDENCY GRAPH

For the enumeration of regular uniform hypergraphs in Chapter 5, the underlying graph will be a complete uniform hypergraph, in which case the conflict graph is always a negative dependency graph [19].

Theorem 2.1 (Lu, M, Székely 2012). *Let \mathcal{M} be any collection of matchings in a complete uniform hypergraph. The conflict graph for the collection $\{A_M \mid M \in \mathcal{M}\}$ of canonical events is a negative dependency graph.*

At this point our interest in hypergraph matchings is two-fold. In Chapter 5, we apply the theorem above (together with other tools) to the asymptotic enumeration of regular uniform hypergraphs. In Chapter 6, we pursue the classification of the underlying hypergraphs for which the theorem above holds.

2.4 POSITIVE DEPENDENCY GRAPH

Let \mathcal{M} be a collection of matchings in the complete s -uniform hypergraph on N vertices with negative dependency graph (\mathcal{M}, E) , and let δ be a positive real number. (For the moment, we may suppose that s and δ are both fixed, but later applications will allow these to grow slowly with N .) Such a collection is δ -**sparse** provided

no matching from \mathcal{M} is a subset of another matching from \mathcal{M} and the following inequalities are satisfied for every matching $M \in \mathcal{M}$ and every edge $e \in E$:

- $\Pr(A_M) < \delta$
- $\sum_{L:LM \in E} \Pr(A_L) + \Pr^2(A_L) < \delta$
- $\sum_{L:e \in L} \Pr(A_L) + \Pr^2(A_L) < \delta$
- $\sum_{L \in \mathcal{M}_M} \Pr_{N-s|M|}(A_L) + \Pr_{N-s|M|}^2(A_L) < \delta,$

where

$$\mathcal{M}_M = \{L \setminus M \mid L \in \mathcal{M}, L \neq M, L \cap M \neq \emptyset, L \text{ does not conflict with } M\}.$$

A collection in which every matching contains at most k edges is **k -bounded**.

Theorem 2.2. *Let \mathcal{M} be a collection of matchings in a complete s -uniform hypergraph. If \mathcal{M} is δ -sparse and k -bounded, then the conflict graph for the canonical events $\{A_M \mid M \in \mathcal{M}\}$ is also an ϵ -near positive dependency graph.*

The precise relationship between the parameters k , s , δ , and ϵ is deferred to Section A.2.

2.5 ASYMPTOTICS FOR AVOIDING MATCHINGS

Let Ω_N be the uniform probability space of maximal matchings of a complete uniform hypergraph. In this space, the expression $\Pr\left(\bigwedge_{M \in \mathcal{M}} \overline{A_M}\right)$ denotes the probability that a maximal matching chosen uniformly at random from Ω_N contains no submatching belonging to the set \mathcal{M} . According to Theorem 2.1, the conflict graph for any collection of canonical matching events is a negative dependency graph. Theorem 2.2 gives some restrictions on \mathcal{M} under which we are assured the conflict graph is also a positive dependency graph. Theorem 1.6 and Corollary 1.8 give further restrictions to

ensure nice asymptotic behavior. We gather here all these conditions into one place to derive an asymptotic expression for $\Pr\left(\bigwedge_{M \in \mathcal{M}} \overline{A_M}\right)$. Take note that expressions such as $\Pr(A_M)$ will depend on N since the size of the ambient probability space Ω_N depends on N .

Theorem 2.3. *Let Ω_N denote the uniform probability space of perfect matchings of K_N^s , the complete s -uniform hypergraph on N vertices. Let r and ϵ both depend on N , where r is a positive integer and ϵ is a real number eventually lying in the interval $(0, \frac{1}{16})$. Let \mathcal{M} be a k -bounded collection of matchings in K_N^s in which no matching is a subset of another. For any matching $M \in \mathcal{M}$, define the canonical event*

$$A_M = \{L \in \Omega_N \mid M \subseteq L\}.$$

Set $\mu = \sum_{M \in \mathcal{M}} \Pr(A_M)$. Finally, suppose the following inequalities are satisfied for every matching $M \in \mathcal{M}$ and every edge e of K_N^s :

- $\Pr(A_M) < \epsilon$
- $\sum_{L: L, M \text{ conflict}} \Pr(A_L) < \epsilon$
- $\sum_{L \in \mathcal{M}: e \in L} \Pr(A_L) < \epsilon$
- $\sum_{L \in \mathcal{M}_M} \Pr_{N-sk}(A_L) < \epsilon$

If, in addition, $ks\epsilon = o(1)$, then

$$\Pr\left(\bigwedge_{M \in \mathcal{M}} \overline{A_M}\right) = e^{-\mu + O(ks\epsilon\mu)}. \quad (2.1)$$

Furthermore, if $ks\epsilon\mu = o(1)$, then

$$\Pr\left(\bigwedge_{M \in \mathcal{M}} \overline{A_M}\right) = (1 + O(ks\epsilon\mu)) e^{-\mu}.$$

The proof (like the statement) is technical, and we relegate it to Section A.2. We will make use of this result in Chapter 5, wherein we establish a bijection between a certain class of matchings and hypergraphs avoiding small cycles.

CHAPTER 3

NEGATIVE DEPENDENCY GRAPHS FOR SET PARTITIONS

Let Ω denote the uniform probability space consisting of all partitions of some underlying set X . (Equivalently stated, Ω contains all perfect matchings of complete *nonuniform* hypergraph on the vertex set X in which edges of the matching are not required to have the same size.) Our primary objective in this section will be to define a certain type of conflict graph for events in Ω and present some conditions under which it is a negative dependency graph.

3.1 INTRODUCTION

A **partial partition** is a collection of disjoint subsets of the underlying set X . (A partial partition may in fact fully partition the set X .) For a particular partial partition P , define A_P to be the collection of all (ordinary) partitions extending P . More precisely,

$$A_P = \{Q \in \Omega \mid P \subseteq Q\}.$$

(We are using the ordinary subset relation, not the refinement relation.) We call the collection A_P the **canonical event** for the partial partition P to emphasize its interpretation as an event in the probability space Ω . Two partial partitions **conflict** whenever their union is not again a partial partition, and two canonical events conflict when the partitions used to define them conflict.

Finally, let \mathcal{P} be any collection of partial partitions of the set X . The **conflict graph** for the collection $\{A_P \mid P \in \mathcal{P}\}$ of canonical events is a simple graph whose vertex set is \mathcal{P} . Two partitions are adjacent in this graph if and only if they conflict.

Table 3.1 Partial partitions of [5] and corresponding canonical events.

Partial Partition	Canonical Event
P 12 3	A_P 12 3 45 12 3 4 5
Q 14 2	A_Q 14 2 35 14 2 3 5
R 12 45	A_R 12 3 45

In this section we are concerned with characterizing the collections $\{A_P \mid P \in \mathcal{P}\}$ of canonical events for which the conflict graph is a negative dependency graph. In other words, for what collections \mathcal{P} of partial partitions does the inequality

$$\Pr \left(A_P \mid \bigwedge_{A_Q \in \mathcal{S}} \overline{A_Q} \right) \leq \Pr(A_P)$$

hold for each $P \in \mathcal{P}$ and every subset \mathcal{S} of $\overline{N(P)}$? (In this instance, $\overline{N(P)}$ is the subset of \mathcal{P} containing precisely those partial partitions that do not conflict with P .)

3.2 EXAMPLE CONFLICT GRAPH

Consider the underlying set $X = [5]$ and the partial partitions P , Q , and R defined in Table 3.2. (The notation, for example, 12 | 3 is shorthand for $\{\{1, 2\}, \{3\}\}$.)

The partial partitions P and Q conflict, since $P \cup Q = 12 \mid 3 \mid 14 \mid 2$ is not a collection of disjoint subsets. Notice the canonical events A_P and A_Q are disjoint. Similarly, the partial partitions Q and R conflict. On the other hand, the partial partitions P and R do not conflict, since $P \cup R = 12 \mid 3 \mid 45$ is itself a partial partition (in fact, it is an ordinary partition). The conflict graph for the associated canonical events is therefore the graph with vertex set $\{P, Q, R\}$ and edge set $\{PQ, QR\}$.

3.3 A CLASS OF COUNTEREXAMPLES

For convenience, let us henceforth assume that the underlying set of the partitions is $[N]$, write Ω_N to denote the space of all partitions of $[N]$, and write $\Pr_N(\cdot)$ to denote

the probability is taken with respect to the space Ω_N .

The conflict graph need not be a negative dependency graph for arbitrary collections \mathcal{P} of partial partitions. Indeed, a counterexample exists for every N , but the construction presented there seems to rely on the fact that each partial partition is quite large with respect to the underlying set. To present the counterexamples, we introduce N^{th} Bell number B_N [4, 5, 24], which is the number of partitions of an N -element set.

Let the size N of the underlying set be given. For each $i \in [N]$, let P_i denote the partition of $[N] \setminus \{i\}$ into singletons. We have

$$\Pr_N \left(\bigwedge_{i=1}^N \overline{A_{P_i}} \right) = \Pr_N \left(\bigvee_{i=1}^N A_{P_i} \right) = \frac{B_N - 1}{B_N},$$

since the only partition extending any of the P_i is the partition of $[N]$ into singletons. On the other hand, the partition of $[N + 1]$ into singletons extends any of the P_i . Each P_i is also extended by the partition containing only singletons except for the block $\{i, N + 1\}$. Thus we have,

$$\Pr_{N+1} \left(\bigwedge_{i=1}^N \overline{A_{P_i}} \right) = \Pr_{N+1} \left(\bigvee_{i=1}^N A_{P_i} \right) = \frac{B_{N+1} - (N + 1)}{B_{N+1}}.$$

Before proceeding, we will need the fact that

$$\Pr_N \left(\bigwedge_{i=1}^N \overline{A_{P_i}} \right) > \Pr_{N+1} \left(\bigwedge_{i=1}^N \overline{A_{P_i}} \right).$$

Begin with the fact that $(N + 1)B_N > B_{N+1}$ for all integers N (see Appendix B).

Now,

$$\begin{aligned} (N + 1)B_N &> B_{N+1} \\ -B_{N+1} &> -(N + 1)B_N \\ B_{N+1}B_N - B_{N+1} &> B_{N+1}B_N - (N + 1)B_N \\ \frac{B_N - 1}{B_N} &> \frac{B_{N+1} - (N + 1)}{B_{N+1}} \\ \Pr_N \left(\bigwedge_{i=1}^N \overline{A_{P_i}} \right) &> \Pr_{N+1} \left(\bigwedge_{i=1}^N \overline{A_{P_i}} \right), \end{aligned}$$

as desired.

Introduce now another partial partition $P = \{\{N + 1\}\}$. We have

$$\begin{aligned} & \Pr_{N+1} \left(\bigwedge_{i=1}^N \overline{A_{P_i}} \mid A_P \right) \\ &= \Pr_N \left(\bigwedge_{i=1}^N \overline{A_{P_i}} \right) \\ &> \Pr_{N+1} \left(\bigwedge_{i=1}^N \overline{A_{P_i}} \right), \end{aligned}$$

violating the condition of negative dependence. Thus, the conflict graph for the events $A_P, A_{P_1}, \dots, A_{P_N}$ is not a negative dependency graph in the space Ω_{N+1} .

3.4 RESULTS

For a collection \mathcal{P} of partitions of $[N]$, let $a(\mathcal{P})$ denote the average number of blocks among the partitions belonging to \mathcal{P} . That is,

$$a(\mathcal{P}) = \frac{\sum_{P \in \mathcal{P}} |P|}{|\mathcal{P}|}.$$

For example, it is known that $a(\Omega_N) = \frac{B_{N+1}}{B_N} - 1$ (see Appendix B). The conflict graph is a negative dependency graph for collections \mathcal{P} that are “coarse” enough in the sense that the average number of blocks of its corresponding canonical events is smaller than the average over all partitions.

The following lemma will be useful for establishing negative dependency graphs in collections of coarse partial partitions. Notice the conclusion is for the *entire* collection \mathcal{P} of partial partitions and says nothing about *subcollections*.

Lemma 3.1. *Let \mathcal{P} be a collection of partial partitions of $[N]$. If*

$$a \left(\bigcap_{P \in \mathcal{P}} \overline{A_P} \right) \geq \frac{B_{N+1}}{B_N} - 1$$

or, equivalently,

$$a \left(\bigcup_{P \in \mathcal{P}} A_P \right) \leq \frac{B_{N+1}}{B_N} - 1,$$

then

$$\Pr_N \left(\bigwedge_{P \in \mathcal{P}} \overline{A_P} \right) \leq \Pr_{N+1} \left(\bigwedge_{P \in \mathcal{P}} \overline{A_P} \right).$$

Proof. For convenience, let $\mathcal{P}^N = \bigcap_{P \in \mathcal{P}} \overline{A_P^N}$ and $\mathcal{P}^{N+1} = \bigcap_{P \in \mathcal{P}} \overline{A_P^{N+1}}$, where the superscript denotes the size of the underlying set. Given any partition $P \in \mathcal{P}^N$, one can form a partition belonging to \mathcal{P}^{N+1} either by adjoining the block $\{N+1\}$ to P or by introducing the element $N+1$ to any existing block of P . From this we see

$$\begin{aligned} |\mathcal{P}^{N+1}| &\geq \sum_{P \in \mathcal{P}^N} (|P| + 1) \\ &= |\mathcal{P}^N| + \sum_{P \in \mathcal{P}^N} |P| \\ &= \left(1 + \frac{\sum_{P \in \mathcal{P}^N} |P|}{|\mathcal{P}^N|} \right) |\mathcal{P}^N| \\ &= (1 + a(\mathcal{P}^N)) |\mathcal{P}^N| \\ &\geq \left(1 + \frac{B_{N+1}}{B_N} - 1 \right) |\mathcal{P}^N| \\ &= B_{N+1} \Pr(\mathcal{P}^N), \end{aligned}$$

so $\Pr(\mathcal{P}^{N+1}) \geq \Pr(\mathcal{P}^N)$. □

A heavy-handed way to ensure sufficient coarseness is to require *every* block appearing in \mathcal{P} be large enough. It turns out a block size of $\log N$ is large enough, and empirical data suggests this is as small as possible. The proof of this fact relies on Canfield's formulation [9] of Moser and Wyman's asymptotic expression for the Bell numbers [23].

Lemma 3.2 (Canfield 1995). *Let r be the unique real solution of the equation $re^r = N$ (that is, $r = \text{LambertW}_0(N)$). The identity*

$$B_{N+h} = \frac{(N+h)!}{r^{N+h}} \cdot \frac{\exp(e^r - 1)}{(2\pi B)^{1/2}} \left(1 + \frac{P_0 + hP_1 + h^2P_2}{e^r} + \frac{Q_0 + hQ_1 + h^2Q_2 + h^3Q_3 + h^4Q_4}{e^{2r}} + O(e^{-3r}) \right)$$

holds uniformly for $h = O(\log N)$ as N tends to infinity, where

$$\begin{aligned}
B &= (r^2 + r)e^r, \\
P_0 &= -\frac{2r^4 + 9r^3 + 16r^2 + 6r + 2}{24r(r+1)^3}, \\
P_1 &= -\frac{r^2 + 3r + 1}{2r(r+1)^2}, \\
P_2 &= -\frac{1}{2r(r+1)}, \\
Q_0 &= \frac{4 + 24r + 100r^2 - 636r^3 - 588r^4 - 384r^5 - 143r^6 - 12r^7 + 4r^8}{1152r^2(r+1)^6}, \\
Q_1 &= \frac{6 + 32r + 56r^2 + 135r^3 + 101r^4 + 37r^5 + 6r^6}{48r^2(r+1)^5}, \\
Q_2 &= \frac{20 + 90r + 190r^2 + 105r^3 + 20r^4}{48r^2(r+1)^4}, \\
Q_3 &= \frac{5 + 15r + 5r^2}{12r^2(r+1)^3}, \text{ and} \\
Q_4 &= \frac{1}{8r^2(r+1)^2}.
\end{aligned}$$

The factor of $r = \text{LambertW}(N)$ proves troublesome in the asymptotic analysis, so we make use of the following bounds [16].

Lemma 3.3 (Hoorfar, Hassani 2008). *For every $x \geq e$, we have*

$$\log x - \log \log x + \frac{1}{2} \frac{\log \log x}{\log x} \leq \text{LambertW}(x) \leq \log x - \log \log x + \frac{e}{e-1} \frac{\log \log x}{\log x}$$

with equality only for $x = e$.

Lemma 3.4. *Set $k = \lceil \log N \rceil$ and let c be any constant. The inequality*

$$\frac{B_{N+1}}{B_N} - \frac{B_{N+1-c\ell}}{B_{N-c\ell}} > 1$$

holds for all sufficiently large N and any $\ell \geq k$.

Proof. The case $\ell = k$ is handled by a Maple worksheet [22] making use of the Moser-Wyman expansion of the Bell numbers and the Hoorfar-Hassani bound on the LambertW function. The sequence $\frac{B_{N+1-c\ell}}{B_{N-c\ell}}$ is nonincreasing in ℓ [12], which gives the desired conclusion for any $\ell \geq k$. \square

Theorem 3.5. *Set $k = \lceil \log N \rceil$ and let c be any constant. If \mathcal{P} is a collection (possibly depending on N) of partial partitions of a sufficiently large set $[N]$ such that every partial partition $P \in \mathcal{P}$ contains at most c blocks and every block contains at least k elements, then the conflict graph for $\{A_P \mid P \in \mathcal{P}\}$ is a negative dependency graph.*

Proof. Let $P \in \mathcal{P}$ and let \mathcal{Q} be any subcollection of \mathcal{P} that does not conflict with P .

Our ultimate goal is to establish

$$\Pr_N \left(\bigwedge_{Q \in \mathcal{Q}} \overline{A_Q} \mid A_P \right) \leq \Pr_N \left(\bigwedge_{Q \in \mathcal{Q}} \overline{A_Q} \right). \quad (3.1)$$

To that end, define $\mathcal{Q}_P = \{Q \setminus P \mid Q \in \mathcal{Q}\}$. Let $\|P\|$ denote the number of *ground elements* (not blocks) appearing in the partial partition P and assume, without loss of generality, that these elements are $N - \|P\| + 1, \dots, N$. Since P does not conflict with \mathcal{Q} , any block of P is either identical to or disjoint from any block of $Q \in \mathcal{Q}$. Hence, the ground elements appearing in partial partitions belonging to \mathcal{Q}_P are elements of the set $[N - \|P\|]$.

Now, if $\emptyset \in \mathcal{Q}_P$, then there is nothing to show (the lefthand side of Inequality 3.1 evaluates to zero). Otherwise, let $\{c_{Q_P} \mid Q_P \in \mathcal{Q}_P\}$ be nonnegative weights such that

$$\sum_{Q_P \in \mathcal{Q}_P} c_{Q_P} = 1$$

and

$$a \left(\bigcup_{Q_P \in \mathcal{Q}_P} A_{Q_P} \right) = \sum_{Q_P \in \mathcal{Q}_P} c_{Q_P} a(A_{Q_P}).$$

Fix an arbitrary partial partition $Q_P \in \mathcal{Q}_P$ and denote its blocks by B_1, \dots, B_j . Each of the B_i contains at least k elements. In what follows, we use a superscript N on a collection of partial partitions to denote the number of ground elements over which the partitions are to be formed. Repeated application of Lemma 3.4 and the

monotonicity of the sequence $\frac{B_{N+1}}{B_N}$ give

$$\begin{aligned}
a(A_{Q_P}^N) &= j + a(\Omega_{N-\|Q_P\|}^{N-\|Q_P\|}) \\
&= j - 1 + \frac{B_{N+1-\|Q_P\|}}{B_{N-\|Q_P\|}} \\
&< j - 2 + \frac{B_{N+1-\|Q_P\|\{B_1\}}}{B_{N-\|Q_P\|\{B_1\}}} \\
&< j - 3 + \frac{B_{N+1-\|Q_P\|\{B_1, B_2\}}}{B_{N-\|Q_P\|\{B_1, B_2\}}} \\
&\vdots \\
&< j - (j + 1) + \frac{B_{N+1-\|Q_P\|\{B_1, \dots, B_j\}}}{B_{N-\|Q_P\|\{B_1, \dots, B_j\}}} \\
&= \frac{B_{N+1}}{B_N} - 1.
\end{aligned}$$

for all sufficiently large N .

As Q_P was selected arbitrarily from \mathcal{Q}_P , we can write

$$\begin{aligned}
a\left(\bigcup_{Q_P \in \mathcal{Q}_P} A_{Q_P}\right) &= \sum_{Q_P \in \mathcal{Q}_P} c_{Q_P} a(A_{Q_P}) \\
&< \sum_{Q_P \in \mathcal{Q}_P} c_{Q_P} \left(\frac{B_{N+1}}{B_N} - 1\right) \\
&= \frac{B_{N+1}}{B_N} - 1.
\end{aligned}$$

Finally, we return to the negative dependency inequality 3.1. Repeated application

of Lemma 3.1 gives

$$\begin{aligned}
\Pr_N \left(\bigwedge_{Q \in \mathcal{Q}} \overline{A_Q} \mid A_P \right) &= \Pr_{N-\|P\|} \left(\bigwedge_{Q_P \in \mathcal{Q}_P} \overline{A_{Q_P}} \right) \\
&\leq \Pr_{N+1-\|P\|} \left(\bigwedge_{Q_P \in \mathcal{Q}_P} \overline{A_{Q_P}} \right) \\
&\leq \Pr_{N+2-\|P\|} \left(\bigwedge_{Q_P \in \mathcal{Q}_P} \overline{A_{Q_P}} \right) \\
&\vdots \\
&\leq \Pr_N \left(\bigwedge_{Q_P \in \mathcal{Q}_P} \overline{A_{Q_P}} \right) \\
&\leq \Pr_N \left(\bigwedge_{Q \in \mathcal{Q}} \overline{A_Q} \right).
\end{aligned}$$

□

3.5 FAILED ATTEMPT VIA INJECTION

Let M be a partial partition and \mathcal{M} be a collection of partial partitions that does not conflict with M . The correlation inequality presented in Section 1.3 would ask us to verify

$$\Pr(A_M) \Pr \left(\bigvee_{L \in \mathcal{M}} A_L \right) \leq \Pr \left(A_M \wedge \left(\bigvee_{L \in \mathcal{M}} A_L \right) \right),$$

which is equivalent to

$$|A_M| \left| \bigcup_{L \in \mathcal{M}} A_L \right| \leq B_N \left| A_M \cap \left(\bigcup_{L \in \mathcal{M}} A_L \right) \right|.$$

A radically different approach to verifying this inequality would be to establish an injection f from the set

$$\left\{ (P, Q) \mid P \in A_M, Q \in \bigcup_{L \in \mathcal{M}} A_L \right\}$$

(which has cardinality equal to the left-hand side) into a set of cardinality no larger than the right-hand side. We describe below one such attempt and demonstrate

where it comes short of achieving the desired goal. Notice that the inequality cannot possibly be verified in full generality, since there are counterexamples for every N . It is the author's hope that a similar approach can be made to work for appropriately restricted collections of partial partitions.

Given a partial partition R , let $\text{supp}(R)$ denote the collection of all ground elements appearing in R . When speaking of the restriction of a partial partition to another, we may write $R \upharpoonright_S$ to mean $R \upharpoonright_{\text{supp}(S)}$. Empty blocks arising as a result of restriction are discarded.

Let $C = [N] \setminus \text{supp}(M)$ ("C" for "complement"). We will say a block B of a partial partition is

- a *type M* block whenever $B \upharpoonright_M = B \upharpoonright_{[N]}$,
- a *type C* block whenever $B \upharpoonright_C = B \upharpoonright_{[N]}$, and
- a *mixed* block if it is neither type M nor type C.

In other words, if we first restrict our attention only to elements of $[N]$ (there will be other kinds of elements to consider later), then type M blocks contain only elements of M , type C blocks contain no elements of M , and mixed blocks contain a mixture of the two kinds of elements.

Let now $P \in A_M$ and $Q \in \bigcup_{L \in \mathcal{M}} A_L$ be given and let $(P, Q) \mapsto (P', Q')$ under f .

To obtain P' from P :

1. Remove all type M blocks from P . (Since $P \in A_M$, this is just $P \setminus M$.)
2. Insert all type M blocks of Q into P .
3. For each mixed type block B of Q , let $\ell_B = \min(B \upharpoonright_C)$ and insert into P the block $(B \upharpoonright_M) \cup \{\widehat{\ell}_B\}$, where $\widehat{\ell}_B$ denotes a duplicate, but distinguishable, copy of ℓ_B .

For the purposes of the function f , duplicate elements will always correspond to elements of C , so we define $\widehat{C} = \{\widehat{\ell} \mid \ell \in C\}$ as the collection of possible duplicate elements.

To obtain Q' from Q :

1. Remove all type M blocks from Q .
2. Replace each mixed type block B of Q with $B \upharpoonright_C$.
3. Insert all blocks of M into Q .

(As an example of the behavior of f , consider $M = 1|23$, $P = 1|23|4|56$, and $Q = 13|26|45$. The procedure described above gives $P' = 13|\widehat{26}|4|56$ and $Q' = 1|23|45|6$.)

Lemma 3.6. *The function f is injective.*

Proof. Let $f(P_1, Q_1) = (P'_1, Q'_1) = (P'_2, Q'_2) = f(P_2, Q_2)$. We show $(P_1, Q_1) = (P_2, Q_2)$.

Since $P'_1 = P'_2$, it follows that $P'_1 \upharpoonright_C = P'_2 \upharpoonright_C$, and so

$$\begin{aligned} P_1 &= P'_1 \upharpoonright_C \cup M \\ &= P'_1 \upharpoonright_C \cup M \\ &= P_2. \end{aligned}$$

Since $Q'_1 = Q'_2$, it follows that $Q'_1 \setminus M = Q'_2 \setminus M$, and so $Q_1 \upharpoonright_C = Q_2 \upharpoonright_C$. Now, every type M block of $P'_1 = P'_2$ is also present in both Q_1 and Q_2 . Every mixed block B of $P'_1 = P'_2$ contains an element $\widehat{\ell} \in \widehat{C}$, and so there is $B_1 \in Q_1$ and $B_2 \in Q_2$ such that ℓ belongs to both B_1 and B_2 and $B_1 \upharpoonright_M = B_2 \upharpoonright_M = B \setminus \{\widehat{\ell}\}$. Since $Q_1 \upharpoonright_C = Q_2 \upharpoonright_C$, we know also that $B_1 \upharpoonright_C = B_2 \upharpoonright_C$, and thus $B_1 = B_2$. As this holds for all mixed blocks of $P'_1 = P'_2$, we have $Q_1 = Q_2$. \square

It remains to verify the inequality

$$|A_M| \left| \bigcup_{L \in \mathcal{M}} A_L \right| \leq B_N \left| A_M \cap \left(\bigcup_{L \in \mathcal{M}} A_L \right) \right|.$$

Since f is injective, we have

$$\left| \left\{ (P, Q) \mid P \in A_M, Q \in \bigcup_{L \in \mathcal{M}} A_L \right\} \right| \leq \left| \left\{ f((P, Q)) \mid P \in A_M, Q \in \bigcup_{L \in \mathcal{M}} A_L \right\} \right|.$$

Observe that, if $(P, Q) \mapsto (P', Q')$ under f , then $Q' \in A_M \cap (\bigcup_{L \in \mathcal{M}} A_L)$, since M does not conflict with \mathcal{M} . Hence, there are at most $|A_M \cap (\bigcup_{L \in \mathcal{M}} A_L)|$ possible Q' . The desired inequality would follow by showing, for each fixed $Q'_0 \in A_M \cap (\bigcup_{L \in \mathcal{M}} A_L)$, we have

$$|\{P' \mid (P', Q'_0) \in \text{im}(f)\}| \leq B_N.$$

Unfortunately, this inequality is not true. Fixing Q'_0 only restricts what elements of \widehat{C} can be applied as labels in a P' under f . The available labels are precisely the minimum elements of type C blocks of Q'_0 . Since Q'_0 can have as many as $N - 1$ type C blocks, the fact that Q'_0 is fixed can be of little help.

For a concrete example, let $M = \{\{1\}\}$, \mathcal{M} contain only the partial partition $\{\{2\}\}$, and Q'_0 be a partition of $[N]$ into singletons. The partition $P \in A_M$ may be any partition having the singleton block $\{1\}$. The partition $Q \in A_{\{\{2\}\}}$ that would map to Q'_0 under f may be the partition of $[N]$ into singletons or any partition that is singletons except for the block $\{1, i\}$ for $i \notin \{1, 2\}$. If Q is of the former type, then the resulting P' may be any partition having the singleton block $\{1\}$, of which there are B_{N-1} possibilities. If Q is of the latter type, then P' may be any partition containing the block $\{1, \widehat{i}\}$ for $i \neq 2$, of which there are $(N - 2)B_{N-1}$ possibilities. In total, this gives $(N - 1)B_{N-1}$ possible P' that may appear with Q'_0 under f . Since $\frac{B_{N-1}}{B_N} \sim \frac{r}{N}$ (where r is the solution to the equation $N = re^r$) [11], we have $\frac{(N-1)B_{N-1}}{B_N} \sim r$, which grows without bound in N . Hence,

$$|\{P' \mid (P', Q'_0) \in \text{im}(f)\}| > B_N$$

for large N . (In fact, it is already true for $N = 5$.)

3.6 FURTHER RESEARCH

In light of the class of counterexamples in Chapter B, one cannot guarantee even asymptotically that the conflict graph for an unrestricted collection $\{A_P \mid P \in \mathcal{P}\}$ of canonical events is a negative dependency graph. The known counterexamples seem to rely on the fact that the set of ground elements of partial partitions in the collection \mathcal{P} is quite a large subset of $[N]$. Empirical evidence suggests that a negative dependency graph always exists when this is not the case.

Conjecture 3.7. Let \mathcal{P} be a collection of partial partitions of $[N_0]$. For sufficiently large N , the conflict graph for $\{A_P \mid P \in \mathcal{P}\}$ is a negative dependency graph in the probability space Ω_N .

While possibly true, this conjecture carries with it the unfortunate restriction that only members of Ω_N whose support lies in $[N_0]$ can be forbidden via the lopsided local lemma.

In Chapter 5, we make use of the lopsided local lemma to derive asymptotics for the number of hypergraphs avoiding small cycles. It was the author's intent to derive a similar expression for the number of partitions avoiding small blocks. For this application, we are interested only in the collection \mathcal{P} defined by

$$\mathcal{P} = \{\{B\} : B \subset [N], |B| \leq m\}$$

for some fixed (or perhaps slowly growing) integer m . Showing that the conflict graph for this collection is a negative dependency graph would be an important step toward proving the following conjecture about partitions having no small blocks.

Conjecture 3.8. The number of partitions of $[N]$ whose smallest block is of size m is asymptotic to

$$B_N \exp\left(-\sum_{k=1}^{m-1} \binom{N}{k} \frac{B_{N-k}}{B_N}\right)$$

(assuming restricted growth of m as a function of N).

The conjecture is correct when $m = 2$, for which it claims the number of singleton-free partitions of $[N]$ is asymptotic to $B_N \exp\left(-N \frac{B_{N-1}}{B_N}\right)$. The number of such partitions can be expressed exactly as $\sum_{i=1}^N (-1)^{N-i} B_i$ using Lemma B.2. Dividing both sides by B_N , we show $\frac{1}{B_N} \sum_{i=1}^N (-1)^{N-i} B_i$ and $\exp\left(-N \frac{B_{N-1}}{B_N}\right)$ converge to the same value. For both calculations, we make use of Canfield's expansion for the Bell numbers.

For the former,

$$\begin{aligned} \frac{1}{B_N} \sum_{i=1}^N (-1)^{N-i} B_i &= \frac{B_{N-1}}{B_N} \left(\sum_{i=1}^N (-1)^{N-i} \frac{B_i}{B_{N-1}} \right) \\ &= \frac{B_{N-1}}{B_N} \left(1 + O\left(\frac{B_{N-2}}{B_{N-1}}\right) \right) \\ &= \frac{B_{N-1}}{B_N} \left(1 + O\left(\frac{r}{n}\right) \right) \\ &= \frac{r}{n} (1 + o(1)). \end{aligned}$$

For the latter,

$$\begin{aligned} \exp\left(-N \frac{B_{N-1}}{B_N}\right) &= \exp\left(-r \cdot \frac{1}{1 + \frac{\text{poly}(r)}{e^r}}\right) \\ &= \exp\left(-r \left(1 - O\left(\frac{\text{poly}(r)}{e^r}\right)\right)\right), \end{aligned}$$

since $\frac{1}{1 \pm x} \sim 1 \mp x$ as $x \rightarrow 0$. Continuing,

$$\begin{aligned} \exp\left(-r \left(1 - O\left(\frac{\text{poly}(r)}{e^r}\right)\right)\right) &= \exp\left(-r + O\left(\frac{\text{poly}(r)}{e^r}\right)\right) \\ &= \exp(-r) \exp\left(O\left(\frac{\text{poly}(r)}{e^r}\right)\right) \\ &= \frac{r}{n} (1 + o(1)). \end{aligned}$$

CHAPTER 4

NEGATIVE DEPENDENCY GRAPHS FOR SPANNING TREES

The probability spaces we defined for hypergraph matchings and set partitions have in common that a partial object (partial matching and partial partition, respectively) does not conflict with any maximal object (maximal matching and partition, respectively) in its corresponding canonical event. This useful property lends itself immediately to the use of induction, since every maximal object splits nicely into the partial object and its extension. The proofs presented on hypergraph matchings rely on the fact that we can extend a partial matching M to a maximal one by taking the disjoint union of M together with *any* maximal matching of the vertices missed by M . Similarly, we extend a partial partition P to a full partition by taking the disjoint union of P together with *any* partition of the ground elements missed by P .

In this section, we define a natural space in which a partial object (forest) conflicts with *every* maximal object (spanning tree) in its canonical event. As a result, the maximal objects do not split nicely into the disjoint union of a partial object together with its extension.

4.1 INTRODUCTION

A **cycle** in a simple graph is a sequence of vertices and edges $v_1, e_1, v_2, e_2, \dots, v_k, e_k$ ($k \geq 3$) in which $e_i = \{v_i, v_{i+1}\}$ for each i (where we understand v_{k+1} to be v_1). A **forest** is a cycle-free graph, and a **tree** is a connected forest. (Notice each connected

component of a forest is a tree.) Given an underlying graph G , we say a tree T is a **spanning tree** of G whenever the vertex set of T coincides with the vertex set of G .

Let Ω be the uniform probability space containing all spanning trees of K_N , the complete graph on N vertices. For any forest F contained in K_N , define the **canonical event** A_F to be the collection of all spanning trees of K_N containing F . That is,

$$A_F = \{T \in \Omega \mid F \subseteq T\}.$$

Two forests F_1 and F_2 (both contained in K_N) **conflict** whenever there are trees $T_1 \subseteq F_1$ and $T_2 \subseteq F_2$ such that T_1 and T_2 are neither identical nor disjoint.

Finally, let \mathcal{F} be any collection of forests contained in K_N . The **conflict graph** for the collection $\{A_F \mid F \in \mathcal{F}\}$ is a simple graph whose vertex set is \mathcal{F} . Two forests are adjacent in this graph if and only if they conflict.

4.2 EXAMPLE CONFLICT GRAPH

Take K_4 , the complete graph on four vertices, to be the underlying graph. Figure 4.2 depicts a forest contained in K_4 composed of two disjoint edges. The canonical event for this forest consists of the four spanning trees of K_4 that contain it as a subgraph. Notice *every* spanning tree in the canonical event conflicts with the forest that defined it.

Unrelated to the previous example, consider the three forests D , E , and F pictured in Figure 4.2 with K_8 as the underlying graph. The forests D and F conflict, since the leftmost component of D is neither identical to nor disjoint from the single component of F . Similarly, the forests E and F conflict in the lower component of E . The forests D and E do not conflict, since any two components (one from D and one from E) are either identical or disjoint. The conflict graph for the associated canonical events is therefore the graph with vertex set $\{D, E, F\}$ and edge set $\{DF, EF\}$.

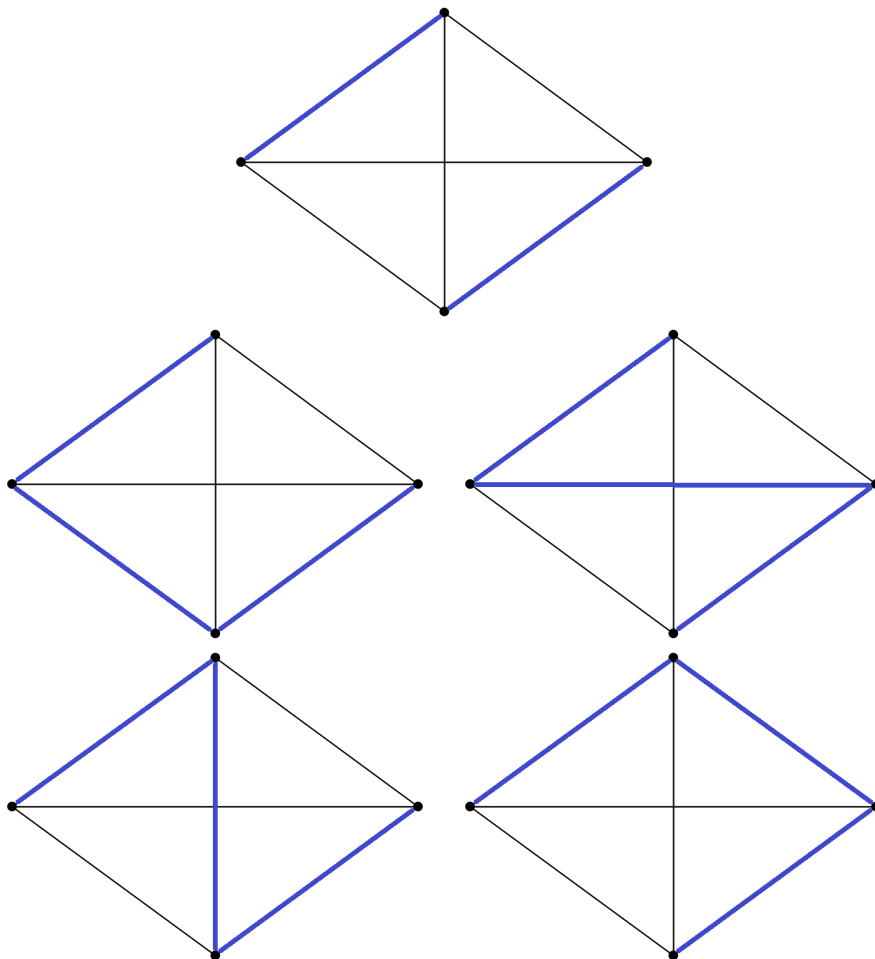


Figure 4.1 Forest in K_4 and its canonical event.

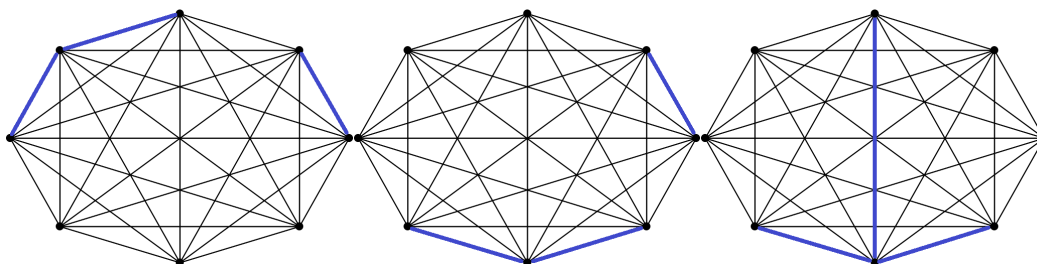


Figure 4.2 Forests D , E , and F , respectively.

4.3 RESULTS

The main result of this section is that the conflict graph for *any* collection of forests is a negative dependency graph.¹

Theorem 4.1. *Let \mathcal{F} be any collection of forests in K_N . The conflict graph for the collection $\{A_F \mid F \in \mathcal{F}\}$ of canonical events is a negative dependency graph.*

We first prove two lemmata. The first counts the number of spanning trees of K_N that contain a given forest (i.e. the size of a canonical event). When the forest is edgeless, the lemma gives the familiar result of Cayley [29] that the number of spanning trees of K_N is N^{N-2} .

Lemma 4.2. *Let F be a forest in K_N with connected components T_1, \dots, T_k on t_1, \dots, t_k vertices, respectively. The number of spanning trees of K_N containing F is given by*

$$N^{N-2} \prod_{i=1}^k \frac{t_i}{N^{t_i-1}}.$$

Proof. Contracting the components of F to single vertices reduces the number of vertices by $t_i - 1$ for each component T_i . Setting $N' = N - \sum_{i=1}^k (t_i - 1)$, this contraction transforms a spanning tree of K_N into a spanning tree of $K_{N'}$. Label the vertices of $K_{N'}$ by v_i , where the vertices v_1, \dots, v_k resulted from the contraction of the components T_1, \dots, T_k , respectively, and the vertices $v_{k+1}, \dots, v_{N'}$ were not covered by F in K_N .

Menon's theorem [18] states that the number of spanning trees of $K_{N'}$ in which each vertex v_i has degree r_i is given by the multinomial coefficient

$$\binom{N' - 2}{r_1 - 1, \dots, r_{N'} - 1}.$$

We must also determine how many spanning trees of K_N contract to a fixed spanning tree T' of $K_{N'}$. Let us fix our attention on a single component, say T_1 , of F .

¹Results in this section are joint work with Lu and Székely.

If the degree of v_1 in $K_{N'}$ is to be r_1 , then there must be r_1 edges from $K_N \setminus T_1$ into T_1 . Each of the r_1 edges has t_1 choices for its endpoint in T_1 , resulting in $t_1^{r_1}$ possible assignments of the edges. Any two such assignments are indistinguishable after the contraction of T_1 . Multiplying now across all components, we find there are $\prod_{i=1}^k t_i^{r_i}$ spanning trees of K_N whose contraction results in T' .

The previous two paragraphs show that the number of spanning trees of K_N containing F is given by

$$\sum_{r_1 + \dots + r_{N'} = 2(N'-1)} \binom{N'-2}{r_1-1, \dots, r_{N'}-1} \prod_{i=1}^k t_i^{r_i}.$$

Assigning $r_i \leftarrow r_i + 1$ for each of the indices of summation, we write instead

$$\begin{aligned} & \sum_{r_1 + \dots + r_{N'} = N'-2} \binom{N'-2}{r_1, \dots, r_{N'}} \prod_{i=1}^k t_i^{r_i+1} \\ &= \prod_{i=1}^k t_i \sum_{r_1 + \dots + r_{N'} = N'-2} \binom{N'-2}{r_1, \dots, r_{N'}} \prod_{i=1}^k t_i^{r_i}. \end{aligned}$$

Let now $t_{k+1} = \dots = t_{N'} = 1$. Invoking the Multinomial Theorem, the expression above becomes

$$\begin{aligned} & \prod_{i=1}^k t_i \sum_{r_1 + \dots + r_{N'} = N'-2} \binom{N'-2}{r_1, \dots, r_{N'}} \prod_{i=1}^{N'} t_i^{r_i} \\ &= \prod_{i=1}^k t_i \left(\sum_{i=1}^{N'} t_i \right)^{N'-2} \\ &= N^{N'-2} \prod_{i=1}^k t_i \\ &= N^{N'-2} \prod_{i=1}^k \frac{t_i}{N^{t_i-1}}. \end{aligned}$$

□

Two forests are in **strong conflict** whenever they are not vertex disjoint. As a preliminary version of Theorem 4.1, we show that the strong conflict graph (i.e. the conflict graph that uses the definition of strong conflict) is a negative dependency graph.

Lemma 4.3. *Let \mathcal{F} be any collection of forests in K_N . The strong conflict graph for the collection $\{A_F \mid F \in \mathcal{F}\}$ of canonical events is a negative dependency graph.*

Proof. Let us be given a forest $F \in \mathcal{F}$ and any subcollection \mathcal{G} of \mathcal{F} containing forests that are not in strong conflict with F . We seek to establish the correlation inequality formulation of negative dependence introduced in Section 1.3, namely

$$\Pr(A_F) \Pr\left(\bigvee_{G \in \mathcal{G}} A_G\right) \leq \Pr\left(A_F \wedge \bigvee_{G \in \mathcal{G}} A_G\right). \quad (4.1)$$

(In fact, we will prove equality.) By inclusion-exclusion, we have

$$\Pr\left(\bigvee_{G \in \mathcal{G}} A_G\right) = \sum_{\substack{\mathcal{H} \subseteq \mathcal{G} \\ \mathcal{H} \neq \emptyset}} (-1)^{|\mathcal{H}|-1} \Pr\left(\bigwedge_{H \in \mathcal{H}} A_H\right)$$

and

$$\Pr\left(A_F \wedge \bigvee_{G \in \mathcal{G}} A_G\right) = \sum_{\substack{\mathcal{H} \subseteq \mathcal{G} \\ \mathcal{H} \neq \emptyset}} (-1)^{|\mathcal{H}|-1} \Pr\left(A_F \wedge \bigwedge_{H \in \mathcal{H}} A_H\right).$$

We claim

$$\Pr(A_F) \Pr\left(\bigwedge_{H \in \mathcal{H}} A_H\right) = \Pr\left(A_F \wedge \bigwedge_{H \in \mathcal{H}} A_H\right) \quad (4.2)$$

for every nonempty subset \mathcal{H} of \mathcal{G} , which will establish the correlation inequality 4.1.

The event $\bigwedge_{H \in \mathcal{H}} A_H$ consists precisely of spanning trees that contain $\bigcup_{H \in \mathcal{H}} H$ as a subgraph. Denote this union by H' , so that $\bigwedge_{H \in \mathcal{H}} A_H = A_{H'}$.

If H' contains a cycle, then the corresponding event is empty, and both sides of Equation 4.2 evaluate to zero.

Assume now that H' is a forest. Let f_1, \dots, f_r denote the sizes of the connected components of F and similarly h_1, \dots, h_s for H' . From Lemma 4.2, we have

$$\Pr(A_F) = \frac{N^{N-2} \prod_{i=1}^r \frac{f_i}{N^{f_i-1}}}{N^{N-2}} = \prod_{i=1}^r \frac{f_i}{N^{f_i-1}}$$

and

$$\Pr(A_{H'}) = \frac{N^{N-2} \prod_{i=1}^s \frac{h_i}{N^{h_i-1}}}{N^{N-2}} = \prod_{i=1}^s \frac{h_i}{N^{h_i-1}}.$$

Since F is not in strong conflict with any of the forests belonging to \mathcal{H} , we know F is vertex disjoint from H' . Thus, $F \cup H'$ has components of size $f_1, \dots, f_r, h_1, \dots, h_s$. Again by Lemma 4.2, we have

$$\Pr(A_F \wedge A_{H'}) = \frac{N^{N-2} \prod_{i=1}^r \frac{f_i}{N^{f_i-1}} \prod_{j=1}^s \frac{h_j}{N^{h_j-1}}}{N^{N-2}} = \prod_{i=1}^r \frac{f_i}{N^{f_i-1}} \prod_{j=1}^s \frac{h_j}{N^{h_j-1}},$$

which establishes Equation 4.2. \square

We now return to the proof of Theorem 4.1.

Proof of Theorem 4.1. Let us be given a forest $F \in \mathcal{F}$ and any subcollection \mathcal{G} of \mathcal{F} containing forests that are not in conflict with F . We seek to establish yet another formulation of negative dependence. Starting with the correlation inequality introduced in Section 1.3, we may derive the equivalent expression

$$\begin{aligned} \Pr(A_F) \Pr\left(\bigvee_{G \in \mathcal{G}} A_G\right) &\leq \Pr\left(A_F \wedge \bigvee_{G \in \mathcal{G}} A_G\right) \\ \Pr(A_F) \left[1 - \Pr\left(\bigwedge_{G \in \mathcal{G}} \overline{A_G}\right)\right] &\leq \Pr\left(A_F \wedge \bigvee_{G \in \mathcal{G}} A_G\right) \\ \Pr(A_F) - \Pr(A_F) \Pr\left(\bigwedge_{G \in \mathcal{G}} \overline{A_G}\right) &\leq \Pr\left(A_F \wedge \bigvee_{G \in \mathcal{G}} A_G\right) \\ \Pr(A_F) - \Pr\left(A_F \wedge \bigvee_{G \in \mathcal{G}} A_G\right) &\leq \Pr(A_F) \Pr\left(\bigwedge_{G \in \mathcal{G}} \overline{A_G}\right) \\ \Pr\left(A_F \wedge \bigwedge_{G \in \mathcal{G}} \overline{A_G}\right) &\leq \Pr(A_F) \Pr\left(\bigwedge_{G \in \mathcal{G}} \overline{A_G}\right). \end{aligned}$$

If F is vertex disjoint from every $G \in \mathcal{G}$, then Lemma 4.3 finishes the proof. Otherwise, write $G_F = G \setminus F$ for each $G \in \mathcal{G}$. Write also $\mathcal{G}_F = \{G_F \mid G \in \mathcal{G}\}$. Since F does not conflict with any $G \in \mathcal{G}$, it follows that F is vertex disjoint from every

member of \mathcal{G}_F . Now,

$$\begin{aligned}
\Pr\left(A_F \wedge \bigwedge_{G \in \mathcal{G}} \overline{A_G}\right) &= \Pr\left(A_F \wedge \bigwedge_{G_F \in \mathcal{G}_F} \overline{A_{G_F}}\right) \\
&= \Pr(A_F) \Pr\left(\bigwedge_{G_F \in \mathcal{G}_F} \overline{A_{G_F}}\right) && \text{(by Lemma 4.3)} \\
&\leq \Pr(A_F) \Pr\left(\bigwedge_{G \in \mathcal{G}} \overline{A_G}\right).
\end{aligned}$$

□

Even though we showed equality in 4.1, this is not enough to conclude the strong conflict graph is a dependency graph. To establish the *mutual independence* required for a dependency graph, we still need to show that A_F is independent of the *event algebra* generated by $\{A_H \mid H \in \mathcal{H}\}$, which would allow events such as $\overline{A_H}$ to appear in 4.1. This can be accomplished with the following lemma.

Lemma 4.4. *Let \mathcal{A} be a collection of events and let A be any event belonging to \mathcal{A} .*

If

$$\Pr\left(A \wedge \bigwedge_{B \in \mathcal{B}} B\right) = \Pr(A) \prod_{B \in \mathcal{B}} \Pr(B)$$

for any subcollection \mathcal{B} of \mathcal{A} , then A is independent of the event algebra generated by \mathcal{A} .

Proof. Since we know A is independent of any subset of \mathcal{A} , it remains to show that we still have independence even when some events from \mathcal{A} are complemented.

We proceed by induction on $|\mathcal{B}|$. If \mathcal{B} contains a single event B , then

$$\begin{aligned}
\Pr(A \wedge \overline{B}) &= \Pr(A) - \Pr(A \wedge B) \\
&= \Pr(A) - \Pr(A) \Pr(B) \\
&= \Pr(A) (1 - \Pr(B)) \\
&= \Pr(A) \Pr(\overline{B}).
\end{aligned}$$

Assume now that A is independent of any subset of size j of the event algebra generated by \mathcal{B} . For convenience, we show A is independent of the event algebra generated by $\{B_1, \dots, B_{j+1}\}$. To accomplish this, we begin a second induction on the number of complemented events. There is nothing to show if none of the events are complemented. Assume now independence holds when there are at most k complemented events. Let $X_i \in \{B_i, \overline{B_i}\}$ and assume in the following that $X_i = \overline{A_i}$ for at most k of the indices. Invoking both induction hypotheses, we have

$$\begin{aligned}
\Pr\left(A \wedge \overline{B_{j+1}} \wedge \bigwedge_{i=1}^j X_i\right) &= \Pr\left(A \wedge \bigwedge_{i=1}^j X_i\right) - \Pr\left(A \wedge B_{j+1} \wedge \bigwedge_{i=1}^j X_i\right) \\
&= \Pr(A) \prod_{i=1}^j \Pr(X_i) - \Pr(A) \Pr(B_{j+1}) \prod_{i=1}^j \Pr(X_i) \\
&= \Pr(A) \prod_{i=1}^j \Pr(X_i) (1 - \Pr(B_{j+1})) \\
&= \Pr(A) \Pr(\overline{B_{j+1}}) \prod_{i=1}^j \Pr(X_i).
\end{aligned}$$

□

CHAPTER 5

ENUMERATION OF REGULAR UNIFORM HYPERGRAPHS

5.1 CONFIGURATION MODEL FOR HYPERGRAPHS

Our concern with maximal matchings stems from its application to the configuration model of Bollobás [7], which allows one to project a perfect matching of a certain collection of points to a **multihypergraph**. A multihypergraph differs from a hypergraph in that a single edge e may contain repeated vertices and the edge e may itself be repeated in the edge set. For example, the collection $\{\{u, u, v\}, \{u, v, w\}, \{u, v, w\}\}$ can be the edge set of a multihypergraph with vertex set $\{u, v, w\}$, but not a hypergraph. We will be concerned only with configurations that result in an r -regular, s -uniform multihypergraph, which we describe below. Note that in a multihypergraph, we count vertices with multiplicity when defining “regular” and “uniform”.

1. Let U be a set containing Nr distinct minivertices partitioned into N classes each of size r . The i th such class ($i \in [N]$) will be associated with the vertex v_i in the hypergraph H after identifying its elements through a **projection**.
2. Choose uniformly at random a perfect s -matching M of the minivertices in U .
3. Each edge of M is a collection of s minivertices, each corresponding to partition classes with (not necessarily distinct) indices i_1, \dots, i_s . For all such edges of M , add the edge $\{v_{i_1}, v_{i_2}, \dots, v_{i_s}\}$ to the hypergraph H .

Figure 5.1 illustrates a perfect 2-matching on $4 \cdot 3$ minivertices, which projects to a 3-regular, 2-uniform hypergraph (that is, a graph) on four vertices. We will

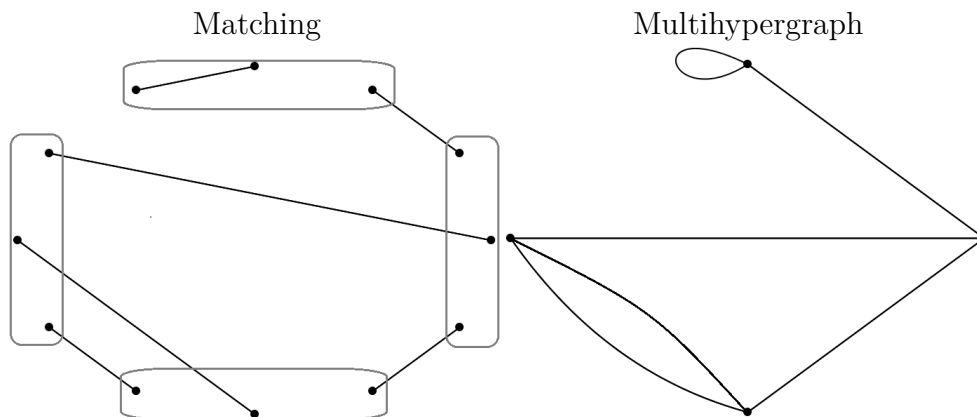


Figure 5.1 Configuration projecting to 3-regular, 2-uniform multihypergraph on four vertices.

later be concerned with configurations that result in a **simple** hypergraph, which is a hypergraph containing no 1-cycles and no pair of edges containing precisely the same vertices (such edges are called repeated edges). The graph in the figure has both a 1-cycle (involving the top vertex) and a pair of repeated edges (involving the left and bottom vertices).

5.2 CYCLES IN HYPERGRAPHS

We define a k -**cycle** in a hypergraph as follows:

- A 1-cycle is a single edge with a repeated vertex.
- A 2-cycle is a pair of edges whose intersection contains at least two vertices.
- For $k \geq 3$, a k -cycle is a collection e_1, \dots, e_k of edges for which there are distinct vertices v_1, \dots, v_k such that $e_i \cap e_{i+1} = \{v_i\}$ for all i (where e_{k+1} is understood to be e_1).

For example, the edges $\{x, a, b, y\}$, $\{y, c, d, z\}$, and $\{z, e, f, x\}$ form a 3-cycle in a 4-uniform hypergraph.

For $k \geq 3$, the definition given above coincides with the usual meaning of “loose cycle”. For $k = 2$, we have a loose cycle whenever the pair of edges intersect in exactly two vertices. We later call such a cycle a “proper” 2-cycle, while the other 2-cycles are “degenerate”.

(Tight cycles and Berge cycles are two other well-studied types of hypergraph cycles [25]. Every tight cycle is a union of 2-cycles, and every Berge cycle either contains a 2-cycle or is itself a k -cycle for some k . As we will see later, the definition given above captures these two notions of cycle so far as regards our present purpose.)

5.3 APPLYING THE LOPSIDED LOCAL LEMMA

For the moment, let r , s , and g be fixed integers. (We will later see that these parameters may be allowed to grow slowly with N .) Fix a set U containing Nr minivertices (with s dividing Nr) partitioned into N classes each of size r . We wish to give an asymptotic expression for the number of r -regular, s -uniform hypergraphs with **girth** at least g , which are the hypergraphs having no k -cycle for $k < g$. (Under this definition of girth, a graph of girth $g \geq 3$ will contain no tight cycles and no Berge cycles that are not themselves k -cycles.) To accomplish this via the lopsided local lemma, let \mathcal{M} contain all matchings whose projection is *precisely* a cycle of size less than g . In the uniform probability space Ω_{Nr} of all perfect s -matchings of U , the expression $\Pr\left(\bigwedge_{M \in \mathcal{M}} \overline{A_M}\right)$ is the probability that a perfect matching does not contain a submatching belonging to the collection \mathcal{M} . From the perspective of the configuration model, we may interpret this instead as the probability that an r -regular, s -uniform multihypergraph on N vertices chosen uniformly at random from among all such multihypergraphs will have girth at least g .

Lu and Székely [21] give a detailed summary of the history of the enumeration of graphs by girth. The count given here introduces three primary advancements over the existing literature. Firstly, we enumerate r -regular, 3-uniform hypergraphs,

while existing results focus on the 2-uniform (that is, the graph) case. Secondly, r and g are allowed to grow slowly with N (as is made precise in the theorem). Finally (and perhaps most importantly), verifying the hypotheses of Theorem 2.3 in this probability space can be accomplished with elementary counting techniques and careful estimation.

In the following result, the exponential factor on the left estimates the probability that a randomly chosen regular uniform multihypergraph has girth at least g , while the quotient of factorials on the right counts exactly the number of such multihypergraphs on N vertices. The proof is deferred to Chapter C.

Theorem 5.1. *In the configuration model, assume $g \geq 1$, $r \geq 3$, and*

$$(2r - 2)^{2g-3} g^3 = o(N). \quad (5.1)$$

The probability that an r -regular, 3-uniform multihypergraph chosen uniformly at random has girth at least g is

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

If $g \geq 3$, then the number of simple r -regular, 3-uniform hypergraphs on N vertices with girth at least g is

$$(1 + o(1)) \exp \left(- \sum_{i=1}^{g-1} \frac{(2r - 2)^i}{2i} \right) \frac{(rN)!}{6^{rN/3} \left(\frac{rN}{3}\right)! (r!)^N}.$$

In fact, letting \mathcal{C} be a subset of $\{3, 4, \dots, g - 1\}$, the number of simple r -regular, 3-uniform hypergraphs whose cycle lengths do not belong to \mathcal{C} is

$$(1 + o(1)) \exp \left(1 - r - (r - 1)^2 - \sum_{i \in \mathcal{C}} \frac{(2r - 2)^i}{2i} \right) \frac{(rN)!}{6^{rN/3} \left(\frac{rN}{3}\right)! (r!)^N}.$$

With extra care in the analysis, the same tools should allow us to enumerate s -uniform multihypergraphs, where s may grow slowly with N .

Conjecture 5.2. In the configuration model, assume $g \geq 1$, $r \geq 3$, $s \geq 2$ and

$$(s-1)^{2g-3}(r-1)^{2g-3}g^3 = o(N). \quad (5.2)$$

The probability that an r -regular, s -uniform multihypergraph chosen uniformly at random has girth at least g is

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

If $g \geq 3$, then the number of simple r -regular, s -uniform hypergraphs on N vertices with girth at least g is

$$(1 + o(1)) \exp \left(- \sum_{i=1}^{g-1} \frac{(s-1)^i (r-1)^i}{2i} \right) \frac{(rN)!}{(s!)^{rN/s} \left(\frac{rN}{s}\right)! (r!)^N}.$$

In fact, letting \mathcal{C} be a subset of $\{3, 4, \dots, g-1\}$, the number of simple r -regular, s -uniform hypergraphs whose cycle lengths do not belong to \mathcal{C} is

$$(1 + o(1)) \exp \left(- \frac{(s-1)(r-1)}{2} - \frac{(s-1)^2(r-1)^2}{4} - \sum_{i \in \mathcal{C}} \frac{(s-1)^i (r-1)^i}{2i} \right) \frac{(rN)!}{(s!)^{rN/s} \left(\frac{rN}{s}\right)! (r!)^N}.$$

5.4 FURTHER RESEARCH

We discussed here only the complete uniform hypergraph, but the configuration model is considerably more flexible. Chapter 6 discusses other hypergraphs for which the conflict graph is always a negative dependency graph. Using an appropriate configuration, one can attempt asymptotic enumeration by girth of any class of graphs for which negative and positive dependency graphs can be found. In particular, it is reasonable to suspect that the complete s -uniform, s -partite hypergraph will support this sort of analysis.

CHAPTER 6

PERFECT MATCHING HOSTS

Throughout this chapter, the ambient probability space Ω will consist only of perfect (rather than simply maximal) matchings of some underlying hypergraph. For a given matching L , the canonical event

$$A_L = \{M \in \Omega \mid L \subseteq M\}$$

will therefore contain only perfect matchings. Recall that the conflict graph for the collection $\{A_M \mid M \in \mathcal{M}\}$ of canonical events is a negative dependency graph, where \mathcal{M} may be any collection of matchings in a complete uniform hypergraph. We call a hypergraph H a **perfect matching host** whenever we can write “ H ” in place of “complete uniform hypergraph” in the previous sentence.

Notice that any hypergraph that has *no* perfect matchings has the property that $|A_L| = 0$ for any partial matching L . For such a hypergraph, the inequality

$$\Pr\left(A_L \mid \bigwedge_{M \in \mathcal{M}} \overline{A_M}\right) \leq \Pr(A_L)$$

is satisfied trivially, since both sides always evaluate to zero. For this reason, we wish only to characterize perfect matching hosts having at least one perfect matching.

6.1 GENERAL RESULTS

A hypergraph is **connected** if, for any two vertices u and v , there is a sequence e_1, \dots, e_k of edges such that $u \in e_1$, $v \in e_k$, and $e_i \cap e_{i+1}$ is nonempty for $i \in [k - 1]$.

When we say that a subhypergraph F of H is a perfect matching host, we mean it is a perfect matching host in its own right (without reference to H) in the probability space containing all perfect matchings of F .

In light of the following theorem, we need only be interested in connected perfect matching hosts.

Lemma 6.1. *A hypergraph is a perfect matching host if and only if each of its connected components is.*

Proof. To establish the leftward implication, let H be the disjoint union of connected components C_1, \dots, C_k , each of which is a perfect matching host. Let L be any partial matching of H and let \mathcal{M} be a collection of partial matchings that does not conflict with L . Given any collection X of perfect matchings, let $X^j = \{M \cap E(C_j) \mid M \in X\}$. We have

$$\begin{aligned} \Pr\left(A_L \mid \bigwedge_{M \in \mathcal{M}} \overline{A_M}\right) &= \frac{|A_L \cap \bigcap_{M \in \mathcal{M}} \overline{A_M}|}{|\bigcap_{M \in \mathcal{M}} \overline{A_M}|} \\ &= \prod_{j=1}^k \frac{|(A_L \cap \bigcap_{M \in \mathcal{M}} \overline{A_M})^j|}{|(\bigcap_{M \in \mathcal{M}} \overline{A_M})^j|} \\ &= \prod_{j=1}^k \frac{|A_L^j \cap \bigcap_{M \in \mathcal{M}} (\overline{A_M})^j|}{|\bigcap_{M \in \mathcal{M}} (\overline{A_M})^j|}. \end{aligned}$$

Now, since each component C_j is a perfect matching host, we have

$$\frac{|A_L^j \cap \bigcap_{M \in \mathcal{M}} (\overline{A_M})^j|}{|\bigcap_{M \in \mathcal{M}} (\overline{A_M})^j|} \leq \frac{|A_L^j|}{|\Omega^j|}$$

for each j . Thus,

$$\begin{aligned}
\Pr\left(A_L \mid \bigwedge_{M \in \mathcal{M}} \overline{A_M}\right) &= \prod_{j=1}^k \frac{|A_L^j \cap \bigcap_{M \in \mathcal{M}} (\overline{A_M})^j|}{|\bigcap_{M \in \mathcal{M}} (\overline{A_M})^j|} \\
&\leq \prod_{j=1}^k \frac{|A_L^j|}{|\Omega^j|} \\
&= \frac{|A_L|}{|\Omega|} \\
&= \Pr(A_L).
\end{aligned}$$

For the rightward direction, suppose some component C_j fails to be a perfect matching host. That is, there is a matching L of C_j and a collection \mathcal{M} of matchings of C_j that does not conflict with L but

$$\Pr\left(A_L \mid \bigwedge_{M \in \mathcal{M}} \overline{A_M}\right) > \Pr(A_L)$$

in the uniform probability space Ω^j of all perfect matchings of C_j .

Let L' be any perfect matching of H that contains L . Such a perfect matching must exist, since there must be a perfect matching of the component C_i extending L (otherwise, the negative dependency inequality could not have been violated) and each component C_i with $i \neq j$ has some perfect matching (since we are only interested

in hypergraphs with at least one perfect matching). We have

$$\begin{aligned}
\Pr\left(A_{L'} \mid \bigwedge_{M \in \mathcal{M}} \overline{A_M}\right) &= \frac{|A_{L'} \cap \bigcap_{M \in \mathcal{M}} \overline{A_M}|}{|\bigcap_{M \in \mathcal{M}} \overline{A_M}|} \\
&= \frac{|A_L^j \cap \bigcap_{M \in \mathcal{M}} (\overline{A_M})^j|}{|\bigcap_{M \in \mathcal{M}} (\overline{A_M})^j|} \\
&> \frac{|A_L^j|}{|\Omega^j|} \\
&= \frac{|A_{L'}|}{|\Omega^j|} \\
&\geq \frac{|A_{L'}|}{|\Omega|} \\
&= \Pr(A_{L'}),
\end{aligned}$$

so H is not a perfect matching host. □

The following lemma is quite useful in showing a hypergraph is *not* a perfect matching host by directly violating the inequality $\Pr(A_L \mid \overline{A_M}) \leq \Pr(A_L)$. We give a more general statement in terms of events in any probability space, and then specialize it to conflict graphs in the space of perfect matchings.

Lemma 6.2. *Let A and B be nonempty events and let G be a graph having (at least) A and B as vertices. If A and B are not adjacent in G and $A \subseteq \overline{B}$, then G is not a negative dependency graph.*

Proof. Since $\emptyset \neq A \subseteq \overline{B}$, the probability $\Pr(A \mid \overline{B})$ is defined. We have

$$\begin{aligned}
\Pr(A \mid \overline{B}) &= \frac{|A \cap \overline{B}|}{|\overline{B}|} \\
&= \frac{|A|}{|\overline{B}|} && \text{(since } A \subseteq \overline{B}\text{)} \\
&> \frac{|A|}{|\Omega|} \\
&= \Pr(A).
\end{aligned}$$

The inequality uses the assumptions that $|A| \neq 0$ and $|\overline{B}| < |\Omega|$ (since $B \neq \emptyset$). \square

Corollary 6.3. *If there are partial matchings L and M of a hypergraph H such that*

- L and M do not conflict,
- $A_L, A_M \neq \emptyset$, and
- $A_L \subseteq \overline{A_M}$,

then H is not a perfect matching host.

The second condition says there is a perfect matching of H that extends L and a perfect matching of H that extends M , respectively. The third condition says every perfect matching that extends L will conflict with M .

A hypergraph is **k -randomly matchable** provided every partial matching containing at most k edges can be extended to a perfect matching. A **randomly matchable** hypergraph is one that is k -randomly matchable for all k .

Lemma 6.4. *If a perfect matching host is 1-randomly matchable, then it is randomly matchable.*

Proof. Let H be a 1-randomly matchable perfect matching host. We show, by induction on k , that it is k -randomly matchable for all k .

The hypergraph H is 1-randomly matchable by hypothesis. Assume now it is k -randomly matchable for $k \geq 1$. Suppose (for contradiction) there is a matching M containing $k + 1$ edges that does not extend to a perfect matching. Choose any edge e of M and write $M = M' \cup \{e\}$. Observe

- M' and $\{e\}$ do not conflict,
- $A_{M'}, A_{\{e\}} \neq \emptyset$ (by inductive hypothesis and by assumption, respectively), and
- $A_{M'} \subseteq \overline{A_{\{e\}}}$ (since no perfect matching extends M).

Having met the conditions of Corollary 6.3, we must conclude that H is not a perfect matching host, which is contrary to our assumption. Hence, M extends to a perfect matching. As M was arbitrary, we see that any partial matching containing $k + 1$ edges extends to a perfect matching, which completes the induction. \square

The preceding lemma is useful because it allows us to partition the edges of a perfect matching host H into the set A of edges belonging to at least one perfect matching and the set B containing edges belonging to no perfect matching. Suppose we are given a matching L and a collection \mathcal{M} of matchings that does not conflict with L . In order for H to be a perfect matching host, it must satisfy

$$\Pr\left(A_L \mid \bigwedge_{M \in \mathcal{M}} \overline{A_M}\right) \leq \Pr(A_L).$$

If L contains an edge of B , then A_L is empty, and so both sides of the inequality are zero. If any of the $M \in \mathcal{M}$ contains an edge of B , then $\overline{A_M} = \Omega$, and so may be omitted from the intersection. Thus, we may assume that the matchings belonging to \mathcal{M} contain only edges from A . Let $H[A]$ denote the subhypergraph whose edge set is A and whose vertex set is the support of the edges in A . Since H is a perfect matching host, we may conclude that $H[A]$ is a (possibly disconnected) perfect matching host, since these are the only edges that have any bearing on the negative dependence inequality. Since $H[A]$ is 1-randomly matchable by construction, Lemma 6.4 implies $H[A]$ is randomly matchable.

Theorem 6.5. *A hypergraph H is a perfect matching host if and only if $H[A]$ is a (possibly disconnected) randomly matchable perfect matching host, where A contains all edges of H that belong to at least one perfect matching.*

6.2 2-UNIFORM PERFECT MATCHING HOSTS VIA RANDOM MATCHABILITY

Sumner [28] has shown the connected randomly matchable 2-uniform hypergraphs (that is, graphs) are precisely K_{2N} and $K_{N,N}$ for all N . Lu and Székely have shown

K_{2N} [21] and $K_{N,N}$ [20] are, in fact, perfect matching hosts. (Indeed, that the former is a perfect matching host is a special case of Theorem 2.1.) Taken together with Lemma 6.4, we may deduce the following corollary.

Corollary 6.6. *A connected, 1-randomly matchable graph is a perfect matching host if and only if it is K_{2N} or $K_{N,N}$.*

We wish to drop the requirement of 1-random matchability. As before, let G be a graph and write $E(G) = A \cup B$, where each edge of A belongs to at least one perfect matching and no edge of B does. In light of Theorem 6.5 and Corollary 6.6, we know that the induced subgraph $G[A]$ is a disjoint union of **even cliques** (a complete graph on an even number of vertices) and **balanced bicliques** (a complete bipartite graph whose partite sets are of equal size). It remains to characterize how the edges of B can be situated between the components of $G[A]$, which is the main result of this section.

Theorem 6.7. *A graph G is a perfect matching host if and only if there is a partition of the edges into sets A and B such that the induced subgraph $G[A]$ is a disjoint union of even cliques and balanced bicliques and there is no subset F of the edges of B such that*

- *F has an even number of vertices in common with each even clique of $G[A]$ and*
- *for any balanced biclique of $G[A]$, F has an equal number of vertices in common with both of its partite sets.*

In the figures, the induced subgraph $G[A]$ is the disjoint union of a K_6 , K_8 , and $K_{5,5}$ (represented as circles and a rectangle, respectively). The edges of B are shown explicitly. The first figure demonstrates a perfect matching host, since no subset of B meets the conditions stated in the theorem above. The second figure fails to be

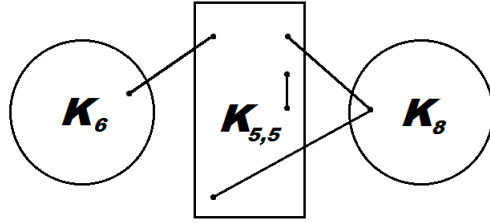


Figure 6.1 Perfect matching host (no F as described in theorem).

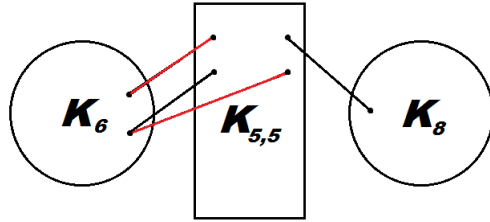


Figure 6.2 Not a perfect matching host, as evidenced by red edges.

a perfect matching host, since the edges highlighted in red meet the K_6 in an even number of vertices and each partite set of the $K_{5,5}$ in one vertex.

6.3 2-UNIFORM PERFECT MATCHING HOSTS VIA COROLLARY 6.3

In the previous section, we characterized the 2-uniform perfect matchings hosts by relying on Sumner's results on randomly matchable graphs. In this section, we derive the conclusion of Corollary 6.6 instead by repeated application of Lemma 6.3.

Lemma 6.8. *If a graph is a connected, 1-randomly matchable perfect matching host of order $2N$, then it contains $K_{N,N}$ as a subgraph.*

Proof. Let G be a connected, 1-randomly matchable perfect matching host of order $2N$ and fix a perfect matching M of G (one must exist, since G is 1-randomly matchable). The idea of the proof will be to look at certain induced subgraphs of G and

conclude which edges must be present in these induced subgraphs based on Lemma 6.3. In the end, we will find these required edges form a $K_{N,N}$ subgraph.

To begin, choose two edges e_1 and e_2 from the perfect matching M . Denote the vertices of each e_i by u_i and v_i . We will choose these edges in such a way that the induced subgraph $G[\{u_1, u_2, v_1, v_2\}]$ is connected (which can always be done, since G is connected). Without loss of generality, assume u_1 is adjacent to u_2 .

Now, for the application of Lemma 6.3, write $R = M \setminus \{e_1, e_2\}$ and $S = \{u_1 u_2\}$.

Observe

- R and S do not conflict,
- $A_R \neq \emptyset$ (since M extends R), and
- $A_S \neq \emptyset$ (since G is 1-randomly matchable).

Now, if $v_1 v_2$ is not an edge of G , then we would have $A_R \subseteq \overline{A_S}$. Taken together with the previous observations, we would conclude that G is not a perfect matching host via Lemma 6.3. In order to avoid contradiction, it must be that the edge $v_1 v_2$ is present in G .

We have shown so far that G contains a $K_{2,2}$ subgraph. We show next how to “grow” this $K_{2,2}$ subgraph into a $K_{3,3}$, from which it will be evident how to proceed from any $K_{s,s}$ to $K_{s+1,s+1}$ until finally all of G contains a $K_{n,n}$ subgraph.

Choose an edge e from the matching M with endpoints u and v such that the induced subgraph $G[\{u, u_1, u_2, v, v_1, v_2\}]$ is connected (which can always be done, since G is connected). Without loss of generality, let u be adjacent to u_1 . For brevity, we show in a table how the vertices u and v are forced to link with $K_{2,2}$ subgraph. As before, we demonstrate two matchings R and S and discover the existence of a new edge in G under threat of Lemma 6.3.

R	S	Edge Gained
$M \setminus \{e, e_1\}$	$\{uu_1\}$	vv_1
$M \setminus \{e, e_1, e_2\} \cup \{uu_1\}$	$\{v_1v_2\}$	vu_2
$M \setminus \{e, e_1, e_2\} \cup \{u_1u_2\}$	$\{vv_1\}$	uv_2

□

Theorem 6.9. *A connected, 1-randomly matchable graph is a perfect matching host if and only if it is K_{2N} or $K_{N,N}$.*

Proof. Let G be a connected, 1-randomly matchable graph of order $2N$ that is also a perfect matching host. We have already shown that $K_{N,N}$ is a subgraph of G , and it is known that $K_{N,N}$ is indeed a perfect matching host [20]. Suppose now $K_{N,N}$ is a proper subgraph of G . Let M be a perfect matching of the $K_{N,N}$ subgraph. Choose two edges e_1 and e_2 from the perfect matching M . Denote the vertices of each e_i by u_i and v_i . We will choose these edges in such a way that the induced subgraph $G[\{u_1, u_2, v_1, v_2\}]$ contains a K_4 minus an edge (which can always be done, since $K_{n,n}$ is a proper subgraph of G). Without loss of generality, assume u_1 is adjacent to u_2 .

Now, for the application of Corollary 6.3, take $R = M \setminus \{e_1, e_2\}$ and take $S = \{u_1u_2\}$. Arguing as before, the corollary implies the edge v_1v_2 is present in G .

We have shown so far that G contains a K_4 subgraph. We show next how to “grow” this K_4 subgraph into a K_6 , from which it will be evident how to proceed from any $K_{2(s-1)}$ to K_{2s} until finally $G = K_{2n}$.

Choose an edge e from the matching M with endpoints u and v such that the induced subgraph $G[\{u, u_1, u_2, v, v_1, v_2\}]$ is connected (which can always be done, since G is connected). Without loss of generality, let u be adjacent to u_1 . For brevity, we show in a table how the vertices u and v are forced to link up with the rest of the vertices, so that $G[\{u, u_1, u_2, v, v_1, v_2\}] = K_6$.

R	S	Edge Gained
$M \setminus \{e, e_1\}$	$\{uu_1\}$	vv_1
$M \setminus \{e, e_1, e_2\} \cup \{uu_1\}$	$\{v_1v_2\}$	u_2v
$M \setminus \{e, e_1, e_2\} \cup \{u_1u_2\}$	$\{vv_1\}$	uv_2
$M \setminus \{e, e_1, e_2\} \cup \{u_1v_2\}$	$\{u_2v\}$	uv_1
$M \setminus \{e, e_1, e_2\} \cup \{u_2v_1\}$	$\{uv_2\}$	u_1v
$M \setminus \{e, e_1, e_2\} \cup \{v_1v_2\}$	$\{u_1v\}$	uu_2
$M \setminus \{e, e_1, e_2\} \cup \{u_1u_2\}$	$\{uv_1\}$	vv_2

□

6.4 FURTHER RESEARCH

This section addressed the characterization of *perfect* matching hosts, for which the ambient probability space contains only perfect matchings. In Theorem 2.1, we showed that a complete uniform hypergraph is a matching host when the ambient probability space contains merely maximal matchings. The first two lemmata hold in this more general space, but the concept of a 1-randomly matchable hypergraph makes no sense in the context of maximal matchings. Indeed, any edge of any hypergraph belongs to a maximal matching trivially. A characterization of matching hosts in the more general space may need a new insight.

Even in the more restrictive setting of perfect matching hosts, we were able to characterize only the graph case aided by Sumner's result on randomly matching graphs. Extending his proof even to 3-uniform hypergraphs has proved challenging. In the figure, we present a seemingly exotic 3-uniform, randomly matchable hypergraph that is a perfect matching host. (In the figure, each solid line indicates a triple of vertices. The colors are merely to aid in recognizing the edges and convey no additional information.) Even if a full characterization of such hypergraphs is unwieldy, it

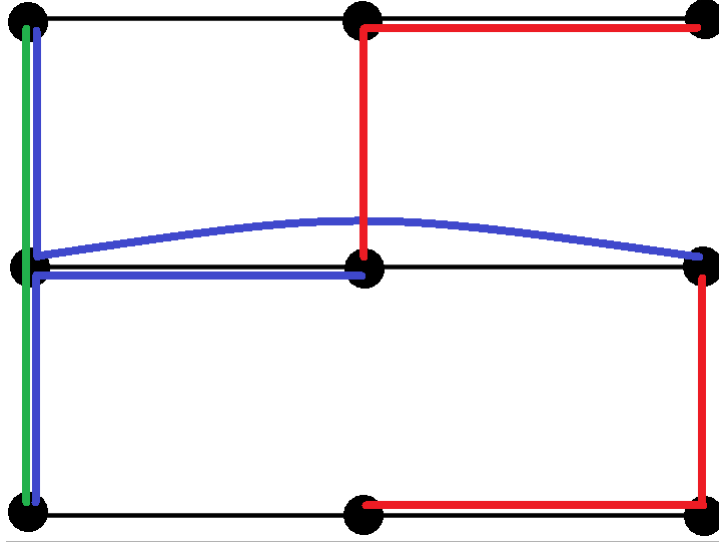


Figure 6.3 3-uniform, randomly matchable perfect matching host.

may be interesting to determine what additional restrictions are necessary to narrow the class to just the complete and complete multipartite hypergraphs.

BIBLIOGRAPHY

- [1] Michael Albert, Alan Frieze, and Bruce Reed, *Multicoloured Hamilton Cycles*, *Electron. J. Combin.* **2** (1995), Research Paper 10, approx. 13 pp. 1327570 (96b:05058)
- [2] Noga Alon and Joel H. Spencer, *The Probabilistic Method*, third ed., Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons Inc., Hoboken, NJ, 2008, With an appendix on the life and work of Paul Erdős. 2437651 (2009j:60004)
- [3] H. D. Becker, *Solution to Problem E 461*, *Amer. Math. Monthly* **48** (1941), 701–702.
- [4] Bell, E. T., *Exponential Numbers*, *Amer. Math. Monthly* **41** (1934), no. 7, 411–419. 1523147
- [5] ———, *Sequence A0001100 in The Online-Encyclopedia of Integer Sequences*, published electronically at <http://oeis.org>, 2013.
- [6] Frank R. Bernhart, *Catalan, Motzkin, and Riordan Numbers*, *Discrete Math.* **204** (1999), no. 1-3, 73–112. 1691863 (2000f:05001)
- [7] Béla Bollobás, *A Probabilistic Proof of an Asymptotic Formula for the Number of Labelled Regular Graphs*, *European J. Combin.* **1** (1980), no. 4, 311–316. 595929 (82i:05045)
- [8] Sadek Bouroubi, *Bell Numbers and Engel’s Conjecture*, *Rostock. Math. Kolloq.* (2007), no. 62, 61–70. 2352650 (2008h:11016)
- [9] E. R. Canfield, *Notes by Ron Canfield on the Moser-Wyman Expansion of the Bell Numbers*, http://www.austinmohr.com/Work_files/bellMoser.pdf.
- [10] ———, *Engel’s Inequality for Bell Numbers*, *J. Combin. Theory Ser. A* **72** (1995), no. 1, 184–187. 1354972 (96m:05012)

- [11] Éva Czabarka, Péter L. Erdős, Virginia Johnson, Anne Kupczok, and László A. Székely, *Asymptotically Normal Distribution of Some Tree Families Relevant for Phylogenetics and of Partitions Without Singletons*, Mosc. J. Comb. Number Theory **1** (2011), no. 3, 12–24. 2988375
- [12] Konrad Engel, *On the Average Rank of an Element in a Filter of the Partition Lattice*, J. Combin. Theory Ser. A **65** (1994), no. 1, 67–78. 1255264 (95c:05013)
- [13] P. Erdős and L. Lovász, *Problems and Results on 3-Chromatic Hypergraphs and Some Related Questions*, Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vol. II, North-Holland, Amsterdam, 1975, pp. 609–627. Colloq. Math. Soc. János Bolyai, Vol. 10. 0382050 (52 #2938)
- [14] Paul Erdős and Joel Spencer, *Lopsided Lovász Local Lemma and Latin Transversals*, Discrete Appl. Math. **30** (1991), no. 2-3, 151–154, ARIDAM III (New Brunswick, NJ, 1988). 1095369 (92c:05160)
- [15] Mehdi Hassani, *Derangements and Applications*, J. Integer Seq. **6** (2003), no. 1, Article 03.1.2, 8 pp. (electronic). 1971432 (2004b:05014)
- [16] Abdolhossein Hoorfar and Mehdi Hassani, *Inequalities on the Lambert W Function and Hyperpower Function*, JIPAM. J. Inequal. Pure Appl. Math. **9** (2008), no. 2, Article 51, 5. 2417333 (2009g:33026)
- [17] C. Y. Ku, *Lovász Local Lemma*, <https://sites.google.com/site/kuchengyeaw/>.
- [18] László Lovász, *Combinatorial Problems and Exercises*, second ed., AMS Chelsea Publishing, Providence, RI, 2007, Problem 4.1. 2321240
- [19] L. Lu, A. Mohr, and L. A. Székely, *Quest for Negative Dependency Graphs*, Recent Adv. in Harm. Anal. and App. (in honor of Konstantin Oskolkov), Springer Proc. in Math. and Stat. **25** (2013), 243–258.
- [20] L. Lu and L. A. Székely, *Using Lovász Local Lemma in the Space of Random Injections*, Electron. J. Combin. **14** (2007), no. 1, Research Paper 63, 13. 2350453 (2008j:60026)
- [21] ———, *A New Asymptotic Enumeration Technique: The Lovász Local Lemma*, arXiv:0905.3983v3 [math.CO], 2011.

- [22] A. Mohr, *Asymptotics Using the Moser-Wyman Expansion for the Bell Numbers*, http://www.austinmohr.com/Work_files/dissertation/avgasymp.mw, 2013.
- [23] Leo Moser and Max Wyman, *An Asymptotic Formula for the Bell Numbers*, Trans. Roy. Soc. Canada. Sect. III. (3) **49** (1955), 49–54. 0078489 (17,1201c)
- [24] Gian-Carlo Rota, *The Number of Partitions of a Set*, Amer. Math. Monthly **71** (1964), 498–504. 0161805 (28 #5009)
- [25] G. Sárközy, *Cycles in Hypergraphs*, <http://web.cs.wpi.edu/~gsarkozy/Cikkek/Rio08.pdf>, 2008.
- [26] Joel Spencer, *Asymptotic Lower Bounds for Ramsey Functions*, Discrete Math. **20** (1977/78), no. 1, 69–76. 0491337 (58 #10600)
- [27] ———, *Ten lectures on the probabilistic method*, second ed., CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 64, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994. 1249485 (95c:05113)
- [28] David P. Sumner, *Randomly Matchable Graphs*, J. Graph Theory **3** (1979), no. 2, 183–186. 530304 (80k:05088)
- [29] Douglas B. West, *Introduction to Graph Theory*, Prentice Hall Inc., Upper Saddle River, NJ, 1996. 1367739 (96i:05001)

APPENDIX A

DETAILS FOR HYPERGRAPH MATCHINGS

A.1 PRELIMINARIES

We review some needed facts from Lu and Székely [21]. Throughout this chapter, expressions of the form $y(\gamma)^k$ should be read as $(y(\gamma))^k$. The notation $\Pr_{N+s}(\cdot)$ means that the event should be considered in the probability space Ω_{N+s} . If no subscript is present, then it is assumed that the event belongs to the probability space Ω_N .

Lemma A.1 (Lu, Székely 2011).

1. For $0 \leq \gamma \leq 1/4$, the equation

$$1 = ye^{-\gamma y}$$

has a unique solution in $1 \leq y \leq 2$ and defines a function $y(\gamma)$.

2. The function $y(\gamma)$ is equal to $-\text{LambertW}_0(-\gamma)/\gamma$, where LambertW_0 is the principal branch of the compositional inverse of xe^x .
3. As the Taylor series of $\text{LambertW}_0(\gamma)$ is convergent for $|\gamma| < 1/e$, so is the Taylor series of $y(\gamma)$.
4. The function $y(\gamma)$ is strictly increasing on $[0, 1/4]$.
5. For $\gamma \rightarrow 0$, we have

$$y(\gamma) = 1 + \gamma + \frac{3}{2}\gamma^2 + \frac{8}{3}\gamma^3 + \frac{125}{24}\gamma^4 + \frac{54}{5}\gamma^5 + O(\gamma^6).$$

6. For $0 \leq \gamma \leq 1/4$, we have

$$1 + \gamma + \frac{3}{2}\gamma^2 \leq y(\gamma) \leq 1 + \gamma + \frac{3}{2}\gamma^2 + 66\gamma^3.$$

Theorem A.2 (Lu, Székely 2011). *Let A_1, \dots, A_n be events with negative dependency graph $([n], E)$. Let us be given any ϵ with $0 < \epsilon < 1/4$. If*

$$\Pr(A_i) < \epsilon \quad \text{and} \quad \sum_{j:i \in E} \Pr(A_j) + 2\Pr^2(A_j) < \epsilon \quad (\text{A.1})$$

for every i , then

$$\Pr\left(\bigwedge_{i \in S} \overline{A_i} \mid \bigwedge_{j \in T} \overline{A_j}\right) \geq \prod_{i \in S} (1 - \Pr(A_i)y(\epsilon))$$

for any disjoint subsets S and T of $[n]$. In particular, we have

$$\Pr\left(\bigwedge_{i=1}^n \overline{A_i}\right) \geq \exp\left(-\sum_{i=1}^n \Pr(A_i)y(\epsilon) - \sum_{i=1}^n \Pr^2(A_i)y(\epsilon)^2\right).$$

Lemma A.3.

1. *An s -matching L belongs to $\overline{A_M}$ if and only if there are edges $e \in L$ and $f \in M$ such that $1 \leq |e \cap f| \leq s - 1$.*
2. *A pair of s -matchings L and M conflict if and only if A_L and A_M are disjoint.*
3. *If the s -matchings L and M do not conflict, then*

$$\overline{A_{M \setminus L}} \subseteq \overline{A_M} \quad \text{and} \quad \overline{A_M} \cap A_L = \overline{A_{M \setminus L}} \cap A_L.$$

Proof. The contrapositive of the first claim says that L belongs to A_M if and only if every pair of edges $e \in L$ and $f \in M$ are either identical or disjoint, which another way of saying that the s -matching L extends M .

The contrapositive of the second claim says that L and M do not conflict if and only if there is an s -matching common to both A_L and A_M . If L and M do not conflict, then their union is again an s -matching, and so there will be a maximal

s -matching extending both. Conversely, if there is an s -matching extending both L and M , it must be that $L \cup M$ is itself a s -matching.

In the third claim, let K be an s -matching that conflicts with $M \setminus L$. By the first claim, there is $e \in K$ and $f \in M \setminus L \subseteq M$ witnessing the conflict, and so K conflicts with M , as well.

In the final claim, we may understand the collections as

$$\overline{A_M} \cap A_L = \{K \in \Omega \mid K \text{ conflicts with } M \text{ but not with } L\}$$

and

$$\overline{A_{M \setminus L}} \cap A_L = \{K \in \Omega \mid K \text{ conflicts with } M \setminus L \text{ but not with } L\}.$$

We have shown the latter is a subset of the former. For the other inclusion, the edge $e \in M$ that witnesses conflict with K does not belong to L , so it must belong to $M \setminus L$. \square

A.2 PROOFS OF THEOREMS 2.2 AND 2.3

Lemma A.4. *Let \mathcal{M} be a collection of s -matchings in K_N^s with negative dependency graph (\mathcal{M}, E) . If there is $\epsilon \in (0, \frac{1}{5})$ such that*

- $\Pr(A_M) \leq \epsilon$,
- $\sum_{L \in \mathcal{M}: LM \in E} \Pr(A_L) + 2 \Pr^2(A_L) < \epsilon$, and
- $\sum_{L \in \mathcal{M}: e \in L} \Pr(A_L) + 2 \Pr^2(A_L) < \epsilon$

for each $M \in \mathcal{M}$ and each $e \in E(K_N^s)$, then

$$\Pr_{N+s} \left(\bigwedge_{M \in \mathcal{M}} \overline{A_M} \right) \leq y(\epsilon)^{2(s-1)} \Pr_N \left(\bigwedge_{M \in \mathcal{M}} \overline{A_M} \right).$$

Proof. Let \mathcal{S} denote the collection of all subsets of $[N + s - 1]$ of size $s - 1$. For each $S \in \mathcal{S}$, define

$$\mathcal{B}_S = \{M \in \mathcal{M} \mid M \text{ does not conflict with the edge } S \cup \{N + s\}\}.$$

Lu and Székely [21] have shown

$$\Pr_{N+s} \left(\bigwedge_{M \in \mathcal{M}} \overline{A_M} \right) = \frac{1}{\binom{N+s-1}{s-1}} \sum_{S \in \mathcal{S}} \Pr_N \left(\bigwedge_{M \in \mathcal{B}_S} \overline{A_M} \right). \quad (\text{A.2})$$

We will apply the first part of Theorem A.2 with the collections $\mathcal{M} \setminus \mathcal{B}_S$ and \mathcal{B}_S . The collection \mathcal{B}_S contains those matchings from \mathcal{M} whose support is disjoint from S , while $\mathcal{M} \setminus \mathcal{B}_S$ contains those matchings whose support meets S . Our goal is to show

$$\frac{\Pr \left(\bigwedge_{M \in \mathcal{M}} \overline{A_M} \right)}{\Pr \left(\bigwedge_{M \in \mathcal{B}_S} \overline{A_M} \right)} = \Pr \left(\bigwedge_{M \in \mathcal{M} \setminus \mathcal{B}_S} \overline{A_M} \mid \bigwedge_{M \in \mathcal{B}_S} \overline{A_M} \right) \geq y(\epsilon)^{-2(s-1)}. \quad (\text{A.3})$$

Theorem A.2 gives

$$\Pr \left(\bigwedge_{M \in \mathcal{M} \setminus \mathcal{B}_S} \overline{A_M} \mid \bigwedge_{M \in \mathcal{B}_S} \overline{A_M} \right) \geq \prod_{M \in \mathcal{M} \setminus \mathcal{B}_S} (1 - \Pr(A_M)y(\epsilon)).$$

If the product on the righthand side is empty, then we have nothing to prove, so we assume otherwise.

Without loss of generality, let $S = [s-1]$. For each vertex i of S , pick an edge e_i belonging to some matching $M_i \in \mathcal{M} \setminus \mathcal{B}_S$ such that i is a vertex of e_i . (If there is a vertex of S with no such edge, then we simply disregard that vertex.) By definition, every matching belonging to $\mathcal{M} \setminus \mathcal{B}_S$ meets S . Hence, every matching belonging to $\mathcal{M} \setminus \mathcal{B}_S$ either contains one of the e_i or conflicts with one. Letting $\mathcal{M}_i = \{M \in \mathcal{M} \setminus \mathcal{B}_S \mid e_i \in M\}$, we have

$$\mathcal{M} \setminus \mathcal{B}_S \subseteq \bigcup_{i=1}^{s-1} \mathcal{M}_i \cup \bigcup_{i=1}^{s-1} N(M_i),$$

where $N(M_i)$ contains the neighbors of M_i in the negative dependency graph (that is, those matchings of \mathcal{M} that conflict with M_i).

Now,

$$\begin{aligned} & \prod_{M \in \mathcal{M} \setminus \mathcal{B}_S} (1 - \Pr(A_M)y(\epsilon)) \\ & \geq \prod_{i=1}^{s-1} \prod_{M \in \mathcal{M}_i} (1 - \Pr(A_M)y(\epsilon)) \prod_{M \in N(M_i)} (1 - \Pr(A_M)y(\epsilon)). \end{aligned}$$

Observe $1 - x \geq e^{-x-x^2}$ for $x \in (0, 1/2)$. Since $\epsilon < \frac{1}{5}$, we have

$$P(A_M)y(\epsilon) < \epsilon \left(1 + \epsilon + \frac{3}{2}\epsilon^2 + 66\epsilon^3 \right) < \frac{1}{2},$$

where $y(\epsilon)$ is bounded by Theorem A.1. We may therefore write

$$\begin{aligned} & \prod_{M \in \mathcal{M}_i} (1 - \Pr(A_M)y(\epsilon)) \\ & \geq \prod_{M \in \mathcal{M}_i} \exp \left(-\Pr(A_M)y(\epsilon) - \Pr^2(A_M)y(\epsilon)^2 \right) \\ & = \exp \left(-\sum_{M \in \mathcal{M}_i} \Pr(A_M)y(\epsilon) - \sum_{M \in \mathcal{M}_i} \Pr^2(A_M)y(\epsilon)^2 \right) \\ & = \exp \left(-y(\epsilon) \left(\sum_{M \in \mathcal{M}_i} \Pr(A_M) + \sum_{M \in \mathcal{M}_i} \Pr^2 y(\epsilon) \right) \right) \\ & \geq \exp \left(-y(\epsilon) \left(\sum_{M \in \mathcal{M}_i} \Pr(A_M) + \sum_{M \in \mathcal{M}_i} 2\Pr^2(A_M) \right) \right) \\ & \geq \exp \left(-y(\epsilon) \left(\sum_{M \in \mathcal{M}: e_i \in M} \Pr(A_M) + \sum_{M \in \mathcal{M}: e_i \in M} 2\Pr^2(A_M) \right) \right) \\ & > \exp(-y(\epsilon)\epsilon) \\ & = y(\epsilon)^{-1} \end{aligned}$$

for each i .

We may similarly derive

$$\prod_{M \in \mathcal{N}(M_i)} (1 - \Pr(A_M)y(\epsilon)) \geq y(\epsilon)^{-1}$$

for each i .

Multiplying the bounds together gives

$$\begin{aligned} & \prod_{M \in \mathcal{M} \setminus \mathcal{B}_S} (1 - \Pr(A_M)y(\epsilon)) \\ & \geq \prod_{i=1}^{s-1} \prod_{M \in \mathcal{M}_i} (1 - \Pr(A_M)y(\epsilon)) \prod_{M \in \mathcal{N}(M_i)} (1 - \Pr(A_M)y(\epsilon)) \\ & \geq \prod_{i=1}^{s-1} y(\epsilon)^{-2} \\ & \geq \prod_{i=1}^{s-1} y(\epsilon)^{-2(s-1)}, \end{aligned}$$

which finally establishes Condition (A.3).

Combining Conditions (A.2) and (A.3), we have

$$\begin{aligned} \Pr_{N+s} \left(\bigwedge_{M \in \mathcal{M}} \overline{A_M} \right) &= \frac{1}{\binom{N+s-1}{s-1}} \sum_{S \in \mathcal{S}} \Pr_N \left(\bigwedge_{M \in \mathcal{B}_S} \overline{A_M} \right) \\ &\leq \frac{1}{\binom{N+s-1}{s-1}} \sum_{S \in \mathcal{S}} \Pr_N \left(\bigwedge_{M \in \mathcal{M}} \overline{A_M} \right) y(\epsilon)^{2(s-1)} \\ &= y(\epsilon)^{2(s-1)} \Pr_N \left(\bigwedge_{M \in \mathcal{M}} \overline{A_M} \right). \end{aligned}$$

□

We now restate Theorem 2.2 in full detail.

Theorem 2.2. *Let \mathcal{M} be a collection of matchings in a complete s -uniform hypergraph. If \mathcal{M} is δ -sparse and k -bounded, then the conflict graph for the canonical events $\{A_M \mid M \in \mathcal{M}\}$ is also an ϵ -near positive dependency graph with*

$$\epsilon = 1 - y(2\delta)^{-2k(s-1)} \exp \left(-\delta y(2\delta) - \delta^2 y(2\delta)^2 \right) \quad (\text{A.4})$$

and therefore

$$\Pr \left(\bigwedge_{M \in \mathcal{M}} \overline{A_M} \right) \leq \prod_{M \in \mathcal{M}} \left(1 - \Pr(A_M) y(2\delta)^{-2k(s-1)} \exp \left(-\delta y(2\delta) - \delta^2 y(2\delta)^2 \right) \right). \quad (\text{A.5})$$

Proof. We show first that the conflict graph G is an ϵ -near positive dependency graph for the prescribed ϵ . Theorem 1.7 together with (A.4) will finish the proof of (A.5).

For the first part of the definition, L is adjacent to M in G if and only if L and M conflict. By Lemma A.3, we have $\Pr(A_L \wedge A_M) = 0$.

Given any $F \in \mathcal{M}$ and a subset S of $\overline{N(F)}$, we need to prove

$$\Pr \left(A_F \mid \bigwedge_{M \in S} \overline{A_M} \right) \geq (1 - \epsilon) \Pr(A_F),$$

which is equivalent to

$$\Pr \left(\bigwedge_{M \in S} \overline{A_M} \mid A_F \right) \geq (1 - \epsilon) \Pr \left(\bigwedge_{M \in S} \overline{A_M} \right).$$

Let $S_F = \{M \setminus F \mid M \in S\}$. Observe that $\emptyset \notin S_F$ since \mathcal{M} is δ -sparse. Note that

$$\Pr\left(\bigwedge_{M \in S} \overline{A_M} \mid A_F\right) = \frac{\Pr\left(\bigwedge_{M \in S} \overline{A_M} \wedge A_F\right)}{\Pr(A_F)} \quad (\text{A.6})$$

$$\begin{aligned} &= \frac{\Pr\left(A_F \wedge \bigwedge_{M \in S} \overline{A_{M \setminus F}}\right)}{\Pr(A_F)} \\ &= \Pr\left(\bigwedge_{M \in S_F} \overline{A_M} \mid A_F\right). \end{aligned} \quad (\text{A.7})$$

Now,

$$\Pr_N\left(\bigwedge_{M \in S_F} \overline{A_M} \mid A_F\right) = \Pr_{N-s|F|}\left(\bigwedge_{M \in S_F} \overline{A_M}\right) \quad (\text{A.8})$$

$$\begin{aligned} &= \Pr_N\left(\bigwedge_{M \in S_F} \overline{A_M}\right) \prod_{j=1}^{|F|} \frac{\Pr_{N-sj}\left(\bigwedge_{M \in S_F} \overline{A_M}\right)}{\Pr_{N-s(j-1)}\left(\bigwedge_{M \in S_F} \overline{A_M}\right)} \\ (\text{by Lemma A.4}) &\geq \Pr_N\left(\bigwedge_{M \in S_F} \overline{A_M}\right) \prod_{j=1}^{|F|} y(2\delta)^{-2(s-1)} \\ &\geq \Pr_N\left(\bigwedge_{M \in S_F} \overline{A_M}\right) y(2\delta)^{-2(s-1)k}. \end{aligned} \quad (\text{A.9})$$

For any M that does not conflict with F , Lemma A.3 gives $\overline{A_{M \setminus F}} \subseteq \overline{A_M}$. Letting $S_F = \{M \setminus F \mid M \in S\}$, we have

$$\begin{aligned} \frac{\Pr\left(\bigwedge_{M \in S_F} \overline{A_M}\right)}{\Pr\left(\bigwedge_{M \in S} \overline{A_M}\right)} &= \frac{\Pr\left(\bigwedge_{M \in S} \overline{A_{M \setminus F}}\right)}{\Pr\left(\bigwedge_{M \in S} \overline{A_M}\right)} \\ &= \frac{\Pr\left(\bigwedge_{M \in S} \overline{A_{M \setminus F}} \wedge \overline{A_M}\right)}{\Pr\left(\bigwedge_{M \in S} \overline{A_M}\right)} \end{aligned} \quad (\text{A.10})$$

$$\begin{aligned} &= \frac{\Pr\left(\left[\bigwedge_{M \in S, M \cap F \neq \emptyset} \overline{A_{M \setminus F}}\right] \wedge \left[\bigwedge_{M \in S} \overline{A_M}\right]\right)}{\Pr\left(\bigwedge_{M \in S} \overline{A_M}\right)} \\ &= \Pr\left(\bigwedge_{M \in S_F \setminus S} \overline{A_M} \mid \bigwedge_{M \in S} \overline{A_M}\right). \end{aligned} \quad (\text{A.11})$$

Now apply the first part of Theorem A.2 with $S_F \setminus S$ and S to obtain

$$\begin{aligned}
& \Pr \left(\bigwedge_{M \in S_F \setminus S} \overline{A_M} \mid \bigwedge_{M \in S} \overline{A_M} \right) \\
& \geq \prod_{M \in S_F \setminus S} (1 - \Pr(A_M) y(2\delta)) \\
& \geq \prod_{M \in S_F} (1 - \Pr(A_M) y(2\delta)) \\
& \geq \exp \left(- \sum_{M \in S_F} \Pr(A_M) y(2\delta) - \sum_{M \in S_F} \Pr^2(A_M) y(2\delta)^2 \right) \\
& \geq \exp \left(-\delta y(2\delta) - \delta^2 y(2\delta)^2 \right). \tag{A.12}
\end{aligned}$$

Finally, we have

$$\begin{aligned}
& \Pr \left(\bigwedge_{M \in S} \overline{A_M} \mid A_F \right) \\
\text{by (A.6-A.7)} &= \Pr \left(\bigwedge_{M \in S_F} \overline{A_M} \mid A_F \right) \\
\text{by (A.8-A.9)} &\geq \Pr \left(\bigwedge_{M \in S_F} \overline{A_M} \right) y(2\delta)^{-2(s-1)k} \\
\text{by (A.10-A.11)} &= \Pr \left(\bigwedge_{M \in S} \overline{A_M} \right) \Pr \left(\bigwedge_{M \in S_F \setminus S} \overline{A_M} \mid \bigwedge_{M \in S} \overline{A_M} \right) y(2\delta)^{-2(s-1)k} \\
\text{by (A.12)} &\geq \Pr \left(\bigwedge_{M \in S} \overline{A_M} \right) e^{-\delta y(2\delta) - \delta^2 y(2\delta)^2} y(2\delta)^{-2(s-1)k}.
\end{aligned}$$

Thus, the negative dependency graph G is also a ϵ -near positive dependency graph.

The proof is finished by Theorem 1.7. \square

The expression $\Pr \left(\bigwedge_{M \in \mathcal{M}} \overline{A_M} \right)$ can be bounded from below by Theorem A.2 and bounded from above by Theorem 2.2, which can be combined to obtain asymptotics under the appropriate conditions.

Theorem 2.3. *Let Ω_N denote the uniform probability space of perfect matchings of K_N^s , the complete s -uniform hypergraph on N vertices. Let r and ϵ both depend on N , where r is a positive integer and ϵ is a real number eventually lying in the interval*

$(0, \frac{1}{16})$. Let \mathcal{M} be a k -bounded collection of matchings in K_N^s in which no matching is a subset of another. For any matching $M \in \mathcal{M}$, define the canonical event

$$A_M = \{L \in \Omega_N \mid M \subseteq L\}.$$

Set $\mu = \sum_{M \in \mathcal{M}} \Pr(A_M)$. Finally, suppose the following inequalities are satisfied for every matching $M \in \mathcal{M}$ and every edge e of K_N^s :

- $\Pr(A_M) < \epsilon$
- $\sum_{L: L, M \text{ conflict}} \Pr(A_L) < \epsilon$
- $\sum_{L \in \mathcal{M}: e \in L} \Pr(A_L) < \epsilon$
- $\sum_{L \in \mathcal{M}_M} \Pr_{N-sk}(A_L) < \epsilon$

If, in addition, $ks\epsilon = o(1)$, then

$$\Pr\left(\bigwedge_{M \in \mathcal{M}} \overline{A_M}\right) = e^{-\mu + O(kse\mu)}.$$

Furthermore, if $ks\epsilon\mu = o(1)$, then

$$\Pr\left(\bigwedge_{M \in \mathcal{M}} \overline{A^M}\right) = (1 + O(kse\mu)) e^{-\mu}.$$

Proof. Let G be the conflict graph for the collection $\{A_M \mid M \in \mathcal{M}\}$ of canonical events. By Theorem 1.1, the graph G is a negative dependency graph. Note that the condition (A.1) in Theorem A.2 is satisfied with 2ϵ instead of ϵ , where ϵ is from the conditions of Theorem 2.3. Applying Theorem A.2, we have

$$\begin{aligned} \Pr\left(\bigwedge_{M \in \mathcal{M}} \overline{A_M}\right) &\geq \exp\left(-\sum_{M \in \mathcal{M}} \Pr(A_M)y(2\epsilon) - \sum_{M \in \mathcal{M}} \Pr^2(A_M)y(2\epsilon)^2\right) \\ &> \exp\left(-\sum_{M \in \mathcal{M}} \Pr(A_M)y(2\epsilon) - \sum_{M \in \mathcal{M}} \Pr(A_M)\epsilon y(2\epsilon)^2\right) \\ &= \exp\left(-\mu\left(1 + 3\epsilon + O(\epsilon^2)\right)\right). \end{aligned}$$

Now we consider the upper bound. Note that \mathcal{M} is 2ϵ -sparse and k -bounded. By Theorem 2.2, we have

$$\begin{aligned}
& \Pr\left(\bigwedge_{M \in \mathcal{M}} \overline{A_M}\right) \\
& \leq \prod_{M \in \mathcal{M}} \left(1 - \Pr(A_M) \exp\left(-2\epsilon y(4\epsilon) - (2\epsilon)^2 y(4\epsilon)^2\right) y(4\epsilon)^{-2(s-1)k}\right) \\
& \leq \exp\left(-\sum_{M \in \mathcal{M}} \Pr(A_M) \exp\left(-2\epsilon y(4\epsilon) - (2\epsilon)^2 y(4\epsilon)^2\right) y(4\epsilon)^{-2(s-1)k}\right) \\
& = \exp\left(-\mu \exp\left(-2\epsilon - O(\epsilon^2)\right) y(4\epsilon)^{-2(s-1)k}\right),
\end{aligned}$$

where we use $y(4\epsilon) = 1 + 4\epsilon + O(\epsilon^2)$.

Focusing now on the factor of $y(4\epsilon)^{-2(s-1)k}$, we have

$$\begin{aligned}
y(4\epsilon)^{-2(s-1)k} &= \exp\left(-2(s-1)k \log(y(4\epsilon))\right) \\
&= \exp\left(-2(s-1)k \log(1 + 4\epsilon + O(\epsilon^2))\right) \\
&= \exp\left(-2(s-1)k(4\epsilon + O(\epsilon^2))\right) \\
&= \exp\left(-8\epsilon(s-1)k - O(\epsilon^2 sk)\right).
\end{aligned}$$

Returning now to the main term, we have

$$\begin{aligned}
\Pr\left(\bigwedge_{M \in \mathcal{M}} \overline{A_M}\right) &= \exp\left(-\mu \exp\left(-2\epsilon - O(\epsilon^2)\right) y(4\epsilon)^{-2(s-1)k}\right) \\
&= \exp\left(-\mu \exp\left(-2\epsilon - O(\epsilon^2) - 8\epsilon(s-1)k - O(\epsilon^2 sk)\right)\right) \\
&= \exp\left(-\mu(1 - (8(s-1)k + 2)\epsilon + O(\epsilon^2 s^2 k^2))\right).
\end{aligned}$$

Combining the lower bound and the upper bound above, we obtain equation (2.1). \square

APPENDIX B

USEFUL FACTS ABOUT BELL NUMBERS

Throughout, let B_N denote the number of partitions of the set $[N]$ (the N^{th} Bell number), B_N^* denote the number of singleton-free partitions on $[N]$, and $S(N, k)$ denote the number of partitions of $[N]$ into exactly k nonempty subsets (the Stirling numbers of the second kind).

Lemma B.1. *The inequality*

$$2B_N < B_{N+1} < (N + 1)B_N.$$

holds for all $N \geq 2$.

Proof. Bouroubi [8] gave a proof via the generating function

$$B_N(x) = \sum_{k=0}^N S(N, k)x^k.$$

We give an argument from first principles instead.

Given any partition P of $[N]$, one can form at least two distinct partitions of $[N + 1]$ by introducing the element $N + 1$ either as a singleton or by inserting it into an existing block of P . Since P has at least one but at most N blocks, we can create at least two but at most $N + 1$ distinct partitions in this way. The result follows by applying this operation to all B_N partitions of $[N]$ and noticing that distinct partitions P and Q cannot be mapped to the same partition of $[N + 1]$. \square

Lemma B.2. *The identity*

$$B_N = B_{N+1}^* + B_N^*$$

holds for all N .

Proof. This identity was first observed by Becker [3] in the context of a certain difference equation. Another interesting treatment of the result comes from Bernhart [6] in the context of noncrossing partitions. We prove the equivalent statement $B_{N+1}^* = B_N - B_N^*$. Notice the lefthand side counts the partitions of the set $[N + 1]$ that do *not* have singletons, and the righthand side counts the partitions of the set $[N]$ that *do* have singletons. Given a partition of the former type, create a partition of the latter type by splitting the block containing the element $N + 1$ into singletons and removing the element $N + 1$. This operation defines a bijection between the two collections, thus establishing the desired identity. \square

Lemma B.3. *The average number of blocks in a partition of $[N]$ is $\frac{B_{N+1}}{B_N} - 1$.*

Proof. Canfield [10] and Engel [12] prove this fact using identities relating the Bell number to the Stirling numbers of the second kind. For example, it is well known that $B_N = \sum_{k=1}^N S(N, k)$ and $S(N + 1, k) = kS(N, k) + S(N, k - 1)$. By convention, $S(N, k) = 0$ whenever $k \notin \{1, \dots, N\}$. Let $a_N(\Omega)$ denote the average number of blocks in a partition of $[N]$. Making use of the aforementioned identities, we find

$$\begin{aligned} a_N(\Omega) &= \sum_{k=1}^{\infty} \frac{kS(N, k)}{B_N} \\ &= \sum_{k=1}^{\infty} \frac{S(N + 1, k) - S(N, k - 1)}{B_N} \\ &= \frac{B_{N+1} - B_N}{B_N} \\ &= \frac{B_{N+1}}{B_N} - 1. \end{aligned}$$

A proof avoiding (direct) reference to Stirling numbers is provided by observing that every partition of $[N + 1]$ is formed in exactly one way from an appropriately chosen partition of $[N]$ with the element $N + 1$ as a singleton or inserted into an

existing block. Thus we have

$$\begin{aligned} B_{N+1} &= \sum_{P \in \Omega_N} (|P| + 1) \\ B_{N+1} - B_N &= \sum_{P \in \Omega_N} |P| \\ \frac{B_{N+1} - B_N}{B_N} &= \sum_{P \in \Omega_N} \frac{|P|}{B_N} \\ \frac{B_{N+1}}{B_N} - 1 &= a_N(\Omega). \end{aligned}$$

□

APPENDIX C

PROOF OF THEOREM 5.1

Lemma C.1. *The number of perfect s -matchings of the complete s -uniform hypergraph on sN vertices is*

$$\frac{(sN)!}{(s!)^N N!}.$$

Proof. For fixed s , let $f_s(N)$ be the number of perfect s -matchings of the complete s -uniform hypergraph on sN vertices. Fixing some vertex v , there are $\binom{sN-1}{s-1}$ ways to form the edge containing v . Since $sN - s = s(N - 1)$ vertices remain to be matched, we have the recurrence

$$\begin{aligned} f_s(N) &= \binom{sN-1}{s-1} f_s(N-1), \\ f_s(0) &= 1. \end{aligned}$$

Iteration gives

$$f_s(N) = \frac{(sN)!}{(s!)^N N!}.$$

□

Theorem 5.1. *In the configuration model, assume $g \geq 1$, $r \geq 3$, and*

$$(2r - 2)^{2g-3} g^3 = o(N). \tag{C.1}$$

The probability that an r -regular, 3-uniform multihypergraph chosen uniformly at random has girth at least g is

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

If $g \geq 3$, then the number of simple r -regular, 3-uniform hypergraphs on N vertices with girth at least g is

$$(1 + o(1)) \exp \left(- \sum_{i=1}^{g-1} \frac{(2r-2)^i}{2i} \right) \frac{(rN)!}{6^{rN/3} \left(\frac{rN}{3}\right)! (r!)^N}.$$

In fact, letting \mathcal{C} be a subset of $\{3, 4, \dots, g-1\}$, the number of simple r -regular, 3-uniform hypergraphs whose cycle lengths do not belong to \mathcal{C} is

$$(1 + o(1)) \exp \left(1 - r - (r-1)^2 - \sum_{i \in \mathcal{C}} \frac{(2r-2)^i}{2i} \right) \frac{(rN)!}{6^{rN/3} \left(\frac{rN}{3}\right)! (r!)^N}.$$

Proof. We prove the first claim. To prove the second claim, only (C.3) has to be adjusted.

Recall the following definition of j -cycle in a hypergraph:

- A 1-cycle is a single edge with a repeated vertex.
- A 2-cycle is a pair of edges whose intersection contains at least two vertices.
- For $j \geq 3$, a j -cycle is a collection e_1, \dots, e_j of edges for which there are distinct vertices v_1, \dots, v_j such that $e_i \cap e_{i+1} = \{v_i\}$ for all i (where e_{j+1} is understood to be e_1).

For $i = 1, \dots, g-1$, let \mathcal{M}_i be the set of (partial) matchings of $U = [rN]$ whose projection gives *precisely* a cycle of length i .

Matchings in \mathcal{M}_1 project to single edges having either exactly two or exactly three repeated vertices. There are $(N)_2 \binom{r}{2} r$ matchings of the former type and $N \binom{r}{3}$ of the latter type.

Matchings in \mathcal{M}_2 project to pairs of edges having either exactly two or exactly three vertices in common. There are $\frac{1}{2} \binom{N}{3} r^3 (r-1)^3$ two-edge matchings of the latter type. The former is the case $i = 2$ in the next paragraph.

The cases so far enumerated are degenerate. Recall that a simple 3-uniform hypergraph is one in which every edge contains three distinct vertices and any pair of

edges intersect in at most two vertices. In a simple 3-uniform hypergraph, a typical i -cycle ($i \geq 2$) contains $2i$ distinct vertices. Of these vertices, i of them belong to exactly one edge, and the other i belong to exactly two edges. The vertices can be selected and placed on a cycle in $\frac{1}{2i}(N)_{2i}$ ways. Each vertex belonging to exactly one edge can arise from r different minivertices. Each vertex belonging to exactly two edges can arise from $r(r-1)$ ordered pairs of minivertices. Putting this together, the number of matchings whose projection gives precisely a nondegenerate i -cycle for $i \geq 2$ is given by $\frac{1}{2i}(N)_{2i}r^{2i}(r-1)^i$.

Summarizing, we have

- $|\mathcal{M}_1| = (N)_2 \binom{r}{2} r + N \binom{r}{3} = \frac{1}{2}(N)_2 r^2 (r-1) \left(1 + \frac{r-2}{3(N-1)r}\right)$,
- $|\mathcal{M}_2| = \frac{1}{4}(N)_4 r^4 (r-1)^2 + \frac{1}{2} \binom{N}{3} r^3 (r-1)^3 = \frac{1}{4}(N)_4 r^4 (r-1)^2 \left(1 + \frac{r-1}{3(N-3)r}\right)$, and
- $|\mathcal{M}_i| = \frac{1}{2i}(N)_{2i} r^{2i} (r-1)^i$

for $i \geq 3$.

The bad events for the negative dependency graph are the union of matchings $\mathcal{M} = \bigcup_{i=1}^{g-1} \mathcal{M}_i$.

Recall rN is divisible by 3. For positive integers j , define

$$(rN)_{3j,3} = \frac{(rN)_{3j}}{\prod_{i=0}^{j-1} (rN - 3i)}.$$

Observe $(rN)_{3j,3}$ is a product of $2j$ integers.

For fixed $1 \leq i \leq g-1$ and each $M \in \mathcal{M}_i$, we have

$$|A_M| = \frac{(rN - 3i)!}{6^{\frac{rN-3i}{3}} \left(\frac{rN-3i}{3}\right)!}.$$

Taken together with the fact that the total number of perfect 3-matchings on rN vertices is

$$\frac{(rN)!}{6^{\frac{rN}{3}} \left(\frac{rN}{3}\right)!},$$

we obtain

$$\Pr(A_M) = \frac{2^i}{(rN)_{3i,3}}. \quad (\text{C.2})$$

Now,

$$\begin{aligned} \sum_{M \in \mathcal{M}} \Pr(A_M) &= \sum_{i=1}^{g-1} \sum_{M \in \mathcal{M}_i} \Pr(A_M) \\ &= \frac{1}{2} (N)_{2r^2} (r-1) \left(1 + \frac{r-2}{3(N-1)r}\right) \cdot \frac{2}{(rN)_{3,3}} \\ &\quad + \frac{1}{4} (N)_{4r^4} (r-1)^2 \left(1 + \frac{r-1}{3(N-3)r}\right) \cdot \frac{4}{(rN)_{6,3}} \\ &\quad + \sum_{i=3}^{g-1} \frac{1}{2i} (N)_{2i} r^{2i} (r-1)^i \cdot \frac{2^i}{(rN)_{3i,3}} \\ &= \sum_{i=1}^{g-1} \frac{(2r-2)^i}{2i} \left(1 + O\left(\frac{i^2}{N}\right)\right) \\ &= \left(1 + O\left(\frac{g^2}{N}\right)\right) \sum_{i=1}^{g-1} \frac{(2r-2)^i}{2i}. \end{aligned} \quad (\text{C.3})$$

The collection \mathcal{M} is $(g-1)$ -bounded by construction. Let also $\epsilon = \frac{K(2r-2)^{g-2}g^3}{N}$ for a large constant K . We verify the conditions of Theorem 2.3.

For item 1,

$$\begin{aligned} \Pr(A_M) &= \frac{2^i}{(rN)_{3i,3}} \\ &\leq \frac{2^{g-1}}{(rN-1)(rN-2)} \\ &< \epsilon. \end{aligned}$$

We verify item 2 in cases.

We first bound the size of the set $\{L \in \mathcal{M}_1 \mid L \text{ and } M \text{ conflict}\}$ for fixed $M \in \mathcal{M}$. Recall matchings in \mathcal{M}_1 project to single edges having at most two distinct vertices. To construct a matching $L \in \mathcal{M}_1$ that conflicts with M , choose the first minivertex of L from among the minivertices of M to ensure conflict, which can be done in at most $3(g-1)$ ways. After this, there are fewer than rN choices for the second minivertex of L and $2(r-1)$ choices for the third (since the minivertices belong to at most two

different vertices). Thus, we have

$$|\{L \in \mathcal{M}_1 \mid L \text{ and } M \text{ conflict}\}| \leq 6(g-1)rN(r-1).$$

We now bound the size of the set $\{L \in \mathcal{M}_i \mid L \text{ and } M \text{ conflict}\}$ for each fixed $i \geq 2$. First, select a minivertex u from among the minivertices of M to ensure conflict, which can be done in at most $3(g-1)$ ways. Select two other minivertices v and w to join with u in at most $rN(N-1)r$ ways. Finally, decide whether v will be a minivertex belonging to exactly one or exactly two edges of L . (In the former case, we will call v the “first” minivertex. In the latter case, we will call u the “first” minivertex. In either case, w is the “last” minivertex.) Multiplying everything together, the first triple of L can be formed in at most $6(g-1)rN(N-1)r$ ways. The second triple carries with it the restriction that its first minivertex must belong to the same vertex as the last minivertex of the previous triple. The second triple can therefore be formed in at most $(r-1)(N-2)r(N-3)d$ ways. Carry on in this way until the i th triple is to be formed, which can be accomplished in at most $(r-1)(N-2i+2)r(r-1)$ ways, since it carries the further restriction that its final minivertex must belong to the same vertex as the first minivertex in the first triple. Multiplying all these choices together gives

$$|\{L \in \mathcal{M}_i \mid L \text{ and } M \text{ conflict}\}| \leq 6(g-1)(N)_{2i-1}r^{2i-1}(r-1)^i$$

for each $i \geq 2$.

Now,

$$\begin{aligned}
\sum_{L: L, M \text{ conflict}} \Pr(A_L) &= \sum_{i=1}^{g-1} \sum_{L \in \mathcal{M}_i: L, M \text{ conflict}} \Pr(A_L) \\
&\leq \sum_{i=1}^{g-1} 6(g-1)(N)_{2i-1} r^{2i-1} (r-1)^i \cdot \frac{2^i}{(rN)_{3i,3}} \\
&\leq \frac{12(g-1)}{rN-1} \sum_{i=1}^{g-1} (2r-2)^i \left(1 + O\left(\frac{i^2}{N}\right)\right) \\
&\leq \frac{12g^2(2r-2)^{g-1}}{rN-1} \left(1 + O\left(\frac{g^2}{N}\right)\right) \\
&< \epsilon
\end{aligned}$$

for large N .

Now we verify item 3. Fix any edge e of K_N^s . We wish to bound the size of the set

$$\{L \in \mathcal{M} \mid e \in L\},$$

which we write as the disjoint union

$$\bigcup_{i=1}^{g-1} \{L \in \mathcal{M}_i \mid e \in L\}.$$

There is at most one matching in the collection \mathcal{M}_1 that contains e , since all such matchings contain exactly one edge.

For degenerate 2-cycles, we must form a second edge using the three vertices appearing in the edge e . The minivertices can thus be chosen in at most $(r-1)^3$ ways. For proper 2-cycles, we can count as in the verification of item 2 to obtain at most $3(N-3)r(r-1)^2$ cycles. (Recall, in a non-degenerate cycle, each edge has two vertices of degree two and one vertex of degree one. The factor of 3 above comes from the freedom to choose any of the three minivertices of e to project to the degree one vertex.) In total, we have

$$|\{L \in \mathcal{M}_2 \mid e \in L\}| \leq 3Nr(r-1)^2.$$

For $i \geq 3$, all i -cycles are proper, so we proceed as before to obtain

$$|\{L \in \mathcal{M}_i \mid e \in L\}| \leq 3(N)_{2i-3} r^{2i-3} (r-1)^i.$$

Now,

$$\begin{aligned} & \sum_{L \in \mathcal{M}: e \in L} \Pr(A_L) \\ = & \sum_{i=1}^{g-1} \sum_{L \in \mathcal{M}_i: e \in L} \Pr(A_L) \\ \leq & \frac{2}{(rN)_{3,3}} + \sum_{i=2}^{g-1} 3(N)_{2i-3} r^{2i-3} (r-1)^i \cdot \frac{2^i}{(rN)_{3i,3}} \\ \leq & \frac{2}{(rN-1)(rN-2)} + \frac{6}{(rN-1)(rN-2)(rN-4)} \sum_{i=2}^{g-1} (2r-2)^i \left(1 + O\left(\frac{i^2}{N}\right)\right) \\ \leq & \frac{2}{(rN-1)(rN-2)} + \frac{6g(2r-2)^{g-1}}{(rN-1)(rN-2)(rN-4)} \left(1 + O\left(\frac{g^2}{N}\right)\right) \\ \leq & \frac{8g(2r-2)^{g-1}}{(rN-1)(rN-2)} \left(1 + O\left(\frac{g^2}{N}\right)\right) \\ < & \epsilon \end{aligned}$$

for large N .

Finally, we verify item 4. For any $F \in \mathcal{M}$, we estimate $\sum_{M \in \mathcal{M}_F} \frac{\Pr(A_M)}{r^{N-3k}}$. Recall,

$$\mathcal{M}_F = \{M \setminus F \mid M \in \mathcal{M}, M \neq F, M \cap F \neq \emptyset, F \text{ does not conflict with } M\}.$$

If the projection of F is a 1-cycle, then $\mathcal{M}_F = \emptyset$ (every matching $M \in \mathcal{M}$ either conflicts with F or is identical to F), so there is nothing to do.

Now we assume the projection of F is an i -cycle C_i with $2 \leq i \leq g-1$. Let $M = M' \setminus F$ be such that $M \in \mathcal{M}_F$ and M' projects to a j -cycle C_j with $2 \leq j \leq g-1$. If M' projects to a degenerate 2-cycle (i.e. two identical edges), then $M' \setminus F = \emptyset$ for any M' under consideration. This this case increases the cardinality of \mathcal{M}_F by at most one, so we may disregard it in our asymptotic analysis.

We need a definition before proceeding. Let $\{e_\alpha\}_{\alpha=1}^p$ be a collection of edges of K_N^s and write $e_\alpha = \{v_1^\alpha, \dots, v_s^\alpha\}$ for each α . The collection forms a **loose path** of

length p (hereinafter, p -**path** or simply **path**) provided the *only* equalities among the vertices are $v_s^\alpha = v_1^{\alpha+1}$ for each $\alpha \in [p-1]$. Thus, there are exactly $p-1$ vertices that belong to two edges of the p -path and the rest belong to only one edge. For example, the edges $\{a, b, c\}$, $\{c, d, e\}$, and $\{e, f, g\}$ form a loose path of length 3 in K_N^3 , where each distinct letter denotes a distinct vertex. Notice that a p -path in a 3-uniform hypergraph contains $2p+1$ distinct vertices.

The minivertices in $M' \cap F$ form, after projection, a collection of loose paths P_1, \dots, P_t in $C_i \cap C_j$. (A path may consist of a single edge.) Let m denote the total number of edges among all the paths. Fixing these paths (and the edges in $M' \cap F$), we must choose some additional ℓ vertices to make C_j . In fact, we can specify the value of ℓ exactly in terms of j , t , and m . The cycle C_j contains a total of $2j$ distinct vertices. Since each p -path has $2p+1$ distinct vertices, it follows that the t paths in total represent $2m+t$ distinct vertices. We conclude $\ell = 2j - 2m - t$.

Momentarily regarding the paths as featureless points, the vertices and paths can be arranged on the cycle in $\frac{1}{2(\ell+t)}(\ell+t)! \leq (\ell+t-1)!$ ways.

Each path may be integrated into the cycle in eight ways. First, we choose which of the paths ends will be the “left” end (that is, the end that will be set adjacent to the paths neighbors to the left on the cycle). Next, the leftmost edge of the path must have a vertex in common with the edge of C_j to its left. There are two free vertices from which to choose. Similarly, there are two choices for the rightmost edge of the path. Multiplying these choices yields the eight possibilities.

Taking all this together, the number of possible cycles C_j with t fixed paths is at most

$$8^t \binom{N}{\ell} (\ell+t-1)!$$

for fixed j , t , and m .

Now, the minivertices defining $M' \cap F$ are fixed, but we have some freedom to choose the minivertices defining M . The t paths of $M' \cap F$ break the edges of M into

t paths, as well (the “gaps” between each path). Let m_q denote the number of edges belonging to the q^{th} gap, so that $\sum_{q=1}^t m_q = j - m$. There are

$$\prod_{q=1}^t r^{2m_q-1}(r-1)^{m_q+1}$$

choices for the minivertices that will form M after projection. We have r choices for each minivertex in the q^{th} path excluding the two endpoints (since the first of these minivertices has already been fixed by $M' \cap F$), of which there are $2m_q - 1$. We have $r - 1$ choices for each minivertex belonging to the intersection of two edges including the two endpoints, of which there are $m_q + 1$. Multiplying the choices together, we have

$$\begin{aligned} \prod_{q=1}^t r^{2m_q-1}(r-1)^{m_q+1} &= r^{2(j-m)-t}(r-1)^{j-m+t} \\ &= r^\ell(r-1)^{j-m+t} \end{aligned}$$

possible matchings M' defining C_j with $M' \cap F$ fixed.

To specify the t fixed paths, fix some orientation of C_i and choose $2t$ vertices v_1, \dots, v_{2t} from among the vertices of C_i that belong to exactly two edges of C_i . There are exactly i such vertices. From this collection, we can specify the paths P_1, \dots, P_t in two ways. One way to specify the paths is to take P_γ to be all edges of C_i between the vertices $v_{2\gamma-1}$ and $v_{2\gamma}$ for $1 \leq \gamma \leq t$ (where “between” means “starting with $v_{2\gamma-1}$ and ending at $v_{2\gamma}$ according to the fixed orientation”). The second way is to take P_γ to be all edges of C_i between the vertices $v_{2\gamma}$ and $v_{2\gamma+1}$ (where we understand v_{2t+1} to be v_1).

There are $j - m$ edges belonging to $M = M' \setminus F$. Since Equation (C.2) is decreasing

in i , we obtain

$$\begin{aligned}
\Pr_{rN-3(g-1)}(A_M) &\leq \frac{2^{j-m}}{(rN-3(g-1))_{3(j-m),3}} \\
&\leq \frac{2^{j-m}}{(rN-3(g-1)-3(j-m)+1)^{2j-2m}} \\
&\leq \frac{2^{j-m}}{(rN-6g)^{2j-2m}} \\
&= \frac{2^{j-m}}{(rN-6g)^{\ell+t}}.
\end{aligned}$$

Summarizing, we have

$$\begin{aligned}
&\sum_{M \in \mathcal{M}_F} \Pr_{rN-3(g-1)}(A_M) \\
&\leq \sum_{j=2}^{g-1} \sum_{t=1}^{\lfloor \frac{j}{2} \rfloor} \sum_{m=t}^{i-1} 2 \binom{i}{2t} 8^t \binom{N}{\ell} (\ell+t-1)! r^\ell (r-1)^{j-m+t} \frac{2^{j-m}}{(rN-6g)^{\ell+t}} \\
&\leq 2 \sum_{j=2}^{g-1} 2^j (r-1)^j \sum_{t=1}^{\lfloor \frac{j}{2} \rfloor} \binom{i}{2t} 8^t (r-1)^t \sum_{m=t}^{i-1} \binom{N}{\ell} (\ell+t-1)! r^\ell (r-1)^{-m} \frac{2^{-m}}{(rN-6g)^{\ell+t}}.
\end{aligned}$$

Since $\ell+t-1 = 2j-2m-1$, we have $(\ell+t-1)! = \ell!(\ell+t-1)_{t-1} \leq \ell!(2j-2m)^{t-1}$,

which gives

$$\begin{aligned}
&\sum_{M \in \mathcal{M}_F} \Pr_{rN-3(g-1)}(A_M) \\
&\leq 2 \sum_{j=2}^{g-1} 2^j (r-1)^j \sum_{t=1}^{\lfloor \frac{j}{2} \rfloor} \binom{i}{2t} 8^t (r-1)^t \sum_{m=t}^{i-1} \binom{N}{\ell} (2j-2m)^{t-1} r^\ell (r-1)^{-m} \frac{2^{-m}}{(rN-6g)^{\ell+t}} \\
&\leq \sum_{j=2}^{g-1} 2^j (r-1)^j \sum_{t=1}^{\lfloor \frac{j}{2} \rfloor} \binom{i}{2t} 2^{4t} (r-1)^t (j-t)^{t-1} \sum_{m=t}^{i-1} \binom{N}{\ell} r^\ell (r-1)^{-m} \frac{2^{-m}}{(rN-6g)^{\ell+t}}.
\end{aligned}$$

There is an absolute upper bound $K_1 > \frac{N^t(N)_\ell r^{\ell+t}}{(rN-6g)^{\ell+t}}$. Making use of this estimate,

we have

$$\begin{aligned}
& \sum_{M \in \mathcal{M}_F} \Pr_{rN-3(g-1)}(A_M) \\
& \leq K_1 \sum_{j=2}^{g-1} 2^j (r-1)^j \sum_{t=1}^{\lfloor \frac{j}{2} \rfloor} \binom{i}{2t} 2^{4t} (r-1)^t (j-t)^{t-1} \sum_{m=t}^{i-1} (rN)^{-t} (r-1)^{-m} 2^{-m} \\
& = K_1 \sum_{j=2}^{g-1} 2^j (r-1)^j \sum_{t=1}^{\lfloor \frac{j}{2} \rfloor} \binom{i}{2t} 2^{4t} (r-1)^t (j-t)^{t-1} (rN)^{-t} \sum_{m=t}^{i-1} (r-1)^{-m} 2^{-m} \\
& = K_1 \sum_{j=2}^{g-1} 2^j (r-1)^j \sum_{t=1}^{\lfloor \frac{j}{2} \rfloor} \binom{i}{2t} 2^{4t} (r-1)^t (j-t)^{t-1} (rN)^{-t} \sum_{m=t}^{i-1} (2r-2)^{-m} \\
& < K_1 \sum_{j=2}^{g-1} 2^j (r-1)^j \sum_{t=1}^{\lfloor \frac{j}{2} \rfloor} \binom{i}{2t} 2^{4t} (r-1)^t (j-t)^{t-1} (rN)^{-t} \sum_{m=t}^{\infty} (2r-2)^{-m} \\
& = K_1 \sum_{j=2}^{g-1} 2^j (r-1)^j \sum_{t=1}^{\lfloor \frac{j}{2} \rfloor} \binom{i}{2t} 2^{4t} (r-1)^t (j-t)^{t-1} (rN)^{-t} (2r-2)^{-t} \sum_{m=0}^{\infty} (2r-2)^{-m} \\
& = K_1 \sum_{j=2}^{g-1} 2^j (r-1)^j \sum_{t=1}^{\lfloor \frac{j}{2} \rfloor} \binom{i}{2t} 2^{3t} (j-t)^{t-1} (rN)^{-t} \cdot \frac{1}{1 - (2r-2)^{-1}} \\
& \leq K_2 \sum_{j=2}^{g-1} 2^j (r-1)^j \sum_{t=1}^{\lfloor \frac{j}{2} \rfloor} \binom{i}{2t} 2^{3t} (j-t)^{t-1} (rN)^{-t} \\
& = K_2 \sum_{j=2}^{g-1} 2^j (r-1)^j \sum_{t=1}^{\lfloor \frac{j}{2} \rfloor} \frac{\binom{i}{2t}}{j-t} \left(\frac{8(j-t)}{rN} \right)^t.
\end{aligned}$$

For large N , the last summation has the largest term at $t = 1$. To see this, write

$$f(t) = \frac{\binom{i}{2t}}{j-t} \left(\frac{8(j-t)}{rN} \right)^t$$

and consider

$$\begin{aligned}
\frac{f(t)}{f(t+1)} &= \frac{(2t+2)(2t+1)Nr(j-t)^{t-1}}{8(i-2t)(i-2t-1)(j-t-1)^t} \\
&\geq \frac{N(j-t)^{t-1}}{8(i-2t)(i-2t-1)(j-t-1)^t} \\
&\geq \frac{N}{8(i-2t)(i-2t-1)(j-t-1)} \\
&\geq \frac{N}{8g^3},
\end{aligned}$$

which is greater than 1 for large N .

Replacing $f(t)$ with $f(1)$ and summing at most g terms, we have

$$\begin{aligned}
& \sum_{M \in \mathcal{M}_F} \Pr_{rN-3(g-1)}(A_M) \\
& \leq K_2 \sum_{j=2}^{g-1} 2^j (r-1)^j g \frac{i(i-1)}{2(j-1)} \frac{8(j-1)}{rN} \\
& = K_3 \sum_{j=2}^{g-1} 2^j (r-1)^j g \frac{i(i-1)}{rN} \\
& \leq K_3 \sum_{j=2}^{g-1} \frac{2^j (r-1)^j g^3}{rN} \\
& = \frac{K_3 g^3}{rN} \sum_{j=2}^{g-1} 2^j (r-1)^j \\
& = \frac{K_3 g^3}{rN} \sum_{j=2}^{g-1} (2r-2)^j \\
& \leq \frac{K_3 g^3}{rN} 2(2r-2)^{g-1} \\
& \leq \frac{K_4 (2r-2)^{g-2} g^3}{N} \\
& < \epsilon.
\end{aligned}$$

To apply Theorem 2.3, we need $k\epsilon = o(1)$ and $k\mu\epsilon = o(1)$. We know $k < g$ and chose $\epsilon = \frac{K(2r-2)^{g-2}g^3}{N}$ (where we take $K = K_4$).

We claim $\mu = O\left(\frac{(2r-2)^{g-1}}{g}\right)$. Starting with $\mu \leq \sum_{i=1}^{g-1} \frac{(2r-2)^i}{2i}$, we write

$$\sum_{i=1}^{g-1} \frac{(2r-2)^i}{2i} = \sum_{i=1}^{\lfloor \frac{g-1}{2} \rfloor} \frac{(2r-2)^i}{2i} + \sum_{i=\lceil \frac{g-1}{2} \rceil}^{g-2} \frac{(2r-2)^i}{2i} + \sum_{i=g-1}^{g-1} \frac{(2r-2)^i}{2i}$$

and show that each piece is $O\left(\frac{(2r-2)^{g-1}}{g}\right)$.

The first summation contains $\lfloor \frac{g-1}{2} \rfloor$ terms, each of which is at most $(2r-2)^{\lfloor \frac{g-1}{2} \rfloor}$,

so

$$\sum_{i=1}^{\lfloor \frac{g-1}{2} \rfloor} \frac{(2r-2)^i}{2i} \leq \left\lfloor \frac{g-1}{2} \right\rfloor (2r-2)^{\lfloor \frac{g-1}{2} \rfloor}.$$

Now,

$$\left\lfloor \frac{g-1}{2} \right\rfloor (2r-2)^{\lfloor \frac{g-1}{2} \rfloor} = O\left(\frac{(2r-2)^{g-1}}{g}\right)$$

if and only if

$$g^2 = O\left((2r-2)^{\lceil \frac{g-1}{2} \rceil}\right).$$

The latter claim holds since g^2 is polynomial in g , while $(2r-2)^{\lceil \frac{g-1}{2} \rceil} \geq 4^{\lceil \frac{g-1}{2} \rceil}$ represents exponential growth in g .

For the second summation,

$$\begin{aligned} \sum_{i=\lceil \frac{g-1}{2} \rceil}^{g-2} \frac{(2r-2)^i}{2i} &\leq \frac{1}{2^{\lceil \frac{g-1}{2} \rceil}} \sum_{i=\lceil \frac{g-1}{2} \rceil}^{g-2} (2r-2)^i \\ &\leq \frac{1}{2^{\lceil \frac{g-1}{2} \rceil}} \cdot \frac{(2r-2)^{g-1} - 1}{(2r-2) - 1} \\ &\leq \frac{1}{2^{\lceil \frac{g-1}{2} \rceil}} \cdot (2r-2)^{g-1} \\ &= O\left(\frac{(2r-2)^{g-1}}{g}\right). \end{aligned}$$

Finally, the last summation contains only the single term $\frac{(2r-2)^{g-1}}{2(g-1)} = O\left(\frac{(2r-2)^{g-1}}{g}\right)$.

Returning to Theorem 2.3, we have

$$k\epsilon < \frac{K(2r-2)^{g-2}g^4}{N}$$

and

$$k\epsilon\mu < \frac{K(2r-2)^{2g-3}g^3}{N}$$

We may therefore apply Theorem 2.3 provided $(2r-2)^{2g-3}g^3 = o(N)$, which is assumed in Condition (C.1). The neglect of error in (C.3) is also allowed by (C.1). \square