

Problems in Extremal and Probabilistic Combinatorics



Natasha Morrison
Merton College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2017

For Kaye, who never knew about this adventure
but was with me every step along the way.

Preface

Years ago my mother used to say to me, she'd say, "In this world, Elwood, you must be" - she always called me Elwood - "In this world, Elwood, you must be oh so smart or oh so pleasant." Well, for years I was smart. I recommend pleasant. You may quote me.

—ELWOOD P. DOWD

Autumn 2012, I arrived in Cambridge to do Part III not knowing what to expect. Coming from Durham, it was hard to imagine being able to stay afloat in a sea of the best, trained by the very best. But I wanted to prove to myself that I could. I thought that, if everything went well, perhaps I would do a PhD in number theory. I'm not sure that I had even come across the word 'combinatorics' before reading the Part III course list.

Never would I have guessed that I would be here four and a half years later, submitting a thesis for a PhD, at the University of Oxford. A thesis about combinatorics, this mysterious subject of which I still have so much to discover. I wasn't sure then how things would turn out. But I am sure that I would not be here if it were not for a number of people whose paths I have been fortunate enough to cross on this journey. Most are probably unaware of the extent of the impact their kindness, time and friendship has had. I don't think I can adequately express my thanks in a few sentences, but I will do my best.

I am incredibly grateful to my supervisor Alex Scott for taking me on as a PhD student; for his time and attention over the past few years; for teaching me that this is a marathon, not a sprint; for his advice about how to run this particular marathon, and; for sharing his mathematical insights, both into which problems are interesting and into how to attack these problems. I am especially thankful for the encouragement to follow my own research interests and the freedom to travel and collaborate with others: without these opportunities Chapters 4 and 5 would not exist.

I would like to particularly thank those who have spent some of their valuable time during the last few years collaborating with me and teaching me. Specifically, Rob Morris for inviting me to IMPA, taking the time to teach me new

maths and for all the helpful advice over the past two years, especially about how to be a proper Carioca. I'd also like to particularly thank Bruce Reed for the opportunities he has given me and for helping me realise that I too can be a penguin.

In the future when I look back on this time, the prevalent memories won't be of the days spent staring at blackboards, the weekends spent editing latex or the nights spent convinced I had *almost* proved something, only to be astonished at what nonsense I had believed the following morning. I won't relive the successes or regret the failures. What will survive the test of time is the memories of the amazing places I have been and the friends I have made all over the world, friends who have supported me unconditionally and who I hope will remain close in the years to come.

In particular, I'd like to thank Marthe for being an excellent walrus and an even better friend; Michał, for pretending to pretend to be my friend and for giving me time and advice especially during the soul-destroying period of applying for jobs; Lena for being incredibly kind whilst I was trying to write this thesis and for sharing my appreciation of juice and the deli; Chloë for our coffee time and always being on my side, even when I am wrong; Matthew for always being there to help with anything that worries me and to make me feel better. Doing a PhD has sometimes been infinitely far from lying on a beach in Rio (despite what some people seem to think) and I'd like to specially thank Jon for being here through all the good parts and all the bad parts.

I am greatly indebted to Imre Leader and Andrew Thomason for introducing me to the world of combinatorics during Part III and for making it such an appealing subject that I should choose to spend the next four years of my life thinking about it. I am incredibly grateful to Imre for his constant encouragement and advice over the past four years. I believe that if it were not for this encouragement, I would not be here in Oxford.

To get here to Oxford, I had to survive Part III. This would have been a much more difficult task if it were not for the friends I made in Cambridge. In particular, Anja, Aldo, Craig, Ewa, Heiki, Joe, Megan, Pedro, for the study sessions until 2am, for teaching me the things you already knew that I hadn't learnt in Durham, and for the magical memories I made in Cambridge. Maths can often be very isolating, but I never felt like I had to go through Part III alone; I had a team and we all had each others backs. And to António I would like to extend

my deepest thanks. In addition to the above, for believing in me more than anyone has ever believed in me, for making the hot chocolate I will never forget and for helping me remember why I like maths when I need to.

I would like to thank John Parker, Herbert Gangl and Jens Funke for their infectious enthusiasm for maths and for the encouragement to continue doing maths after Durham. I would also like to thank Eva Schumacher-Reid for her care and support throughout my time in Durham.

Also, thanks are due to Sandy Patel and Lynn Featherstone, for answering every question I have ever had (no matter how stupid or how incompetent I must have seemed at the time) and as a result making my life easier in dozens of ways.

I would also like to express my deepest gratitude to my parents for supporting me in everything I have chosen to do and for being proud of me however it has turned out. Thank you to my sisters Karina and Sas for loving me unconditionally over the mountains and through the molehills. Special thanks to Karina for her unbeatable enthusiasm about hypercubes (check out Chapters 3 and 4!), and Sas for the Grey's marathons which help me forget about maths when I need to.

I have spent most of the last 5 years trying so so hard to be oh so smart, often at the expense of being oh so pleasant. I have missed holidays, I have spent weekends in the department rather than visiting my sisters because I did not feel as if I had done enough. I am sorry that this meant that I could not always be around. I think it is hard to understand how consuming maths can be from the outside, but I would like to thank my family for trying and for participating enthusiastically in episodes of the Great Morrison Bakeoff, which provided some much needed light relief during the past year. I think though that if this is now enough, then from now on I don't need to choose to be oh so smart or oh so pleasant. I can be both.

Abstract

In this thesis we consider some problems in extremal and probabilistic combinatorics.

In Chapter 2 we determine the maximum number of induced cycles that can be contained in a graph on $n \geq n_0$ vertices, and show that there is a unique graph that achieves this maximum. This answers a question of Tuza.

Let Q_d denote the hypercube of dimension d . Given $d \geq m$, a spanning subgraph G of Q_d is said to be (Q_d, Q_m) -saturated if it does not contain Q_m as a subgraph but adding any edge of $E(Q_d) \setminus E(G)$ creates a copy of Q_m in G . In Chapter 3, we show that for every fixed $m \geq 2$ the minimum number of edges in a (Q_d, Q_m) -saturated graph is $\Theta(2^d)$. This answers a question of Johnson and Pinto.

We also answer another question of Johnson and Pinto about weak saturation. Given graphs F and H , a spanning subgraph G of F is said to be *weakly* (F, H) -saturated if the edges of $E(F) \setminus E(G)$ can be added to G one at a time so that each additional edge creates a new copy of H . We determine the minimum number of edges in a weakly (Q_d, Q_m) -saturated graph for all $d \geq m \geq 1$. More generally, we determine the minimum number of edges in a subgraph of the d -dimensional grid P_k^d which is weakly saturated with respect to ‘axis aligned’ copies of a smaller grid P_r^m .

In Chapter 4 we consider the *r-neighbour bootstrap process* in the hypercube. This process starts with an initial set A_0 of *infected* vertices in a graph G and, at each step of the process, a healthy vertex becomes infected if it has at least r infected neighbours (once a vertex becomes infected, it remains infected forever). If every vertex of G eventually becomes infected, then we say that A_0 *percolates*.

We prove a conjecture of Balogh and Bollobás which says that, for fixed r and $d \rightarrow \infty$, every percolating set in the d -dimensional hypercube has cardinality at least $\frac{1+o(1)}{r} \binom{d}{r-1}$. We also prove an analogous result for multidimensional rectangular grids. Our proofs exploit a connection between bootstrap percolation and weak-saturation. In addition, we improve on the best known upper bound for the minimum size of a percolating set in the hypercube. In particular,

when $r = 3$, we prove that the minimum cardinality of a percolating set in the d -dimensional hypercube is $\left\lceil \frac{d(d+3)}{6} \right\rceil + 1$ for all $d \geq 3$.

Finally, we consider a more general bootstrap process in a hypergraph setting. Given an r -uniform hypergraph \mathcal{H} , the \mathcal{H} -bootstrap process starts with an initial set of infected vertices of \mathcal{H} and, at each step, a healthy vertex becomes infected if there exists a hyperedge of \mathcal{H} in which it is the only healthy vertex. The initial set of infected vertices is said to *percolate* if every vertex of \mathcal{H} is eventually infected.

In Chapter 5, for fixed r and large d , we obtain a sharp threshold for the probability that a p -random set of vertices in a q -random subhypergraph of \mathcal{H} percolates when $p, q = \Theta(d^{-1/(r-1)})$ and \mathcal{H} is any nearly d -regular r -uniform hypergraph with at most $d^{O(1)}$ vertices which satisfies certain ‘codegree’ conditions. As it turns out, for this wide class of hypergraphs, the threshold depends only on r and not on the underlying structure of the hypergraph. We apply this result to obtain a sharp threshold for a variant of the graph bootstrap process for strictly 2-balanced graphs. This result generalises a theorem of Korándi, Peled and Sudakov and the proof involves an application of the differential equations method.

Statement of Originality

Chapter 2 is based on a joint paper with Alex Scott [105]. Chapter 3 is based on a joint paper with Jonathan A. Noel and Alex Scott [104]. Chapter 4 is based on a joint paper with Jonathan A. Noel [102]. Chapter 5 is based on a forthcoming paper with Robert Morris and Jonathan A. Noel. I have not submitted material from this thesis for any other degree or qualification.

Contents

1	Introduction	1
1.1	Maximising the Number of Induced Cycles in a Graph	1
1.2	Minimum Saturation and Weak Saturation in the Hypercube	3
1.3	Extremal Bounds for Bootstrap Percolation in the Hypercube	5
1.4	Bootstrap Percolation in a Random Hypergraph	7
2	Maximising the Number of Induced Cycles in a Graph	11
2.1	Introduction	11
2.2	Induced Paths Between a Pair of Vertices	13
2.3	Proof of Theorem 1.3	22
2.3.1	Outline of the Proof	23
2.3.2	Preliminary Lemmas	24
2.3.3	The Local Structure Around a ‘Typical’ Vertex	26
2.3.4	Bounding the Number of ‘Atypical’ Vertices	35
2.3.5	Concluding the Proof	41
2.4	Proof of Theorem 1.6	51
2.5	Conclusion	53
3	Minimum Saturation and Weak Saturation in the Hypercube	55
3.1	Introduction	55
3.2	Minimum Saturation	57
3.2.1	Preliminaries	58
3.2.2	The Johnson and Pinto Construction	59
3.2.3	Definitions and Proof Outline	61
3.2.4	Proof of Theorem 1.4	63
3.3	Weak Saturation	70
3.3.1	Outline of the Proof	71
3.3.2	A Lovely Linear Algebraic Lemma	71
3.3.3	Proof of Theorem 3.1	73

3.4	Open Problems	77
3.4.1	Minimum Saturation	77
3.4.2	Weak Saturation	78
3.4.3	Semi-saturation	79
4	Extremal Bounds for Bootstrap Percolation in the Hypercube	81
4.1	Introduction	81
4.2	Preliminaries and Proof Outline	84
4.3	The Hypercube Case	86
4.4	General Grids	92
4.5	Upper Bound Constructions	97
4.6	An Explicit Linear Algebraic Construction	101
4.7	Concluding Remarks	103
5	Bootstrap Percolation in a Random Hypergraph	105
5.1	Introduction	105
5.2	Outline of the Proof	109
5.2.1	The First Phase and the Differential Equations Method	110
5.2.2	The Second Phase	124
5.2.3	The Subcritical Case	125
5.2.4	The Supercritical Case	128
5.3	Probabilistic Tools	129
5.3.1	Standard Concentration Inequalities	129
5.3.2	Kim–Vu and Janson Inequalities	130
5.3.3	Martingales and Concentration	132
5.3.4	The Probability Space	133
5.4	Preliminaries	134
5.5	Concentration at Time Zero	141
5.5.1	Proof of Lemma 5.3	145
5.6	The First Phase After Time Zero	148
5.6.1	Tracking the W Configurations	150
5.6.2	Tracking the X Configurations	155
5.6.3	Tracking the Y Configurations	158
5.7	The Second Phase in the Subcritical Case	171
5.7.1	Defining \mathcal{A}_m and proof of Lemma 7.9	176
5.7.2	Proof of Theorem 7.10	178
5.8	The Second Phase in the Supercritical Case	197

<i>CONTENTS</i>	iii
5.9 Strictly k -Balanced Hypergraphs	204
Bibliography	209

List of Figures

2.1	The graph H_{24}	12
2.2	An example of a braid in \mathcal{F}_{10}	15
2.3	An illustration of v_1, \dots, v_k and the partition V_1, \dots, V_k . There may be additional edges within each V_i and between clusters (not depicted here). By definition, for any i and $j > i + 1$, there are no edges from v_i to a vertex in V_j . No vertex in $\bigcup_{i=1}^{k-2} V_i$ is a neighbour of y	20
2.4	The graph induced by $N^2[v]$ for most vertices v in H_n	23
2.5	The local structure around a typical vertex z . Note that Lemma 3.9 does not determine whether there are edges within the sets B_1, B_2, B_3 and B_4	30
2.6	There may or may not be edges within the sets A_1, \dots, A_6	30
2.7	There may or may not be edges within the sets A_1, \dots, A_5 . The edges illustrating $N(r_3) \cap N(z)$ are darker to highlight the purpose of Claim 3.11 in the case $i = 3$	31
2.8	There may or may not be edges within the sets A_1, \dots, A_6 . The darker edges illustrate Claim 3.12.	32
2.9	There may or may not be edges within the sets A_1, \dots, A_6 . The darker edges illustrate the edges given by Claim 3.14.	33
2.10	There may or may not be edges within the sets A_1, \dots, A_6 . The darker edges illustrate the edges given by Claim 3.16.	34
2.11	The shaded area is the part of the graph whose structure we wish to determine. The dashed line represents the rest of the braid B	49
3.1	An illustration of the proof of Lemma 2.15 in the case that $m = 3$ and $k = 1$	66
3.2	Two non-isomorphic weakly $([6]^2, [3]^2)$ -saturated graphs with $\text{wsat}([6]^2, [3]^2)$ edges.	77
4.1	A weakly (Q_5, S_4) -saturated graph F constructed inductively from a weakly (Q_4, S_4) -saturated graph F_0 and a weakly (Q_4, S_3) -saturated graph F_1 , each of which is also constructed inductively.	90

4.2	The edge labels where $G = \prod_{i=1}^2 [3]$. The lower label of each pair corresponds to the vertex on the left of the pair. We have $I_{(1,1)}^G = \{1, 3\}$, $I_{(2,2)}^G = \{1, 2, 3, 4\}$ and $I_{(3,3)}^G = \{2, 4\}$, to give three examples of I_v^G	93
4.3	An illustration of the set A_0 constructed in the proof of Theorem 5.1 in the case $r = 3$. Each node represents a copy of Q_{d-3} . The set A_0 consists of a copy of B_i on each node labelled $i \in \{1, 2, 3\}$	98
5.1	The red curve gives an example of the trajectory of $Q(m)$ when $r = 3$ and $c^{r-2}\alpha > \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$. The blue curve is an example of the trajectory of $Q(m)$ when $r = 3$ and $c^{r-2}\alpha < \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$	113
5.2	Some examples of Y configurations in the case $r = 4$. The marked vertices are shaded, the neutral vertices are white, the root is labelled with an R and the central hyperedge is drawn with a thick outline.	118
5.3	Some examples of copies of Y configurations rooted at v in the case $r = 6$. The infected vertices are shaded, the healthy vertices are white and the central hyperedge is drawn with a thick outline. Each copy is labelled by which configuration it is a copy of. Notice that a copy of a configuration could have more infections than marked vertices in the configuration itself, as is demonstrated by the third example.	118
5.4	Here $r = 4$. We present four examples of ways that a copy of $Y^{1,2}$ rooted at v can be created. Such a copy is created if the open hyperedge containing u (drawn with a thick outline) is successfully sampled.	121
5.5	Three examples of secondary configurations X in the case $r = 4$. The marked vertices are shaded, the neutral vertices are white, the root is labelled with an R and the central hyperedge is drawn with a thick outline.	121
5.6	Some examples of W configurations in the case $r = 6$. The marked vertices are shaded and the roots are labelled with an R	124
5.7	Some examples of possible members of $Z_v^{2,2}(m)$ in the case $r = 6$. The root is labelled by R , the infected vertices are shaded, the healthy vertices are white and the central hyperedge is drawn with a thick outline.	127
5.8	Here $r = 6$. Copies of this hypergraph are counted by S only when w is healthy. Copies are counted by S_2 whether w is infected or not. The (unlabelled) vertices shaded dark grey are infected, the unshaded vertices are healthy.	136

5.9 Here $r = 6$. Some examples of ways that a copy of $Y^{3,1}$ rooted at v can be created via (1) when the hyperedge shaded light grey becomes successfully sampled and x becomes infected. The central hyperedge of the copy of $Y^{3,1}$ created is drawn with a thick outline. Infected vertices are shaded dark grey and healthy vertices are unshaded. 163

5.10 Here $r = 6$. Some examples of ways that a copy of $Y^{3,1}$ rooted at v can be created via (2) when the hyperedge shaded light grey becomes successfully sampled and x becomes infected. The central hyperedge of the copy of $Y^{3,1}$ created is drawn with a thick outline. Infected vertices are shaded dark grey and healthy vertices are unshaded. 164

5.11 Here $r = 6$, $i = 3$ and $j = 1$. Two examples of hypergraphs that are not counted in $C(\mathcal{F}', u, \mathcal{G}'')$. If the shaded hyperedge were successfully sampled, this would create a copy of $Y^{3,1}$ whose central hyperedge is drawn with a thick outline. Infected vertices are shaded dark grey and healthy vertices are unshaded. 167

5.12 Here $r = 6$, $i = 3$ and $j = 2$. Three examples of hypergraphs that are unwantedly counted in $C(\mathcal{F}', u, \mathcal{G}'')$. Observe that $|\mathcal{F}' \cap \mathcal{G}'| \geq 2$ and $\mathcal{F}' \cup \mathcal{G}'$ consists of a copy X of a secondary configuration (whose edges are drawn with a thick outline) and some additional copies of W^1 rooted at healthy vertices of X . The central hyperedge of X is the hyperedge containing v . In any of these cases, if the shaded hyperedge were successfully sampled, it would not turn $\mathcal{F}' \cup \mathcal{G}'$ into a copy of $Y^{3,2}$ 167

5.13 Here $\Upsilon := Y^{2,2}$ and $r = 6$. For a particular choice of fruitful $U \subseteq D$ and $U' \subseteq U$, we have an example of some configuration $X' \in \Upsilon^U$ and a configuration $\Upsilon^{U,U'}$. The vertices of U' are labelled with a U' and the vertices of $U \setminus U'$ are labelled with a U . The root is labelled by R , the marked vertices are dark grey, the neutral vertices are white and the central hyperedge is drawn with a thick outline. The hyperedges $\bigcup_{u \in U} e_u$ are shaded light grey. Observe that if each of these hyperedges in a copy of X' is successfully sampled in the $(m + 1)$ th step, then this creates a new copy of $Y^{2,2}$ in $\mathcal{H}(m + 1)$ 181

5.14 A loose 3-uniform triangle. The black circles represent vertices and the ellipses represent hyperedges. 206

Chapter 1

Introduction

In this thesis we consider four broad topics in extremal and probabilistic combinatorics. We will give a short prelude to each topic below, to provide some history and motivation to the problems presented in each of the chapters. A more detailed background to the general areas discussed will be given in the relevant chapters. The history related to bootstrap percolation will be split between Chapters 4 and 5, to place our results in the appropriate context.

The chapters will be self-contained, and hence any results needed from the introduction will be restated there. Each chapter is numbered independently. On the occasions where we wish to reference a result from another chapter, it will be highlighted that the result is stated in a separate chapter. In particular, every result stated in the introduction is treated in this manner. For the benefit of the reader, relevant discussion from this chapter is also repeated where it would be useful to do so.

1.1 Maximising the Number of Induced Cycles in a Graph

One of the main aims of extremal combinatorics is to determine which n -vertex graphs satisfying a particular property contain the maximum number of copies of a fixed subgraph. Perhaps the earliest result of this flavour is Mantel's theorem [97], which was proved in 1907. It states that the unique n -vertex triangle-free graph with the maximum number of edges is the complete bipartite graph with classes of size $\lfloor n/2 \rfloor$ or $\lceil n/2 \rceil$. In 1941, Turán [124] proved a more general statement. This celebrated theorem states that the unique n -vertex K_{k+1} -free graph with the maximum number of edges is the complete k -partite graph with all classes of size $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$.

Since these foundational results, a vast area of mathematics has emerged to develop the theory to answer related questions. For example, tools such as the regularity method (see the excellent surveys [90, 91]) and the use of flag algebras (see e.g. [67, 68, 73, 113, 116])

have been instrumental to the advancement of our understanding of both this area and to other areas of combinatorics. More recently, a particular direction of interest has been to ask these questions about *induced* subgraphs, where a subgraph H of a graph G is *induced* if any edge of G with both endpoints in $V(H)$ is contained in H .

An interesting question to consider is: For a fixed family of graphs \mathcal{H} , what is the maximum possible number of induced copies of members of \mathcal{H} in an n -vertex graph G , and what can one say about the structure of the extremal graphs? This question would not be interesting to ask in the case of non-induced subgraphs, as clearly K_n would contain the maximum number of copies of any fixed subgraph. In this case, the problem only becomes interesting when we require additional properties to be satisfied by G . However, in the induced case, for H a non-complete graph, the very nature of an induced copy of H suggests that the answer to this question is highly non-trivial: something as simple as adding an edge could increase or decrease the number of induced copies of H , with no restrictions on the properties of G . In Chapter 2 we consider this question, where \mathcal{H} is the family of cycles.

In 1989, Tuza (see [128] and [129]) asked: What is the maximum possible number of induced cycles in a graph with n vertices? This question was investigated independently in unpublished work of Robson, who showed in the 1980s that a graph on n vertices has at most $3^{(1+o(1))n/3}$ induced cycles ([99, 118]). In Chapter 2 we answer this question completely. We determine, for sufficiently large n , the maximum number of induced cycles that can be contained in a graph on n vertices and show that for each n there is a unique graph that achieves this maximum.

The main result proved in Chapter 2 is the following.

Theorem 1.3 of Chapter 2. *Let $m(n)$ be the maximum number of induced cycles that can be contained in a graph on n vertices. There exists n_0 such that, for all $n \geq n_0$, there is a unique n -vertex graph H_n containing $m(n)$ induced cycles.*

We also determine the precise structure of H_n (which will be described in Chapter 2 itself). Knowing the structure of H_n gives the exact value for $m(n)$.

Corollary 1.4 of Chapter 2.

$$m(n) = \begin{cases} 3^{n/3} + 12n & \text{for } n \equiv 0 \text{ modulo } 3; \\ 4 \cdot 3^{(n-4)/3} + 12n + 51 & \text{for } n \equiv 1 \text{ modulo } 3; \\ 2 \cdot 3^{(n-2)/3} + 12n - 36 & \text{for } n \equiv 2 \text{ modulo } 3. \end{cases}$$

1.2 Minimum Saturation and Weak Saturation in the Hypercube

The d -dimensional hypercube, which we denote by Q_d , is the graph on vertex set $\{0, 1\}^d$, where two vertices are joined by an edge if and only if they differ in exactly one coordinate. In Chapter 3, we will consider two related extremal problems in the setting of the hypercube, and more generally in the k -grid of dimension d (which we will define in the chapter itself).

Given a graph F and a family of graphs \mathcal{H} , a spanning subgraph G of F is said to be (F, \mathcal{H}) -saturated if it does not contain any $H \in \mathcal{H}$ as a subgraph, but for every edge $e \in E(F) \setminus E(G)$, the graph $G + e$ contains a copy of some $H \in \mathcal{H}$. The *saturation number* $\text{sat}(F, \mathcal{H})$ is the minimum number of edges in an (F, \mathcal{H}) -saturated graph. When \mathcal{H} consists of a single graph H , we write $\text{sat}(F, H)$ for $\text{sat}(F, \mathcal{H})$ and say that a graph is (F, H) -saturated.

In this language, the classical Turán problem asks for the maximum size of an (F, H) -saturated graph; this number is known as the *extremal number*, denoted $\text{ex}(F, H)$. Turán's theorem, mentioned above, resolves this question when F and H are complete graphs. Determining $\text{ex}(F, H)$ is a major area of interest in extremal combinatorics, particularly where F is a complete graph (see e.g. [50, 93]), but also in other settings (see the survey [63]) and in hypergraphs (see the survey [86]). The problem of determining $\text{sat}(F, H)$ can be thought of as a dual problem to the Turán problem and has also been very well-studied in a number of different settings.

A related notion is that of *weak-saturation*. A spanning subgraph G of F is said to be *weakly* (F, \mathcal{H}) -saturated if the edges of $E(F) \setminus E(G)$ can be added to G , one edge at a time, in such a way that every added edge creates a new copy of some $H \in \mathcal{H}$. Note that we do not require here that G contains no copy of H . The minimum number of edges in a weakly (F, \mathcal{H}) -saturated graph is known as the *weak saturation number* and is denoted $\text{wsat}(F, \mathcal{H})$. As above, when \mathcal{H} consists of a single graph H , we write $\text{wsat}(F, H)$ for $\text{wsat}(F, \mathcal{H})$ and say that a graph is *weakly* (F, H) -saturated.

Note that every (F, \mathcal{H}) -saturated graph is also weakly (F, \mathcal{H}) -saturated and so

$$\text{wsat}(F, \mathcal{H}) \leq \text{sat}(F, \mathcal{H}).$$

The study of weak saturation is intimately linked with that of *bootstrap percolation*. The relationship between these concepts will be discussed in more detail in Chapters 3, 4 and 5.

In Chapter 3, we are interested in determining $\text{sat}(Q_d, Q_m)$ and $\text{wsat}(Q_d, Q_m)$ for $d \geq m$. With regards to $\text{sat}(Q_d, Q_m)$, the special case $m = 2$ was studied by Choi and Guan [41] who constructed a (Q_d, Q_2) -saturated graph with at most $(\frac{1}{4} + o(1)) |E(Q_d)|$ edges. Santolupo (see [53]) conjectured that this construction is best possible. In their recent paper [80],

Johnson and Pinto disproved this conjecture (in a strong sense) by showing that, for every fixed m , there exists a (Q_d, Q_m) -saturated graph with $o(|E(Q_d)|)$ edges. More precisely, they proved the following.

Theorem 1.2 of Chapter 3 (Johnson and Pinto [80]). *For every fixed $m \geq 2$, there exists $0 < \varepsilon_m < 1$ such that*

$$\text{sat}(Q_d, Q_m) = O\left(d^{(1-\varepsilon_m)2^d}\right) = O(|E(Q_d)|/d^{\varepsilon_m}).$$

In the case $m = 2$, Johnson and Pinto [80] obtained a stronger bound; namely,

$$\text{sat}(Q_d, Q_2) < 10 \cdot 2^d.$$

That is, for every d , there exists a (Q_d, Q_2) -saturated graph with bounded average degree. Motivated by this result, they asked the following: For which fixed values of m is $\text{sat}(Q_d, Q_m) = O(2^d)$? We provide a construction, which shows that this is the case for every $m \geq 2$.

Theorem 1.4 of Chapter 3. *For every fixed $m \geq 2$,*

$$\text{sat}(Q_d, Q_m) \leq (1 + o(1))72m^22^d.$$

We remark that, previously, the best known lower bound on $\text{sat}(Q_d, Q_m)$ was $(m + 1 - o(1))2^{d-1}$, due to Johnson and Pinto [80]. In particular, this implies that for fixed m , Theorem 1.4 of Chapter 3 is tight up to a (constant) factor of $O(m)$.

In the case of weak saturation, Johnson and Pinto [80] proved for all $d \geq 2$ that

$$\text{wsat}(Q_d, Q_2) = 2^d - 1,$$

by exhibiting a spanning tree of Q_d which is weakly (Q_d, Q_2) -saturated. They asked about the value of $\text{wsat}(Q_d, Q_m)$ for general $d \geq m \geq 1$. We determine this value exactly.

Theorem 1.5 of Chapter 3. *For $d \geq m \geq 1$,*

$$\text{wsat}(Q_d, Q_m) = (m - 1)2^d - \sum_{j=0}^{m-2} (m - 1 - j) \binom{d}{j}.$$

We remark that this theorem is actually a special case of a more general result (Theorem 3.1 of Chapter 3), which we state later. The lower bound for $\text{sat}(Q_d, Q_m)$ mentioned above can be improved to $(m - 1 - o(1))2^d$ by using Theorem 3.1 of Chapter 3 and recalling from above that $\text{wsat}(Q_d, Q_m) \leq \text{sat}(Q_d, Q_m)$.

To give an upper bound on a weak saturation number $\text{wsat}(F, H)$, all that is required is to give a construction of a weakly saturated graph with a certain number of edges. The

difficulty, it seems, lies in proving a matching lower bound, as one must show that *no* graph on fewer edges is weakly (F, H) -saturated. To prove the lower bound in Theorem 3.1 of Chapter 3, we apply a lovely linear algebraic lemma (Lemma 3.2 of Chapter 3) of Balogh, Bollobás, Morris and Riordan [16]. This lemma is particularly useful as it converts the problem of finding a lower bound for $\text{wsat}(F, H)$ into a problem of constructing a vector space satisfying particular properties. In this way it circumvents some of the difficulty associated with proving lower bounds on weak saturation numbers as the lower bound can also be proved ‘constructively’ (however, it is often very difficult to find an appropriate construction). We remark that Lemma 3.2 of Chapter 3 will also be used to prove our main results in Chapter 4.

1.3 Extremal Bounds for Bootstrap Percolation in the Hypercube

Given a positive integer r and a graph G , the r -neighbour bootstrap process begins with an initial set of *infected* vertices of G and, at each step of the process, a vertex becomes infected if it has at least r infected neighbours. More formally, if A_0 is the initial set of infected vertices, then the set of vertices that are infected after the j th step of the process for $j \geq 1$ is defined by

$$A_j := A_{j-1} \cup \{v \in V(G) : |N_G(v) \cap A_{j-1}| \geq r\},$$

where $N_G(v)$ denotes the neighbourhood of v in G . We say that A_0 *percolates* if $\bigcup_{j=0}^{\infty} A_j = V(G)$.

There are two main types of problem that are studied about the r -neighbour process. The first concerns probabilistic questions, where the initial set of infected vertices is chosen at random. One commonly considered setting is where each vertex is included in A_0 independently with some fixed probability p . We will discuss this setting in greater detail in Chapters 4 and 5. In particular, in Chapter 5 we analyse a bootstrap process on hypergraphs in a random setting. The second type of problem concerns extremal questions: this is the area to which we turn our attention in Chapter 4.

The main extremal problem in bootstrap percolation is to determine the minimum cardinality of a set which percolates under the r -neighbour bootstrap process on G , we denote this by $m(G, r)$. An important case is when G is the d -dimensional hypercube Q_d (defined in the previous section). Balogh and Bollobás [10] made the first progress in this case by determining $m(Q_d, 2)$.

Theorem 1.1 of Chapter 4 (Balogh, Bollobás [10]).

$$m(Q_d, 2) = \left\lceil \frac{d}{2} \right\rceil + 1.$$

Motivated by the study of the r -neighbour process in a random setting, they also made the following conjecture (see also [14, 16]): For fixed $r \geq 3$ and $d \rightarrow \infty$,

$$m(Q_d, r) = \frac{1 + o(1)}{r} \binom{d}{r-1}.$$

As shown in [14], the upper bound of this conjecture is not difficult to prove via a construction based on *Steiner systems*. For $0 \leq i \leq d$, define L_i to be the subset of Q_d containing all vertices with exactly i ones in their coordinate representation. We say a family $\mathcal{S} \subseteq L_i$ is a *Steiner system* if every member of L_{i-1} is contained in exactly one member of \mathcal{S} . As each member of \mathcal{S} contains i distinct members of L_{i-1} , we have $|\mathcal{S}| = \frac{1}{i} \binom{d}{i-1}$. An *approximate Steiner system* is a family $\mathcal{A} \subseteq L_i$ of cardinality $\frac{1+o(1)}{i} \binom{d}{i-1}$ such that every member of L_{i-1} is contained in some member of \mathcal{A} . The existence of approximate Steiner systems was proved by Rödl [119].

To prove the upper bound of the conjecture, let A_0 be the union of L_{r-2} and an approximate Steiner system $\mathcal{A} \subseteq L_r$. We show that when $d \geq 2r$, the family A_0 percolates under the r -neighbour process on Q_d . Observe that every set in L_{r-1} has r neighbours in A_0 : $(r-1)$ in L_{r-2} and one in \mathcal{A} . Thus every vertex in L_{r-1} becomes infected at the first time step. As $d \geq 2r$, for $r \leq i \leq d$, every $x \in L_i$ has at least r neighbours in L_{i-1} and for $0 \leq i \leq r-3$, every $x \in L_i$ has at least r neighbours in L_{i+1} . Thus L_{r-1} percolates under the r -neighbour process on Q_d and so does A_0 .

We have

$$|A| = |A_0| + |\mathcal{A}| = \binom{d}{r-2} + \frac{1+o(1)}{r} \binom{d}{r-1} = \frac{1+o(1)}{r} \binom{d}{r-1}, \quad (3.1)$$

which proves the upper bound of the conjecture. Expanding the penultimate part of (3.1) gives

$$m(Q_d, r) \leq \frac{d^{r-1}}{r!} + \frac{d^{r-2}(r+2)}{2r(r-2)!} + O(d^{r-3}). \quad (3.2)$$

Note that, under certain ‘divisibility’ conditions on d and r , the approximate Steiner system in this construction can be replaced with an exact Steiner system (using, for example, the celebrated result of Keevash [87]). In this special case, the percolating set has cardinality $\frac{1}{r} \binom{d}{r-1} + \binom{d}{r-2}$.

Lower bounds have been far more elusive; previously, the best known lower bound on $m(Q_d, r)$ for fixed $r \geq 3$ was only linear in d (see [14]). In Chapter 4, we prove the conjecture of Balogh and Bollobás and hence substantially improve the best known lower bound on $m(Q_d, r)$ when $r \geq 3$.

Theorem 1.3 of Chapter 4. For $d \geq r \geq 1$,

$$m(Q_d, r) \geq 2^{r-1} + \sum_{j=1}^{r-1} \binom{d-j-1}{r-j} \frac{j2^{j-1}}{r}$$

where, by convention, $\binom{a}{b} = 0$ when $a < b$.

For fixed $r \geq 3$, this theorem implies

$$m(Q_d, r) \geq \frac{d^{r-1}}{r!} + \frac{d^{r-2}(6-r)}{2r(r-2)!} + \Omega(d^{r-3}),$$

which differs from the upper bound in (3.2) by an additive term of order $\Theta(d^{r-2})$. In Chapter 4, we will also provide a recursive upper bound on $m(Q_d, r)$, which improves on the second order term of (3.2). For $r = 3$, we combine this recursive bound with some additional arguments to show that our bound is tight in this case.

Theorem 1.8 of Chapter 4. For $d \geq 3$, we have $m(Q_d, 3) = \left\lceil \frac{d(d+3)}{6} \right\rceil + 1$.

In order to prove our general lower bound on $m(Q_d, r)$, we will exploit a relationship between bootstrap percolation and weak saturation (the notion of weak saturation was defined in the previous section). We will apply Lemma 3.2 of Chapter 3 to solve a related weak saturation problem, this in turn will imply the required lower bound for the conjecture of Balogh and Bollobás.

1.4 Bootstrap Percolation in a Random Hypergraph

In our final chapter, we consider a probabilistic question about a different bootstrap process to the one introduced in the previous section. Given a hypergraph \mathcal{H} , the \mathcal{H} -bootstrap process begins with an initial set of *infected* vertices of \mathcal{H} (a vertex that is not infected is *healthy*) and, at each step, a healthy vertex becomes infected if there exists a hyperedge of \mathcal{H} in which it is the unique healthy vertex. The set of initially infected vertices is said to *percolate* if every vertex of \mathcal{H} is eventually infected. This process was first studied by Balogh, Bollobás, Morris and Riordan [16] and is motivated by numerous connections to other variants of bootstrap percolation.

In particular it is a generalisation of the *graph bootstrap process*. Given graphs F and H , the F -bootstrap process on H starts with an initial set E_0 of infected edges of H and, at each step, a healthy edge becomes infected if there exists a copy of F in H in which it is the unique healthy edge. If every edge of H becomes infected, we say that E_0 *percolates*. Notice that the set E_0 percolates if and only if the subgraph G of H with $V(G) = V(H)$ and $E(G) = E_0$ is weakly (H, F) -saturated. Thus, in some sense, the study of the graph bootstrap process

goes back to the original paper of Bollobás [29] on weak saturation. Determining the minimum size of such a set E_0 is equivalent to determining $\text{wsat}(H, F)$. More about weak saturation can be found in Chapters 3 and 4. In Chapter 5 we are concerned with the study of bootstrap processes in the random setting, and so to be consistent with the literature, we will refer to the ‘graph bootstrap process’, rather than using terminology which relates to weak-saturation.

For a finite set X and $p \in (0, 1)$ let X_p denote a random subset of X obtained by including each element of X with probability p independently of one another. The *critical probability* for the F -bootstrap process on H is defined by

$$p_c(H, F) := \inf \{p \in (0, 1) : \mathbb{P}(E(H)_p \text{ percolates}) \geq 1/2\}.$$

Balogh, Bollobás and Morris [15] initiated the study of $p_c(F, H)$, determining $p_c(K_n, K_r)$ up to a polylogarithmic factor and up to a constant when $r = 4$. They also found upper bounds for a wider family of graphs. This process has since been studied by Gunderson, Koch and Przykucki [69]. Work subsequently done by Angel and Kolesnik [8] and Kolesnik [89] improves on [15], determining $p_c(K_n, K_4)$ up to a factor of $(1 + o(1))$.

The F -bootstrap process on H is equivalent to the $\mathcal{H}_{H,F}$ -bootstrap process where $\mathcal{H}_{H,F}$ is the hypergraph in which each vertex of $\mathcal{H}_{H,F}$ corresponds to an edge of H and the hyperedges of $\mathcal{H}_{H,F}$ are precisely the edge sets of copies of F in H . An analogous relationship between the \mathcal{H} -bootstrap process and a ‘vertex version’ of the graph bootstrap process will be discussed in Chapter 3.

Our goal in Chapter 5 is to estimate the analogous notion of the *critical probability* of the \mathcal{H} -bootstrap process, denoted $p_c(\mathcal{H})$, which is the infimal density at which a random subset of $V(\mathcal{H})$ is likely to percolate. More formally,

$$p_c(\mathcal{H}) := \inf \{p \in (0, 1) : \mathbb{P}(V(\mathcal{H})_p \text{ percolates}) \geq 1/2\}.$$

We will be interested in estimating the quantity $p_c(\mathcal{H}_q)$, where \mathcal{H} is a ‘sufficiently well behaved’ hypergraph and, for $q \in (0, 1)$, \mathcal{H}_q denotes the hypergraph obtained from \mathcal{H} by including each hyperedge of \mathcal{H} independently with probability q .

The main result of Chapter 5 (Theorem 1.6) requires some technical definitions to state and so we will postpone the full statement until that chapter. To give a flavour of the result, we will discuss its main application, which is a sharp threshold for a variant of the graph bootstrap process for *strictly 2-balanced graphs*. In order to state this result, we require some definitions.

Definition 1.7 of Chapter 5. For a graph F with at least two edges, the *2-density* of F is defined to be

$$d_2(F) := \frac{|E(F)| - 1}{|V(F)| - 2}.$$

Definition 1.8 of Chapter 5. A graph F with at least two edges is said to be *2-balanced* if $d_2(F) \geq d_2(F')$ for every proper subgraph F' of F with at least two edges. If the inequality is strict for every such F' , then we say that F is *strictly 2-balanced*.

For F a strictly 2-balanced graph, we will apply Theorem 1.6 of Chapter 5 to the hypergraph $\mathcal{H}_{K_n, F}$. This hypergraph is $|E(F)|$ -uniform and $d(n, F)$ -regular, where $d(n, F) = \Theta(n^{|V(F)|-2})$ (the constant factor is related to the number of automorphisms of F which fix an edge). We will prove the following.

Theorem 1.9 of Chapter 5. *Let F be a strictly 2-balanced graph with at least three edges and define $r := |E(F)|$, $\mathcal{H} := \mathcal{H}_{K_n, F}$ and $d := d(n, F)$. Let $q := \alpha d^{-1/(r-1)}$ for some fixed $\alpha > 0$. Then*

$$p_c(\mathcal{H}_q) = \left(\frac{r-2}{\alpha^{1/(r-2)}(r-1)^{(r-1)/(r-2)} + o(1)} \right) \cdot d^{-1/(r-1)},$$

with high probability as $n \rightarrow \infty$.

We remark that, for F strictly 2-balanced, this theorem can be viewed as a threshold for a variant of the graph bootstrap process where we first ‘activate’ each copy of F in K_n with probability $\Theta(n^{-1/d_2(F)})$ independently of one another (as $d(n, F) = \Theta(n^{|V(F)|-2})$, we have that $n^{-1/d_2(F)}$ is of the same order as $d^{-1/(r-1)}$) and then, given an initial random set of infected edges of K_n , at each step of the process a healthy edge becomes infected if it is the unique healthy edge in an active copy of F .

We also mention that our work on this topic was initially inspired by a paper of Korándi, Peled and Sudakov [92] whose main result is essentially equivalent to the special case of Theorem 1.9 of Chapter 5 where $F = K_3$ and $\alpha = 1/2$. Our result generalises theirs to a wider class of graphs. Moreover, Theorem 1.9 of Chapter 5 is actually just the first case of the result we do prove (Theorem 9.3 of Chapter 5) about strictly k -balanced hypergraphs (defined in Definition 9.2 of Chapter 5).

Our main theorem in Chapter 5 is a more general result (Theorem 1.6). For fixed r and $\alpha > 0$, we obtain a sharp threshold for the probability that a p -random set of vertices in a q -random subhypergraph of \mathcal{H} percolates when $q = \alpha d^{-1/(r-1)}$ and \mathcal{H} is any nearly d -regular r -uniform hypergraph with at most $d^{O(1)}$ vertices which satisfies certain ‘codegree’ and ‘neighbourhood similarity’ conditions. As it turns out, for the wide class of hypergraphs that we consider, the main term in the threshold depends only on r and α and, in particular,

is independent of the underlying structure of the hypergraph. Our approach for proving Theorem 1.6 of Chapter 5 involves an application of the differential equations method.

Chapter 2

Maximising the Number of Induced Cycles in a Graph

Joint work with Alex Scott

2.1 Introduction

What is the maximum number of induced cycles in a graph on n vertices? For cycles of *fixed* length, this problem has been extensively studied. Indeed, for any fixed graph H , let the *induced density* of H in a graph G be the number of induced copies of H in G divided by $\binom{|G|}{|H|}$; let $I(H; n)$ be the maximum induced density of H over all graphs G on n vertices; and let the *inducibility* of H , denoted $\text{ind}(H)$ be the limit $\lim_{n \rightarrow \infty} I(H; n)$. In 1975, Pippinger and Golumbic [114] made the following conjecture.

Conjecture 1.1 (Pippinger and Golumbic [114]). *For $k \geq 5$, $\text{ind}(C_k) = k!/(k^k - k)$.*

In support of this conjecture, Pippinger and Golumbic [114] showed that an iterated blow-up construction gives a lower bound of $\text{ind}(C_k) \geq k!/(k^k - k)$. They also obtained an upper bound of $\text{ind}(C_k) \leq 2e \cdot k!/k^k$. The only case for which the conjecture is known is when $k = 5$. This was proved by Balogh, Hu, Lidický and Pfender [18] via a flag algebra method. In addition they showed that the maximum density is achieved uniquely by an iterated blow-up of a C_5 . Very recently, Hefetz and Tyomkin [74] improved the upper bound of Pippinger and Golumbic to $(128e/81) \cdot k!/k^k$. Beyond these results, nothing more is known (though see [31, 33, 35, 51, 52, 67, 72, 75] for results on inducibility of other graphs).

In this chapter, we consider the total number of induced cycles in a graph (without restriction on length). As mentioned in Chapter 1, Tuza (see [128] and [129]) asked: What is the maximum possible number of induced cycles in a graph with n vertices?

In this chapter we answer this question. We determine, for sufficiently large n , the maximum number of induced cycles that can be contained in a graph on n vertices and show that for each n there is a unique graph that achieves this maximum (Theorem 1.3).

In order to state our results, it is helpful to have a couple of definitions. As usual, for G a graph define the *neighbourhood* of x to be $N_G(x) := \{y \in V(G) : xy \in E(G)\}$.

Definition 1.2. A graph B is called a *cyclic braid* if there exists $k \geq 3$ and a partition B_1, \dots, B_k of $V(B)$ such that for every $1 \leq i \leq k$ and every $x \in B_i$, we have $B_{i-1} \cup B_{i+1} \subseteq N_B(x) \subseteq B_{i-1} \cup B_i \cup B_{i+1}$, where indices are taken modulo k . For such a partition, the notation $B = (B_1, \dots, B_k)$ is used. The sets B_i are called *clusters* of B ; the *length* of the cyclic braid is the number of clusters. If a cyclic braid contains no edges within its clusters, it is called an *empty cyclic braid*. Observe that an empty cyclic braid is k -partite and, when $k > 3$, is triangle free.

As it turns out, the structure of the n -vertex graph containing the maximum number of induced cycles depends on the value of n modulo 3. For $n \geq 8$, we now define an n -vertex graph H_n separately for each value of n modulo 3. Let $k \geq 3$. Define H_{3k} to be the empty cyclic braid of length k where every cluster has size 3. Define H_{3k+1} to be the empty cyclic braid containing $k - 1$ clusters of size 3 and one of size 4. Finally, define H_{3k-1} to be the empty cyclic braid containing $k - 1$ clusters of size 3 and one of size 2. See Figure 2.1 for an example.

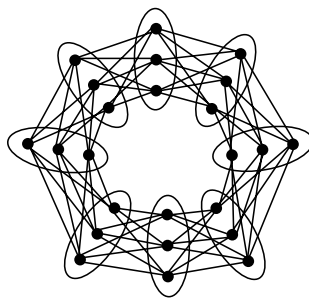


Figure 2.1: The graph H_{24} .

Recall from Chapter 1 that we denote the maximum number of induced cycles that can be contained in a graph on n vertices by $m(n)$. The main result we prove in this chapter is the following.

Theorem 1.3. *There exists n_0 such that, for all $n \geq n_0$, H_n is the unique graph on n vertices containing $m(n)$ induced cycles.*

There are two types of induced cycles in H_n . The first type contains a vertex from every cluster. The second type are induced C_4 s contained within a pair or a triple of clusters. Counting the induced cycles in H_n immediately gives the following corollary to Theorem 1.3.

Corollary 1.4. *There exists n_0 such that, for all $n \geq n_0$:*

$$m(n) = \begin{cases} 3^{n/3} + 12n & \text{for } n \equiv 0 \text{ modulo } 3; \\ 4 \cdot 3^{(n-4)/3} + 12n + 51 & \text{for } n \equiv 1 \text{ modulo } 3; \\ 2 \cdot 3^{(n-2)/3} + 12n - 36 & \text{for } n \equiv 2 \text{ modulo } 3. \end{cases}$$

This implies that $m(n) = \Theta(3^{n/3})$. A *hole* is an induced cycle with at least 4 vertices. For $n \geq 10$, the graph H_n is triangle free and each induced cycle is a hole, so Theorem 1.3 implies the following.

Corollary 1.5. *There exists n_0 such that, for all $n \geq n_0$, H_n is the unique graph on n vertices with the maximum number of holes.*

Using similar arguments to those in the proof of Theorem 1.3 we also prove a stability-type result.

Theorem 1.6. *Fix $0 < \alpha < 1$. There exist constants $C = C(\alpha)$ and $n_0 = n_0(\alpha)$ such that for any $n \geq n_0$, if a graph F on n vertices contains at least $\alpha \cdot m(n)$ induced cycles, then by adding or deleting edges incident to at most $C(\alpha)$ vertices of F , the graph F can be transformed into a cyclic braid with the same cluster sizes as H_n .*

We have also adapted the arguments used to prove Theorem 1.3 to give results about induced cycles of given parity. The details are not presented here, but the results can be found in [105].

The chapter is structured as follows. In Section 2.2 we prove a preliminary result (Theorem 2.6) determining the structure of the n -vertex graphs that maximise the number of induced paths between a particular pair of vertices. The proof of the main theorem (Theorem 1.3) is given in Section 2.3. The proof uses Theorem 2.6 from Section 2.2, but otherwise is entirely contained in Section 2.3. A number of lemmas proved in Section 2.3 are proved in more generality than is strictly needed. This is because the more general versions will be used in Section 2.4, where we prove Theorem 1.6. Finally, in Section 2.5 we conclude by discussing some open questions.

2.2 Induced Paths Between a Pair of Vertices

Let G be a finite graph and let x and y be distinct vertices in $V(G)$. Define $\mathcal{P}(x, y)$ to be the set of all induced paths from x to y in G . Also define:

$$p_2(G) := \max\{|\mathcal{P}(x, y)| : x, y \in V(G)\},$$

and

$$p_2(n) := \max\{p_2(G) : |V(G)| = n\}.$$

We use the notation $p_2(\cdot)$, as it indicates that we are counting the maximum number of induced paths between *two* fixed vertices.

Our first goal in this section is to determine the structure of the n -vertex graphs that contain $p_2(n)$ induced paths between some pair of vertices. We show that these extremal graphs have a particular structure that depends on the value of n modulo 3.

Definition 2.1. A graph B is a *braid* if there exists $k \geq 2$ and a partition B_1, \dots, B_k of $V(B)$ such that:

- (i) for all $1 \leq i \leq k - 1$, every edge between B_i and B_{i+1} is present;
- (ii) if $k \geq 3$ there are no edges between B_1 and B_k ; and
- (iii) for all $1 < i < k$ and all $x \in B_i$,

$$N(x) \subseteq B_{i-1} \cup B_i \cup B_{i+1}.$$

We sometimes use the notation (B_1, \dots, B_k) for B . The sets B_i are called *clusters*. If $i \in \{1, k\}$ we say B_i is an *end cluster*; otherwise we say B_i is a *central cluster*. The *length* of a braid is the number of clusters it contains. For a graph G , we say that a subgraph B is a *braid in G* if:

- (i) B is an induced subgraph of G that is a braid; and
- (ii) no vertex of a central cluster of B has a neighbour in $G \setminus B$.

Note that the end clusters B_1 and B_k are allowed to have neighbours outside of the braid in G but there are no edges between B_1 and B_k . So a cyclic braid *is not* a braid.

Let $n \geq 4$. Define \mathcal{F}_n to be the set of all braids B with the following properties:

- $|V(B)| = n$.
- B has end clusters of size 1.
- If $n \equiv 0$ modulo 3, then either one central cluster has size 4 and the rest have size 3, or two have size 2 and the rest have size 3.
- If $n \equiv 1$ modulo 3, then one central cluster has size 2 and the rest have size 3.
- If $n \equiv 2$ modulo 3, all central clusters have size 3.

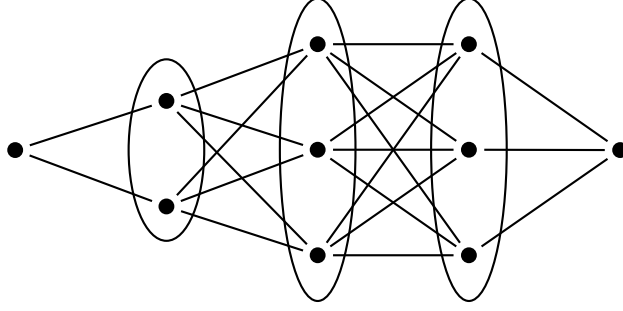


Figure 2.2: An example of a braid in \mathcal{F}_{10} .

Observe that there are no conditions on whether clusters in B contain edges and (as $n \geq 4$) that every braid in \mathcal{F}_n has at least 3 clusters. See Figure 2.2 for an example of a braid in \mathcal{F}_{10} . The following claim will be useful throughout this section.

Claim 2.2. *Let $B = (B_1, \dots, B_k)$ be a braid in a graph G . Let $x \in B_1$ and $y \in B_k$. Then:*

- (i) *If $k = 2$, there is exactly one induced path in G from x to y contained within B .*
- (ii) *If $k \geq 3$, then there are precisely $\prod_{i=2}^{k-1} |B_i|$ induced paths in G from x to y contained within B .*

Proof. When $k = 2$ the result follows by definition of B , as x is a neighbour of y . When $k \geq 3$, by definition of B any induced path $P \subseteq B$ between x and y contains at least one vertex from each cluster. If P were to contain more than one vertex from some cluster, then P would no longer be induced (as every edge between a pair of adjacent clusters is present). So such a path contains exactly one vertex from each cluster and the result follows. \square

We immediately get the following.

Lemma 2.3. *Let $B = (B_1, \dots, B_k)$ be a braid in a graph G with $B_k = \{y\}$. Then for any $1 \leq j \leq k - 1$ and $x \in B_j$, there are precisely $\prod_{i=j+1}^k |B_i|$ induced paths in G from x to y contained within B .*

Proof. As B is a braid in G , for any $1 \leq j \leq k - 1$, (B_j, \dots, B_k) is also a braid in G . Note that any induced path within B from x to y is contained within (B_j, \dots, B_k) . As $|B_k| = 1$, applying Claim 2.2 to the braid (B_j, \dots, B_k) gives the required result. \square

By Lemma 2.3 and definition of \mathcal{F}_n , we have that every F in \mathcal{F}_n with end clusters $\{x\}$ and $\{y\}$ contains the same number of induced paths between x and y . Therefore we define

$$f_2(n) = \begin{cases} 3^{(n-2)/3} & \text{for } n \equiv 2 \text{ modulo } 3 \\ 4 \cdot 3^{(n-6)/3} & \text{for } n \equiv 0 \text{ modulo } 3 \\ 2 \cdot 3^{(n-4)/3} & \text{for } n \equiv 1 \text{ modulo } 3. \end{cases}$$

Note that for all $n \geq 4$,

$$f_2(n) \leq 3^{\frac{n-2}{3}}. \quad (2.4)$$

Claim 2.5. *Let $n \geq 4$ and let $F \in \mathcal{F}_n$. Then*

$$p_2(F) = f_2(n),$$

where $\{x\}$ and $\{y\}$ are the end clusters of F .

Proof. Let $F = (B_1, \dots, B_k) \in \mathcal{F}_n$ and for any distinct $u, v \in V(F)$, consider $|\mathcal{P}(u, v)|$. We will show that $|\mathcal{P}(x, y)| \geq |\mathcal{P}(u, v)|$. Suppose first that $\{u, v\} \subseteq B_i$, for some $2 \leq i \leq k-1$. If $uv \in E(F)$, then $|\mathcal{P}(u, v)| = 1 < f_2(n) = |\mathcal{P}(x, y)|$. So suppose $uv \notin E(F)$. In this case, $|\mathcal{P}(u, v)| \leq |B_{i-1}| + |B_{i+1}|$, as the only induced paths from u to v are paths $u w v$, where w is a neighbour of u and v . But by Lemma 2.3, definition of F and as $|B_i| \geq 2$, we have

$$|\mathcal{P}(x, y)| = \prod_{j=2}^{k-1} |B_j| \geq 2 \cdot |B_{i-1}| \cdot |B_{i+1}| \geq |B_{i-1}| + |B_{i+1}| \geq |\mathcal{P}(u, v)|.$$

So now suppose that $u \in B_i$ and $v \in B_j$ with $i < j$. In this case, any induced path between u and v contains a single vertex from each of B_i, \dots, B_j and no vertex outside of these clusters. So in particular, any induced path between u and v can be extended to an induced path from x to y , and thus again $|\mathcal{P}(x, y)| \geq |\mathcal{P}(u, v)|$. Therefore, we have $p_2(F) = |\mathcal{P}(x, y)| = f_2(n)$ for all $F \in \mathcal{F}_n$. \square

The main aim of this section is to prove the following.

Theorem 2.6. *Let G be a finite graph on $n \geq 4$ vertices and let x and y be distinct vertices of G . Suppose that G contains $p_2(n)$ induced paths between x and y . Then G is isomorphic to a graph in \mathcal{F}_n with end clusters $\{x\}$ and $\{y\}$.*

This along with Lemma 2.3 and Claim 2.5 gives the following.

Corollary 2.7. *For all $n \geq 4$, we have $p_2(n) = f_2(n)$.*

In order to prove Theorem 2.6 (counting induced paths), we will use the following claim, the proof of which is a simple exercise that will be performed below. Let U be a set. Say that \mathcal{X} is a *partition* of U if $X \neq \emptyset$ for all $X \in \mathcal{X}$, $\bigcup_{X \in \mathcal{X}} X = U$ and $X_1 \cap X_2 = \emptyset$ for all $X_1, X_2 \in \mathcal{X}$. Say that \mathcal{X} is *optimal* if $\prod_{X \in \mathcal{X}} |X| \geq \prod_{Y \in \mathcal{Y}} |Y|$ for any partition \mathcal{Y} of U .

Claim 2.8. *Let $n \geq 4$ and let U be a set of cardinality $n-2$. Let \mathcal{X} be an optimal partition of U . Then \mathcal{X} satisfies one of the following cases:*

- *If $n \equiv 2$ modulo 3, we have $|X| = 3$ for all $X \in \mathcal{X}$.*

- If $n \equiv 1$ modulo 3, we have $|Y| = 2$ for some $Y \in \mathcal{X}$ and $|X| = 3$ for all $X \in \mathcal{X} \setminus \{Y\}$;
- If $n \equiv 0$ modulo 3, we have either $|Y| = 4$ for some $Y \in \mathcal{X}$ and $|X| = 3$ for all $X \in \mathcal{X} \setminus \{Y\}$; or $|Y_1| = |Y_2| = 2$ for some $Y_1 \neq Y_2 \in \mathcal{X}$, and $|X| = 3$ for all $X \in \mathcal{X} \setminus \{Y_1, Y_2\}$.

In particular, $\prod_{X \in \mathcal{X}} |X| = f_2(n)$.

Proof. Let \mathcal{X} be an optimal partition of U . Suppose there exists some $X \in \mathcal{X}$ such that $|X| \geq 5$. Let X_1 be a subset of X with cardinality 2 and let $X_2 := X \setminus X_1$. Then as $2 \cdot (|X| - 2) > |X|$, the partition $\mathcal{X}' := \{X_1, X_2\} \cup \mathcal{X} \setminus \{X\}$ (which is still a partition of U) satisfies $\prod_{Y \in \mathcal{X}'} |Y| > \prod_{Y \in \mathcal{X}} |Y|$, a contradiction. So no member of \mathcal{X} has cardinality greater than 4.

Now, suppose there is $X \in \mathcal{X}$ with $|X| = 1$. By taking the partition \mathcal{X}' obtained from \mathcal{X} by removing X and replacing some $Y \neq X$ in \mathcal{X} with $Y \cup X$, we see that again $\prod_{Y \in \mathcal{X}'} |Y| > \prod_{Y \in \mathcal{X}} |Y|$, a contradiction.

Therefore $2 \leq |X| \leq 4$, for all $X \in \mathcal{X}$. If \mathcal{X} contains some distinct X, Y with $|X| = |Y| = 4$, then letting Z_1, Z_2, Z_3 be a partition of $X \cup Y$ such that $|Z_1| = |Z_2| = 3$ and $|Z_3| = 2$ and considering the partition $\mathcal{X}' := \{Z_1, Z_2, Z_3\} \cup \mathcal{X} \setminus \{X, Y\}$ gives $\prod_{Y \in \mathcal{X}'} |Y| > \prod_{Y \in \mathcal{X}} |Y|$, a contradiction. Similarly, replacing three sets of cardinality 2 with two of cardinality 3 or replacing a set of cardinality 4 and a set of cardinality 2 with a pair, each of size 3, gives a partition \mathcal{X}' with $\prod_{Y \in \mathcal{X}'} |Y| \geq \prod_{Y \in \mathcal{X}} |Y|$, a contradiction. The result follows. \square

To prove Theorem 2.6, it is helpful to introduce the following notation and definitions.

Definition 2.9. For a vertex $v \in V(G)$, let $N^i(v)$ be the set of vertices at distance exactly i from v . Define $N^k[v]$ to be the set of vertices within distance k of v (for example $N^3[v] = \{v\} \cup \bigcup_{i=1}^3 N^i(v)$), and define $N[v] := N^1[v] = N(v) \cup \{v\}$. Also, for a set $X \subseteq V(G)$, let $N(X) := \bigcup_{x \in X} N(x)$, and $N[X] := \bigcup_{x \in X} N[x]$. Note that $X \subseteq N[X]$, and $X \cap N(X)$ may or may not be empty. For a subgraph $H \subseteq G$, define $N(H) := N(V(H))$ and $N[H] := N[V(H)]$.

A definition similar to the following is given in Definition 3.19 to prove Theorem 1.3 (maximising induced cycles).

Definition 2.10. Let G be a graph and let x and y be distinct vertices of G . For an induced path $P := x, x_2, \dots, x_k \in G$, define $L_y(P)$ to be the number of paths in $\mathcal{P}(x, y)$ with initial segment P . Say that P is *extendible to y* if $L_y(P) > 0$. We say an induced path $P' \in \mathcal{P}(x, y)$ with initial segment P is an *extension* of P . Define $C_y(P)$ to be the set of $w \in V(G)$ such that x_1, \dots, x_k, w is extendible to y and let $c_y(P) := |C_y(P)|$. For a vertex $x_{k+1} \in N(x_k) \setminus P$, we may write $P \cup x_{k+1}$ to denote the path x_1, \dots, x_{k+1} .

For ease of notation, we may omit the subscript in the notation $L_y(P)$, $C_y(P)$ and $c_y(P)$ when it is clear which vertex plays the role of y .

Remark 2.11. Fix x_1 and x_k distinct vertices of G and some path $x_1, x_2, \dots, x_{k-1}, x_k \in \mathcal{P}(x_1, x_k)$. Note that, for $j < k$ and $w \in C_{x_k}(x_1, \dots, x_j)$, the path x_1, \dots, x_j, w is induced and hence $w \notin N(x_i)$ for any $i < j$. As $C_{x_k}(x_1, \dots, x_i) \subseteq N(x_i)$, we have $w \notin C_{x_k}(x_1, \dots, x_i)$. Therefore, for any $i < j$ the sets $C_{x_k}(x_1, \dots, x_i)$ and $C_{x_k}(x_1, \dots, x_j)$ are disjoint.

We will also use the following lemma.

Lemma 2.12. *Let G be a graph on $n \geq 4$ vertices. Let x and y be distinct vertices in $V(G)$ and let $\mathcal{P} := \mathcal{P}(x, y) \neq \emptyset$. Let $x_1 := x$ and $P := x_1 \dots, x_{k-1}, y$ be any path in \mathcal{P} . Then:*

(i) $|\mathcal{P}| \leq f_2(n)$.

(ii) *If $|\mathcal{P}| = f_2(n)$, then:*

(a) *the set $\{C(x_1, \dots, x_j) : 1 \leq j \leq k-2\}$ is an optimal partition of $V(G) \setminus \{x, y\}$;*

(b) *for all $1 \leq j \leq k-2$ and for all $u, w \in C_y(x_1, \dots, x_j)$, we have*

$$L_y(x_1, \dots, x_j, u) = L_y(x_1, \dots, x_j, w).$$

Proof. Let $v_1 := x$. If $xy \in E(G)$ then (i) follows immediately. So suppose $xy \notin E(G)$. Sequentially choose a path v_1, \dots, v_{k-1}, y in \mathcal{P} in the following way. At vertex v_j , choose v_{j+1} to be some $z \in C(v_1, \dots, v_j)$ such that

$$L(v_1, \dots, v_j, z) = \max\{L(v_1, \dots, v_j, w) : w \in C(v_1, \dots, v_j)\}.$$

Define $P_i := v_1, \dots, v_i$ and let $\mathcal{P}' \subseteq \mathcal{P}$ be the set of paths that can be obtained in this manner.

For $1 \leq i \leq k-1$, we have

$$\begin{aligned} L(P_i) &= \sum_{z \in C(P_i)} L(P_i \cup z) \\ &\leq c(P_i) \max\{L(P_i \cup z) : z \in C(P_i)\} \\ &= c(P_i) L(P_i \cup v_{i+1}). \end{aligned} \tag{2.13}$$

By repeatedly applying (2.13) we get

$$|\mathcal{P}| \leq c(P_1) \max\{L(P_1 \cup z) : z \in C(P_1)\} \leq \dots \leq L(P_{k-1}) \prod_{i=1}^{k-2} c(P_i). \tag{2.14}$$

As by construction y is a neighbour of v_{k-1} , we have $L(P_{k-1}) = 1$. Thus

$$|\mathcal{P}| \leq \prod_{i=1}^{k-2} c(P_i), \quad (2.15)$$

where $\sum_{i=1}^{k-2} c(P_i) \leq n - 2$ (as by remark 2.11, $C(P_1), \dots, C(P_{k-2})$ are disjoint subsets of $V(G) \setminus \{x, y\}$). So by Claim 2.8 we see that the maximal possible value of $\prod_{i=1}^{k-2} c(P_i)$ is $f_2(n)$, and so $|\mathcal{P}| \leq f_2(n)$ as required for (i).

When $|\mathcal{P}| = f_2(n)$ we have

$$\prod_{i=1}^{k-2} c(P_i) = f_2(n). \quad (2.16)$$

So using Claim 2.8, we have that $\{C(P_i) : 1 \leq i \leq k - 2\}$ is an optimal partition of $V(G) \setminus \{x, y\}$ (as by Remark 2.11, $C(P_1), \dots, C(P_{k-2})$ are disjoint subsets of $V(G) \setminus \{x, y\}$). In addition, we have equality in (2.15) and hence in (2.13) for each value of $1 \leq i \leq k - 1$. Therefore, for each i and for all $z, w \in C(P_i)$, we have

$$L(P_i \cup z) = L(P_i \cup w). \quad (2.17)$$

Now we will show that $\mathcal{P}' = \mathcal{P}$. The statements (ii a) and (ii b) will follow from this and the arguments of the previous paragraph (as the deductions are valid for any path in \mathcal{P}). So suppose, for a contradiction, that there exists a path $X := x_1, \dots, x_{k-1}, y$ in $\mathcal{P} \setminus \mathcal{P}'$. Choose $Y := y_1, \dots, y_{\ell-1}, y$ in \mathcal{P}' so that it coincides with X on the longest possible initial segment. Since neither X nor Y is an extension of the other, there exists t such that $x_i = y_i$ for all $i < t$, but $x_t \neq y_t$. Setting $P_i := y_1, \dots, y_i = x_1, \dots, x_i$, by (2.17) we have $L(P_i \cup x_{i+1}) = L(P_i \cup y_{i+1})$, contradicting the choice of Y (there exists some path in \mathcal{P}' that coincides with X on a longer initial segment than Y does). So $X \in \mathcal{P}'$, as required. This completes the proof of the lemma. \square

We will now complete the proof of Theorem 2.6.

Proof of Theorem 2.6. Let $\mathcal{P} := \mathcal{P}(x, y)$. By Lemma 2.12 (i) we know $|\mathcal{P}| \leq f_2(n)$ and moreover we know there exist n -vertex graphs H such that $p_2(H) = f_2(n)$ (just pick $H \in \mathcal{F}_n$). So $|\mathcal{P}| = f_2(n)$. We will show that G is in \mathcal{F}_n . First we will show that G is a braid.

Let $v_1 := x$ and let $P := v_1, \dots, v_{k-1}, y$ be a path in \mathcal{P} such that k is minimal. For $1 \leq i \leq k - 1$, define $P_i := v_1, \dots, v_i$ and $V_{i+1} := C(P_i)$ (note that $v_{i+1} \in V_{i+1}$). Therefore, $V_k = \{y\}$. Define $V_1 := \{x\}$. The sets V_i are disjoint by Remark 2.11. By choice of P , no vertex in $\bigcup_{i=1}^{k-2} V_i$ is a neighbour of y . We also have by Lemma 2.12 (ii a) that $\bigcup_{i=1}^k V_i = V(G)$. See Figure 2.3 for an illustration of the path P and how the vertices are partitioned into the sets V_i .

Theorem 2.6 will follow almost immediately from the next claim.

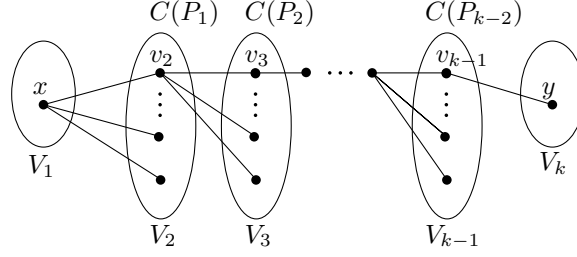


Figure 2.3: An illustration of v_1, \dots, v_k and the partition V_1, \dots, V_k . There may be additional edges within each V_i and between clusters (not depicted here). By definition, for any i and $j > i + 1$, there are no edges from v_i to a vertex in V_j . No vertex in $\bigcup_{i=1}^{k-2} V_i$ is a neighbour of y .

Claim 2.18. G is the braid $(\{x\}, V_2, \dots, V_{k-1}, \{y\})$.

Let us briefly see why the claim implies the proof of the theorem. By Lemma 2.12 (ii a), the set $\{C(P_j) : 1 \leq j \leq k - 2\}$ is an optimal partition of $V(G) \setminus \{x, y\}$. As $|V_{i+1}| = c(P_i)$, using Claim 2.8, the braid G is in \mathcal{F}_n . So it remains to prove Claim 2.18.

Proof of Claim 2.18. We prove by reverse induction on s that for all $1 \leq s \leq k - 1$:

- (S1) All edges between V_s and V_{s+1} are present.
- (S2) There are no edges between V_{s-1} and $\bigcup_{i=s+1}^k V_i$ (where $V_0 = \emptyset$).

It follows from this that $(\{x\}, V_2, \dots, V_{k-1}, \{y\})$ is a braid in G .

For the base case $s = k - 1$, we must show that $N(y) \supseteq V_{k-1}$. By definition of $C(P_{k-2})$, for every vertex $z \in V_{k-1}$, the path v_1, \dots, v_{k-2}, z is extendible to y . But as by Lemma 2.12(ii a) we have $\bigcup_{i=1}^{k-2} C(v_1, \dots, v_i) = V(G) \setminus \{x, y\}$, each such z is a neighbour of y (there is no other possible vertex through which to extend the path to y). Therefore $N(y) \supseteq V_{k-1}$. By choice of P as the shortest induced path in $\mathcal{P}(x, y)$, there are no edges between y and V_{k-2} . This completes the proof of the base case.

Now suppose that (S1) and (S2) hold for $s \leq k - 1$ and we wish to show that they hold for $s - 1 \geq 1$. First consider (S1). We must show that every edge between V_{s-1} and V_s is present. Suppose there exist some $u \in V_{s-1}$ and $v \in V_s$ such that $uv \notin E(G)$. If $s - 1 = 1$, then by definition of V_2 we have that (S1) holds, so we may assume $s - 1 \geq 2$. Consider the path $P' := v_1, \dots, v_{s-2}, u$. Our goal is to show that $L(P') < L(P_{s-1})$ (a contradiction). As $u \in C(v_1, \dots, v_{s-2})$, we have that P' is extendible to y . Let P'' be an extension of P' to y . As $\bigcup_{i=1}^{s-1} V_i \subseteq N[P_{s-2}]$, we have that $P'' \setminus P'$ is contained in $\bigcup_{i=s}^k V_i$. Using this and (S2) from our inductive hypothesis gives that $C(P') \subseteq V_s$. Pick $w \in C(P') \subseteq V_s$. As (V_s, \dots, V_k) is a braid in $G[\bigcup_{i=s}^k V_i]$, by applying Lemma 2.3 (recalling that $s \leq k - 1$) there

are $\prod_{i=s+1}^k |V_i|$ induced paths from w to y in $G[\bigcup_{i=s}^k V_i]$. Any extension of $P' \cup w$ consists of the path $P' \cup w$ and one of these induced paths. Therefore

$$L(P' \cup w) \leq \prod_{i=s+1}^k |V_i|. \quad (2.19)$$

Also, by definition of v_1, \dots, v_{k-1} , our inductive hypothesis and Lemma 2.3, we have

$$L(P_{s-1}) = \prod_{i=s}^k |V_i|. \quad (2.20)$$

Using (2.19), (2.20) and the fact that, by hypothesis, u has a non-neighbour in V_s but $N(v_{s-1}) \supseteq V_s$ (so $|C(P')| < |V_s|$), we have

$$L(P') \leq |C(P')| \cdot \prod_{i=s+1}^k |V_i| < \prod_{i=s}^k |V_i| = L(P_{s-1}),$$

contradicting Lemma 2.12 (ii b). Thus $\{uw : u \in V_{s-1}, w \in V_s\} \subseteq E(G)$, as required for (S1).

Now consider (S2). We must show that no vertex of V_{s-2} has a neighbour in $\bigcup_{i=s}^k V_i$. When $s \leq 2$ this holds vacuously, so we may assume that $s > 2$. Suppose, in order to obtain a contradiction, there exists some $j \geq s$, some $u \in V_{s-2}$ and $w \in V_j$ such that $uw \in E(G)$ (by choice of P , $w \neq y$ and so $j < k$). Let j be maximal with this property. So u has no neighbours in $\bigcup_{i=j+1}^k V_i$. Let $P_u := P_{s-3} \cup u$. Our aim is to show that $L(P_u) < L(P_{s-2})$ (a contradiction).

Consider $L(P_u)$. Every vertex in $\bigcup_{i=1}^{s-2} V_i$ is contained in $N[P_{s-3}]$, so for any extension P' of P_u to y , we have that $P' \setminus P_u$ is contained in $\bigcup_{i=s-1}^k V_i$. By our inductive hypothesis and (S1) for $s-1$ (just proved above), we have that (V_{s-1}, \dots, V_k) is a braid in $G' := G[\bigcup_{i=s-1}^k V_i]$.

We now see that the only way to extend $P_u \cup w$ to y is via an induced path contained entirely in (V_j, \dots, V_k) . As (V_{s-1}, \dots, V_k) is a braid in G' , there are no edges between $\bigcup_{i=s-1}^{j-1} V_i$ and $\bigcup_{i=j+1}^k V_i$, w is adjacent to every vertex in V_{j+1} and the only neighbours of w in G' lie in $V_{j-1} \cup V_j \cup V_{j+1}$. Therefore $L(P_u \cup w \cup x) = 0$ for any $x \in V_{j-1} \cup V_j$. So $C(P_u \cup w) \subseteq V_{j+1}$ and, as required, any extension of $P_u \cup w$ consists of P_u and an induced path containing w in (V_j, \dots, V_k) .

For any neighbour w of u in V_j , by Lemma 2.3 (as $j < k$) there are at most $\prod_{i=j+1}^k |V_i|$ induced paths from w to y contained in (V_j, \dots, V_k) . Therefore

$$L(P_u \cup w) \leq \prod_{i=j+1}^k |V_i|. \quad (2.21)$$

Similarly to the deduction of (2.20), using the definition of v_1, \dots, v_{k-1} , our inductive hypothesis, the knowledge that every edge between V_{s-1} and V_s is present and Lemma 2.3 we have

$$L(P_{s-2}) = \prod_{i=s-1}^k |V_i|. \quad (2.22)$$

By Lemma 2.12 (ii b), we have $L(P_u \cup z_1) = L(P_u \cup z_2)$, for any $z_1, z_2 \in C(P_u)$. So by (2.21), for all $z \in C(P_u)$:

$$L(P_u \cup z) = \prod_{i=j+1}^k |V_i|. \quad (2.23)$$

As j is maximal such that u has a neighbour in V_j , we have $C(P_u) \subseteq \bigcup_{i=s-1}^j V_i$. So by this and (2.23):

$$L(P_u) = \sum_{z \in C(P_u)} L(z) \leq \left(\sum_{i=s-1}^j |C(P_u) \cap V_i| \right) \cdot \prod_{i=j+1}^k |V_i|. \quad (2.24)$$

By Lemma 2.12 (ii b), we have $L(P_u) = L(P_{s-2})$, which is equal to $\prod_{i=s-1}^k |V_i|$ by (2.22). Therefore by (2.24) we have $\sum_{i=s-1}^j |C(P_u) \cap V_i| = \prod_{i=s-1}^j |V_i|$. As by Lemma 2.12(i a) and Claim 2.8 we have $|V_i| > 1$ for all $s-1 \leq i \leq j$ (recall that by assumption $s > 2$, so $s-1 \geq 2$), this is only possible when $j = s$, $C(P_u) \supseteq V_{s-1} \cup V_s$ and $|V_{s-1}| = |V_s| = 2$.

However, if $V_s \subseteq N(u)$ then for any $w \in V_{s-1}$ there is no extension P'_u of $P_u \cup w$ to y (such a P'_u must satisfy $(P'_u \setminus P_u) \cap V_s \neq \emptyset$, as no vertex of $\bigcup_{i=s+1}^k V_i$ has a neighbour outside of $\bigcup_{i=s}^k V_i$, hence P'_u is not an induced path). Therefore (S2) holds for $s-1$ as required and Claim 2.18 now follows by induction. \square

This completes the proof of Theorem 2.6. \square

2.3 Proof of Theorem 1.3

Fix a large constant n_0 and let G_{\max} be a graph on $n \geq n_0$ vertices that contains $m(n)$ induced cycles. In what follows we will take n_0 (and thus n) to be sufficiently large when required and we will make no attempts to optimise the constants in our arguments. We will show that the graph G_{\max} is isomorphic to H_n . As it turns out, Theorem 1.6 (the stability result) will follow almost immediately from the arguments required for the proof of Theorem 1.3. Therefore, in this section several lemmas are proved in more generality than is needed for the proof of Theorem 1.3; they will be used in their more general form in the next section.

Given a graph H , let $f(H)$ denote the number of induced cycles in H and for a vertex $v \in H$, let $f_v(H)$ denote the number of induced cycles in H containing v . As described in Section 2.1, there are two types of induced cycles in H_n . The first type contains a vertex

from every cluster. The second type are induced C_4 s contained within a pair of adjacent clusters or a triple of consecutive clusters. There are $O(n)$ induced cycles of the second type. By counting induced cycles of the first type we have:

$$f(G_{\max}) = m(n) \geq f(H_n) \geq \begin{cases} 3^{n/3} & \text{if } n \equiv 0 \text{ modulo } 3 \\ 4 \cdot 3^{(n-4)/3} & \text{if } n \equiv 1 \text{ modulo } 3 \\ 2 \cdot 3^{(n-2)/3} & \text{if } n \equiv 2 \text{ modulo } 3. \end{cases} \quad (3.1)$$

Any G_{\max} is connected (if it were disconnected we could add edges between two components to increase the number of induced cycles).

2.3.1 Outline of the Proof

Before proceeding with the details of the proof, let us give a brief outline of what is to come.

Heuristically, we wish to show that the local structure of a vertex looks like the local structure around most vertices in H_n (Figure 2.4 depicts this ‘local structure’).

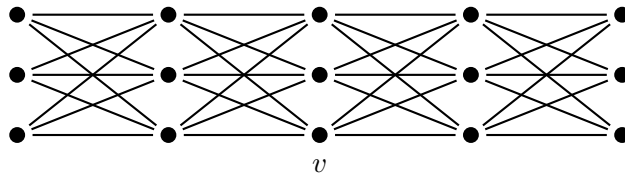


Figure 2.4: The graph induced by $N^2[v]$ for most vertices v in H_n .

Given this, at the heart of the proof lies a notion of a ‘typical’ vertex in G_{\max} . The precise definition of ‘typical’ is fairly complicated (see Definition 3.7). But given an appropriate definition of ‘typical’, we will show that:

1. A ‘typical’ vertex has the required local structure. (See Lemma 3.9.)
2. All but a bounded number of vertices in G_{\max} are ‘typical’. (See Lemma 3.22.)

Once we know these two pieces of information about G_{\max} we will see that most of the vertices of G_{\max} are contained in braids with clusters of size 3. In fact, all but a bounded number of vertices are in such braids, and so (in Lemma 3.45) we will see that this implies there is some braid B containing $\Omega(n)$ vertices such that $(1 - o(1))f(H_n)$ induced cycles ‘pass through’ B . Given this, we can then express the total number of cycles of G_{\max} in the number of ways we can take a path through B and the number of paths between the endpoints x and y of this path outside of B . This is the part of the proof (Lemma 3.46) where we will apply Theorem 2.6 (to a ‘reduced graph’) to get that G_{\max} is a cyclic braid.

The result will follow by ‘cleaning’ the graph to show that the cyclic braid is empty and by determining the precise cluster sizes.

2.3.2 Preliminary Lemmas

We begin by proving some simple lemmas about some structural properties of G_{\max} that will be used throughout the proof.

Lemma 3.2. *Let F be an n -vertex graph. For $v \in V(F)$, we have*

$$f_v(F) \leq \binom{d(v)}{2} 3^{(n-d(v)-1)/3}.$$

Proof. Each induced cycle containing v contains exactly two neighbours of v . Fix a pair of vertices $u, w \in N(v)$. If $uw \in E(F)$, then there is at most one induced cycle in F containing $\{u, v, w\}$. If $uw \notin E(F)$, then by Corollary 2.7 and (2.4) there are at most $3^{(n-d(v)-1)/3}$ induced paths between u and w in $(F \setminus N[v]) \cup \{u, w\}$. Thus there can be at most $3^{(n-d(v)-1)/3}$ induced cycles in F containing $\{v, u, w\}$. As there are $\binom{d(v)}{2}$ distinct pairs of neighbours of v , we have,

$$f_v(F) \leq \binom{d(v)}{2} 3^{(n-d(v)-1)/3}$$

as required. \square

The next lemma tells us that any vertex in G_{\max} is contained in at least a constant proportion of $f(G_{\max})$ induced cycles. In what follows, the *length* of a path or cycle will be the number of vertices it contains.

Lemma 3.3. *Let $0 < c \leq 1$ and let $\gamma = 0.11$. Let F be an n -vertex graph with $f(F) \geq c \cdot m(n)$. Then:*

(i) $(1 - o(1))f(F)$ induced cycles in F have length at least γn .

(ii) F contains a vertex v such that $f_v(F) \geq \frac{c}{10}m(n)$.

(iii) Every vertex $w \in V(G_{\max})$ satisfies $f_w(G_{\max}) \geq \frac{c}{20}m(n)$.

Proof. F contains at most $\sum_{i=1}^{\lfloor \gamma n \rfloor} \binom{n}{i}$ induced cycles of length at most γn (for any $W \subseteq V(F)$, there exists at most one induced cycle C such that $V(C) = W$). Using Stirling's approximation, we get

$$\sum_{i=1}^{\lfloor \gamma n \rfloor} \binom{n}{i} \leq \gamma n \cdot \binom{n}{\gamma n} \leq (1 + o(1)) \frac{\sqrt{\gamma n}}{\sqrt{2\pi(1-\gamma)}} \left[\frac{1}{\gamma^\gamma (1-\gamma)^{1-\gamma}} \right]^n.$$

As

$$\frac{1}{\gamma^\gamma (1-\gamma)^{1-\gamma}} < 3^{1/3},$$

we get

$$\sum_{i=1}^{\lfloor \gamma n \rfloor} \binom{n}{i} = o\left(3^{n/3}\right).$$

By (3.1), $f(F) = \Omega(3^{n/3})$, so $(1 - o(1))f(F)$ induced cycles in F have length at least γn , as required for (i).

Provided n_0 is sufficiently large, we have for all $n > n_0$,

$$\sum_{i=1}^{\lfloor \gamma n \rfloor} \binom{n}{i} < \frac{c}{1000} \cdot 3^{n/3} < \frac{c}{100} m(n),$$

where the second inequality follows from (3.1). It follows from this that at least $99c/100 \cdot m(n)$ induced cycles of F contain at least γn vertices. So

$$\sum_{v \in V(F)} f_v(F) \geq \frac{99\gamma c}{100} n \cdot m(n),$$

and by the pigeonhole principle, there exists a vertex v such that

$$f_v(F) \geq \frac{99\gamma c}{100} m(n) \geq \frac{c}{10} m(n),$$

proving (ii).

Now suppose that there exists some vertex $w \in V(G_{\max})$ with $f_w(G_{\max}) < \frac{c}{20} m(n)$. Consider the graph G' obtained from G_{\max} by removing the vertex w and duplicating the vertex v . We have that

$$f(G') \geq f(G_{\max}) - f_w(G_{\max}) + (f_v(G_{\max}) - f_w(G_{\max})) > f(G_{\max}),$$

a contradiction. This proves (iii). □

Now let v be the vertex given by Lemma 3.3(ii) and consider the graph obtained from G_{\max} by duplicating v . We have

$$m(n+1) \geq \left(1 + \frac{1}{10}\right) m(n). \tag{3.4}$$

We now use this to show that the graph G_{\max} has maximum degree bounded by a constant.

Lemma 3.5. *Let F be a graph with $f(F) = m(n)$, Then $\Delta(F) < 30$.*

Proof. Let v be a vertex of maximal degree. Given v , split the induced cycles in F into those contained in $F \setminus \{v\}$, and those containing v . Using Lemma 3.2 we have

$$m(n) \leq m(n-1) + \binom{d(v)}{2} 3^{(n-d(v)-1)/3}.$$

Using (3.4) to bound $m(n-1)$ gives,

$$m(n) \leq m(n) \left(1 + \frac{1}{10}\right)^{-1} + \binom{d(v)}{2} 3^{(n-d(v)-1)/3}.$$

This expression rearranges to give

$$m(n) \leq 11 \binom{d(v)}{2} 3^{(n-d(v)-1)/3}.$$

For $d(v) \geq 30$, this implies $m(n) < 3^{(n-6)/3}$, a contradiction (as $m(n) \geq f(H_n) > 3^{(n-6)/3}$). \square

Combining Lemma 3.2 with Lemma 3.5 gives that, for any $v \in V(G_{\max})$, we have $f_v(G_{\max}) \leq \binom{30}{2} 3^{(n-3)/3}$. By Lemma 3.3 (i), we know $(1 - o(1))m(n)$ induced cycles in G_{\max} have length at least $0.11n$. Thus

$$(1 - o(1))m(n) \leq \frac{1}{0.11n} \sum_{v \in V(G)} f_v(G_{\max}) = O(3^{n/3}).$$

This along with (3.1) implies

$$m(n) = \Theta(3^{n/3}). \quad (3.6)$$

2.3.3 The Local Structure Around a ‘Typical’ Vertex

The next stage of our proof involves showing that the local structure around a ‘typical’ vertex is the same as the local structure around a vertex in H_n (see Figure 2.4).

Let us now discuss heuristically how we will proceed. Imagine choosing an induced path $P = u_1, u_2, \dots, u_k$ sequentially in H_n (where n is large). Start from any vertex u_1 , let u_2 be in an adjacent cluster and let u_3 be in a different cluster to u_1 (so we begin moving around the cycle of clusters). After this start, at all but one or two steps no matter what choices we have made so far for P , we would have exactly three choices of which vertex to add next to P . Similarly, once we can no longer add a vertex to the path (because every choice would result in a non-induced path) every vertex of H_n is in $N[P]$.

We wish to exploit these properties to show that vertices in G_{\max} must behave, in this fashion, ‘like’ the vertices in H_n . So we will define a notion of ‘typicality’ where a vertex z in F is ‘typical’ if it satisfies the following. Starting at some $v \in V(F)$ that is ‘not too close’ in the graph to z , then for any way of picking a ‘long’ induced path P in F :

- (1) whenever we are at a vertex in $N^5[z]$ we have exactly 3 choices for the next vertex to add to the path;
- (2) when we can no longer choose a new vertex to add to P , every vertex of $N^5[z]$ is in $N[P]$.

It is natural to express these properties in the form of a game, where two players are picking an induced path through F . The player *Builder* is trying to build an induced path that satisfies these properties, but the player *Adversary* is trying to stop this. Adversary will get to choose u_{i+1} when u_i is ‘close’ to z , and Builder will get to choose otherwise. A vertex will be typical when there exists a winning strategy for Builder.

We remark that it is important that we begin at a vertex that is ‘not too close’ to the vertices we wish to prove are typical. When we start building a path in H_n , after placing u_1 we usually (if it is surrounded by clusters of size 3) have 6 choices initially for u_2 , and then 5 choices for u_3 , which would violate (1).

Note also that this notion of a vertex z being typical depends only on what happens locally around z . This will be important in the next subsection, when we wish to show that G_{\max} only contains a constant number of atypical vertices. We will be able to play a game determining the typicality of many vertices simultaneously and will be penalised for every atypical vertex. More details of this will be given in the next subsection, this is just a flavour of what is coming.

We are now ready to formally define the game that determines when a vertex is typical.

Definition 3.7. Let F be a finite graph, let $u_1 \in V(F)$ and let $z \in V(F) \setminus N^5[u_1]$. We define the z -*typical-game* on (F, u_1) as follows. There are two players, *Adversary* and *Builder*. The players choose a sequence of vertices u_2, \dots, u_k under the following set of rules. For $i \geq 1$, at vertex u_i :

- If $u_i \in N^5[z]$, then Adversary is the *active player*, otherwise Builder is.
- The active player chooses a neighbour u_{i+1} of u_i that is not in $N[\{u_1, \dots, u_{i-1}\}]$. We call this a *new neighbour* of u_i (with respect to u_1, \dots, u_{i-1}).
- The game terminates when a vertex u_j is chosen such that there is no possible choice for u_{j+1} .
- We call u_j *good* if there are exactly 3 choices for u_{j+1} ; we call u_j *bad* otherwise.
- Adversary wins if either:
 - for some j , the vertex u_j is in $N^5[z]$ and is bad; or
 - upon termination of the game at vertex u_k , there exists a vertex in $N^5[z]$ that is not in $N[\{u_1, \dots, u_k\}]$.
- Builder wins if both:
 - at every vertex $u_j \in N^5[z]$ chosen, there are exactly 3 choices for u_{j+1} ; and

– upon termination of the game at vertex u_k , we have $N^5[z] \subseteq N[\{u_1, \dots, u_k\}]$.

A vertex $z \in V(F) \setminus N^5[u_1]$ is *typical in* (F, u_1) if there exists a winning strategy for Builder in the z -typical-game on (F, u_1) . A vertex is *atypical in* (F, u_1) otherwise (so vertices in $N^5[u_1]$ are atypical). When it is clear which vertex has been chosen to play the role of u_1 , we simply say that z is (a)typical. Note that the set of vertices $\{u_1, \dots, u_k\}$ chosen during the game induces a path in F . Also, if we play this game on H_n starting at any vertex, most of the chosen vertices are good.

The next proposition will be helpful in determining the local structure around a typical vertex.

Proposition 3.8. *Let F be a graph and let $u_1 \in V(F)$. Let z be typical in (F, u_1) . Let u_1, \dots, u_j be the first j vertices of any path obtained by playing the z -typical game on (F, u_1) where Builder plays with a winning strategy σ . Let $x_1 \in V(F)$ be such that $x_1 \in N^3[z]$, and such that u_j is the first vertex chosen in $N^2(x_1)$. Let Y be the set of new neighbours of u_j . Then $N(u_j) \cap N(x_1) = Y$.*

Proof. We play as Adversary in the z -typical-game on (F, u_1) . We assume that Builder uses strategy σ , and deduce information about the structure of F from the results of our choices of vertices as Adversary (we know that, *whatever we do*, we cannot win so the choices we make have certain consequences). As z is typical, $d(u_1, z) > 5$ and hence $d(u_1, x) > 2$, so no vertex in u_1, \dots, u_j is a neighbour of x_1 .

As $N^2[x_1] \subseteq N^5[z]$, we have that u_j has 3 new neighbours $Y := \{y_1, y_2, y_3\}$. As $u_j \in N^2(x_1)$, for some i we have $y_i \in N(u_j) \cap N(x_1)$. Without loss of generality suppose $y_1 \in N(u_j) \cap N(x_1)$. Choosing u_{j+1} to be y_1 gives that the vertex y_1 has 3 new neighbours, one of which is x_1 , so let the set of new neighbours of y_1 be $X := \{x_1, x_2, x_3\}$. As u_j is the first vertex chosen in $N^2(x_1)$ and y_1 the first in $N(x_1)$, we have $N(u_j) \cap N(x_1) \subseteq Y$.

If we choose $u_{j+1} := y_1$ and then $u_{j+2} := x_1$, we have that x_1 has 3 new neighbours (as $x_1 \in N^5[z]$ and z is typical). Thus x_1 has at most 5 neighbours not in $N[\{u_1, \dots, u_j\}]$ (as x_1 could be adjacent to x_2 or x_3).

We now prove that $N(x_1) \cap N(u_j) = Y$. Suppose otherwise, so without loss of generality we have $y_3 \notin N(x_1)$. Now we describe the set of choices we make for the remainder of the game (recall that Builder always plays by strategy σ). Choose $u_{j+1} := y_3$. Now consider a later step in the game, but before any vertex in $N[x_1]$ has been chosen. We play as follows:

- (i) At a vertex in $N^5[z] \setminus N^2[x_1]$, we play arbitrarily.
- (ii) At a vertex $u_i \in N^2(x_1)$, where $i \geq j + 1$:

- If, for some r , the vertex x_1 has r neighbours not in $N[\{u_1, \dots, u_{i-1}\}]$ (by the discussion about $N(x_1)$, $r \leq 5$) and u_i is adjacent to $t \geq 1$ of these neighbours:
 - (C1) If $r - t < 3$, choose $u_{i+1} \in N(u_i) \cap N(x_1)$ and $u_{i+2} := x_1$. As x_1 now has at most 2 new neighbours we reach a contradiction.
 - (C2) If $r - t \geq 3$, then $t \leq 2$ and u_i has a new neighbour s that is not adjacent to x_1 (otherwise u_i is bad, which contradicts the fact that Builder has a winning strategy). Choose $u_{i+1} := s$. Observe that the vertex x_1 is not in $N[\{u_1, \dots, u_{i+1}\}]$.
- If u_i has no new neighbour that is a neighbour of x_1 (i.e. $t = 0$):
 - (C3) Pick u_{i+1} arbitrarily.

Once we have chosen x_1 , we play arbitrarily. As the first time we choose a neighbour of x_1 we then choose x_1 , we have described a strategy for any step of the game in which we can play.

We now analyse the results of making these choices. As z is typical, Builder has a winning strategy and so at some point a vertex in $N(x_1)$ must be chosen (so that $x_1 \in N[\{u_1, \dots, u_k\}]$). As the game starts outside of $N[x_1]$, there is a first time we chose a vertex of $N(x_1)$. This vertex must be chosen as a new neighbour of a vertex in $N^2(x_1)$. When this happens, we cannot be in case (C3) (otherwise we couldn't move to $N(x_1)$), we are also not in case (C2) (as there we do not move to a neighbour of x_1), so we must be in case (C1) and hence reach a contradiction. Thus $N(u_j) \cap N(x_1) = Y$, as required. \square

The next lemma shows a typical vertex has the required local structure (see Figure 2.5). The proof is fairly technical. To aid the reader we have broken it up into small simple claims.

Lemma 3.9. *Let F be a graph and let u_1 be a vertex in F . Suppose that $z \in V(F) \setminus N^5[u_1]$ is typical in (F, u_1) . Then there exist disjoint sets of vertices B_1, B_2, B_3, B_4 such that $z \in B_3$, $|B_i| = 3$ for all $i \in \{1, \dots, 4\}$, and (B_1, B_2, B_3, B_4) is a braid in F .*

Proof. As in the proof of Proposition 3.8, we play as Adversary in the z -typical-game on (F, u_1) and assume that Builder uses the winning strategy σ . Let u_1, u_2, \dots be the path chosen during the game. Suppose that u_k is the first vertex chosen such that $u_k \in N^5(z)$ (as z is typical, at some point such a vertex will be chosen). Let A_1 be the set of new neighbours of u_k . Sequentially for $i \in \{1, \dots, 4\}$, we (as Adversary) arbitrarily choose $u_{k+i} \in N^{5-i}(z) \cap N(u_{k+i-1})$ and let A_{i+1} be the set of new neighbours of u_{k+i} with respect

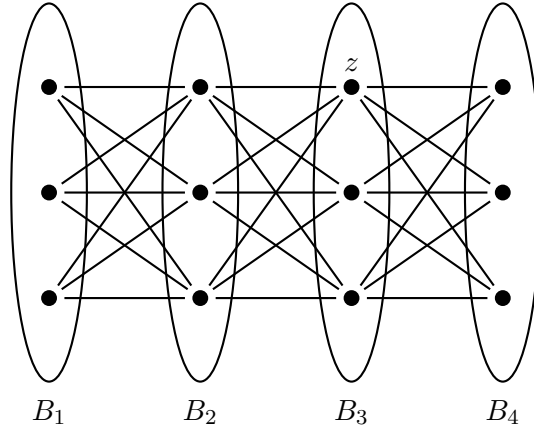


Figure 2.5: The local structure around a typical vertex z . Note that Lemma 3.9 does not determine whether there are edges within the sets B_1, B_2, B_3 and B_4 .

to u_1, \dots, u_{k+i-1} . Such choices for u_{k+1}, \dots, u_{k+4} exist as, for each $j \in \{0, \dots, 4\}$, u_{k+j} is the first vertex chosen in $N^{5-j}(z)$. For $i \in \{1, \dots, 4\}$ define $r_i := u_{k+i}$. As $d(r_i, z) = 5 - i$, we have that $r_3 \notin N(z)$. As z is typical, $|A_i| = 3$ for $i \in \{1, \dots, 4\}$ and r_4 also has 3 new neighbours, the set A_5 containing $r_5 := z$. Choosing $u_{k+5} := z$ gives that z also has 3 new neighbours (as z is typical). Call this set of new neighbours A_6 . See Figure 2.6 for an illustration of what we have deduced so far about the structure of F around z .

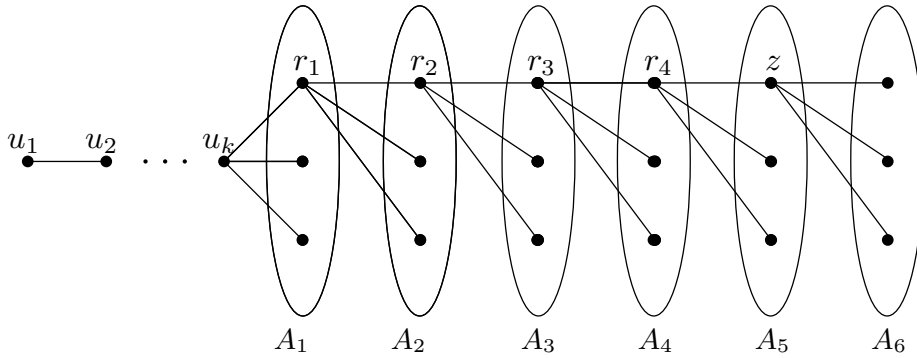


Figure 2.6: There may or may not be edges within the sets A_1, \dots, A_6 .

In what follows, we will analyse several different ways that the game could play out based on what we (as Adversary) choose to do. So for $j \geq k$ we will use the phrase *restarting from u_j* to indicate that we are returning to the point in the game with our initial set-up and fixed choices of u_1, \dots, u_j , in order to play a different set of moves from here on and analyse the consequences of those moves. In practice we will have $j \geq k$, so the vertices u_1, \dots, u_k will remain fixed for the remainder of the proof.

We will show that:

- (i) For all $i \in \{4, 5\}$ and $u \in A_i$, we have $A_{i-1} \cup A_{i+1} \subseteq N(u) \subseteq A_{i-1} \cup A_i \cup A_{i+1}$.

(ii) There are no edges between A_3 and A_6 .

Conditions (i) and (ii) imply that (A_3, A_4, A_5, A_6) is a braid in F . As $z \in A_5$, this gives the statement of the lemma.

Claim 3.10. For $i \in \{2, 3\}$, $N(r_i) \cap N(r_{i+2}) = A_{i+1}$.

Proof. By definition, r_i is the first vertex chosen in $N^2(r_{i+2})$ and $r_{i+2} \in N^3[z]$. By applying Proposition 3.8 with $u_j := r_i$ and $x_1 := r_{i+2}$ (so $Y = A_{i+1}$), we have that $N(r_i) \cap N(r_{i+2}) = A_{i+1}$. \square

See Figure 2.7 for an illustration of what Claim 3.10 tells us about the structure of F close to z .

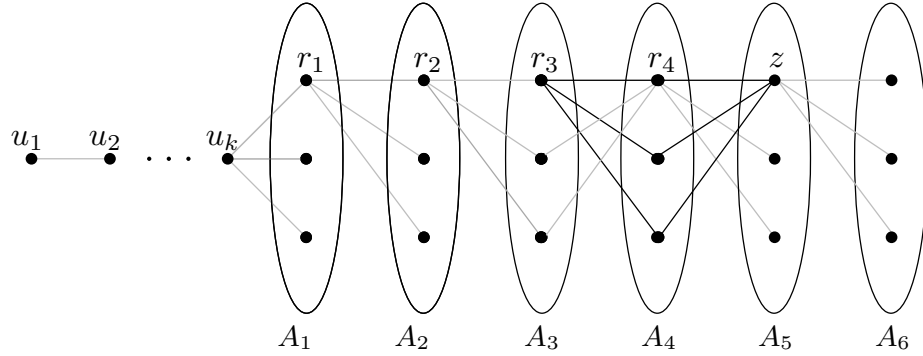


Figure 2.7: There may or may not be edges within the sets A_1, \dots, A_5 . The edges illustrating $N(r_3) \cap N(z)$ are darker to highlight the purpose of Claim 3.11 in the case $i = 3$.

Claim 3.11. For $i \in \{4, 5\}$,

$$A_{i-1} \cup A_{i+1} \subseteq N(r_i) \subseteq A_{i-1} \cup A_i \cup A_{i+1}.$$

Proof. By definition of A_{i+1} and Claim 3.10, we have $N(r_i) \supseteq A_{i-1} \cup A_{i+1}$. As $r_{i-2} \in N^2(r_i)$ is the first vertex chosen in $N^2[r_i]$, no vertex in $N[\{u_1, \dots, u_k, r_1, \dots, r_{i-3}\}]$ is in $N(r_i)$. Any neighbour of r_i not in $N[\{u_1, \dots, u_k, r_1, \dots, r_{i-1}\}] = N[\{u_1, \dots, u_k, r_1, \dots, r_{i-3}\}] \cup A_{i-1} \cup A_i$ would have been a new neighbour of r_i and hence in A_{i+1} . \square

Claim 3.12. $N(r_4) \cap N(w) = A_5$ for all $w \in A_6$.

Proof. Pick $w \in A_6$. We wish to apply Proposition 3.8, so first we need to show that r_4 is the first vertex chosen in $N^2(w)$. Restarting from r_4 , let y be the first vertex chosen in $N^2(w)$. Observe that as $w \in N(z)$, we have $N^2(w) \subseteq N^3[z]$. So $y \in \{r_2, r_3, r_4\}$. We will show that $y = r_4$.

In any of these cases, as y is the first vertex chosen in $N^2(w)$ and $w \in N^3[z]$, we may apply Proposition 3.8. Suppose that $y = r_2$. Then, by Proposition 3.8, $N(r_2) \cap N(w)$ is the set of new neighbours of r_2 , and in particular w is adjacent to r_3 . This is a contradiction as $w \notin A_4$. Now suppose that $y = r_3$. Again, Proposition 3.8 implies that w is adjacent to r_4 , a contradiction. Therefore $y = r_4$ and using Proposition 3.8 we have $N(r_4) \cap N(w) = A_5$ as required. \square

See Figure 2.8 for an illustration of what we now know about the local structure around z .

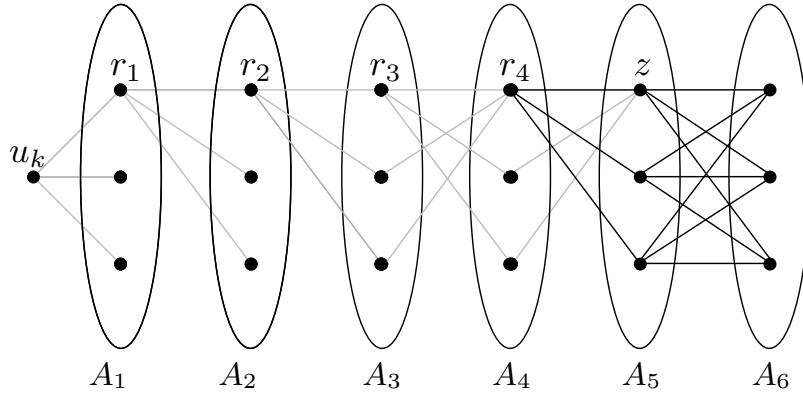


Figure 2.8: There may or may not be edges within the sets A_1, \dots, A_6 . The darker edges illustrate Claim 3.12.

Claim 3.13. For $u \in A_5 \setminus \{z\}$, we have $N(u) \subseteq \bigcup_{j=2}^6 A_j$.

Proof. This claim follows similarly to the proof of Claim 3.11. Restart from u_k and choose $u_{k+1} := r_1$. As u_k is the first vertex chosen in $N^5[z]$ and $u \in N^2[z]$, no vertex of $N[\{u_1, \dots, u_k\}]$ is in $N(u)$. Any vertex of $N(u)$ not in $N[\{u_1, \dots, u_k, r_1, \dots, r_4\}]$ will be a new neighbour of u with respect to the path $u_1, \dots, u_k, r_1, \dots, r_4, u$. As z is typical, u has 3 new neighbours. By definition, no vertex of A_6 is in $N[\{u_1, \dots, u_k, r_1, \dots, r_4\}]$ but by Claim 3.12 we have $A_6 \subseteq N(u)$, so the set of new neighbours of u is precisely the set A_6 and the claim follows. \square

Claim 3.14. For $v \in A_4 \setminus \{r_4\}$ and $u \in A_5 \setminus \{z\}$ we have $uv \in E(F)$.

Proof. Restarting from $u_{k+3} = r_3$ choose $u_{k+4} := v$. Observe that as z is a new neighbour of v , there are at most two vertices of A_6 in $N(v)$ (as z is typical v has exactly 3 new neighbours). Let $w \in A_6 \setminus N(v)$. Choose $u_{k+5} := z$ next (this is possible as $N(r_3) \cap N(z) = A_4$). If z is adjacent to u , pick $u_{k+6} := u$. Otherwise, pick $u_{k+6} := w$ and $u_{k+7} := u$. In both cases, using Claim 3.13 we have that u has at most one new neighbour (the vertex in $A_5 \setminus \{z, u\}$). This contradicts z being typical. Thus we have $uv \in E(F)$ as required. \square

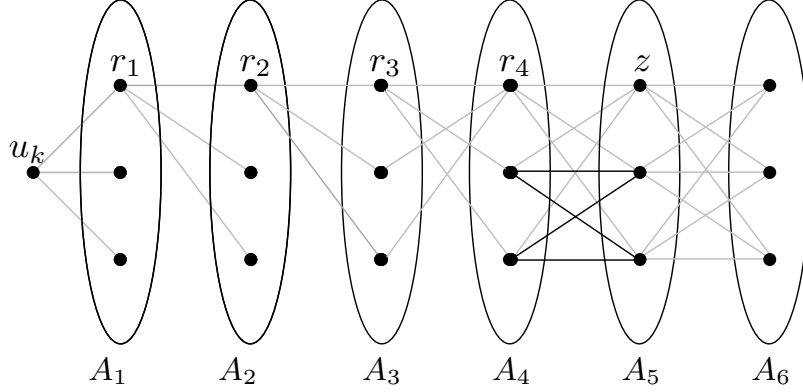


Figure 2.9: There may or may not be edges within the sets A_1, \dots, A_6 . The darker edges illustrate the edges given by Claim 3.14.

Claim 3.14 is illustrated by Figure 2.9. We now prove a pair of claims, analogous to Claims 3.13 and 3.14, whose proofs follow identically to those of Claims 3.13 and 3.14. We remark that in order for the following proofs to work, we need the previous two claims. This is why we did not prove them simultaneously with the previous two claims.

Claim 3.15. For $u \in A_4 \setminus \{r_4\}$, we have $N(u) \subseteq \bigcup_{j=1}^5 A_j$.

Proof. This proof is essentially identical to the proof of Claim 3.15, but is included for completeness. Restart from u_k . As u_k is the first vertex chosen in $N^5[z]$ and $u \in N^2[r_4]$, no vertex of $N[\{u_1, \dots, u_{k-1}\}]$ is in $N(u)$. Any vertex of $N(u)$ not contained within $N[\{u_1, \dots, u_k, r_1, \dots, r_3\}]$ will be a new neighbour of u with respect to $u_1, \dots, u_k, r_1, \dots, r_3$. As z is typical, u has 3 new neighbours. As no vertex of A_5 is in $N[\{u_1, \dots, u_k, r_1, \dots, r_4\}]$ but (using Claim 3.14) we have $A_5 \subseteq N(u)$, the set of new neighbours of u is precisely the set A_5 and the claim follows. \square

Claim 3.16. For $v \in A_3 \setminus \{r_3\}$ and $u \in A_4 \setminus \{r_4\}$ we have $uv \in E(F)$.

Proof. This proof is essentially identical to the proof of Claim 3.14, but is included for completeness. Restarting from $u_{k+2} = r_2$, choose $u_{k+3} := v$. As above, there are at most two vertices of A_5 in $N(v)$. Let $w \in A_5 \setminus N(v)$. Choose $u_{k+4} := r_4$ next. If r_4 is adjacent to u , pick $u_{k+5} := u$. Otherwise, pick $u_{k+5} := w$ and $u_{k+6} := u$ (this is possible by Claim 3.14). In both cases, using Claim 3.15 we have that u has at most one new neighbour (the vertex in $A_4 \setminus \{r_4, u\}$). This contradicts z being typical. Thus we have $uv \in E(F)$ as required. \square

See Figure 2.10 for an illustration of Claim 3.16 and what we now know about the local structure around z .

Claim 3.17. For $i \in \{4, 5\}$, no vertex of A_{i-1} has a neighbour in A_{i+1} .

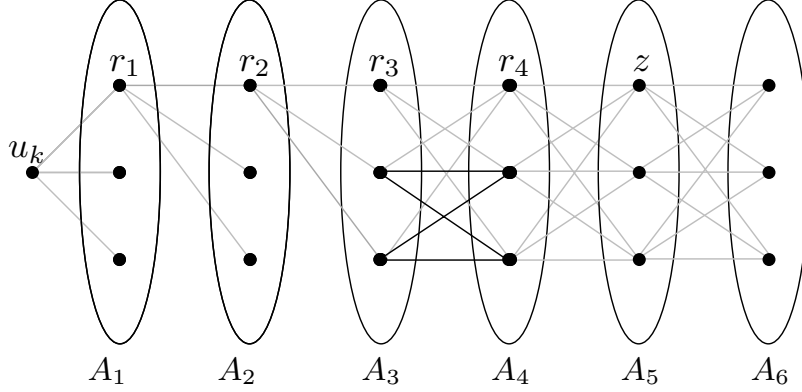


Figure 2.10: There may or may not be edges within the sets A_1, \dots, A_6 . The darker edges illustrate the edges given by Claim 3.16.

Proof. By definition of A_{i-1} and A_{i+1} , r_{i-1} has no neighbour in A_{i+1} . Pick $u \in A_{i-1} \setminus \{r_{i-1}\}$. We show that u has no neighbour in A_{i+1} . Restarting from $u_{k+i-2} = r_{i-2}$, choose $u_{k+3} = u$. By Claims 3.10, 3.14 and 3.16 we have $A_i \subseteq N(u)$ and by definition $(A_i \cup A_{i+1}) \cap N[\{u_1, \dots, u_k, r_1, \dots, r_{i-2}\}] = \emptyset$. Therefore the set S of new neighbours of u contains any vertex of $A_i \cup A_{i+1}$ that is a neighbour of u . As z is typical, u has exactly 3 new neighbours and as $A_i \subseteq S$ and $|A_i| = 3$, no vertex of A_{i+1} can be a neighbour of u . \square

When $i = 4$, replacing $i + 1$ with $i + 2$ in this proof shows that no vertex of A_3 has a neighbour in A_6 , as required for (ii). To prove (i), it remains to show that for $i \in \{4, 5\}$, for all $u \in A_i \setminus \{r_i\}$, u has no neighbours outside of $A_{i-1} \cup A_i \cup A_{i+1}$. For r_4 and $r_5 = z$ this is given by Claim 3.10.

Claim 3.18. For $i \in \{4, 5\}$ and $u \in A_i \setminus \{r_i\}$, $N(u) \subseteq A_{i-1} \cup A_i \cup A_{i+1}$.

Proof. By Claims 3.13 and 3.15, $N(u) \subseteq \bigcup_{j=i-3}^{i+1} A_j$. To complete the claim, we must show that u has no neighbour in $A_{i-3} \cup A_{i-2}$. Suppose $a \in (A_{i-3} \cup A_{i-2}) \cap N(u)$. If $N(u) \cap A_{i-3} \neq \emptyset$, then fix $a \in N(u) \cap A_{i-3}$. Otherwise, fix $a \in A_{i-2}$. Supposing that $a \in A_j$, restart from u_k , next choose r_1, \dots, r_{j-1} and then choose a . Let S be the set of new neighbours of a . Next pick u . As $A_{i-1} \cup A_i \subseteq N(u)$, the set of new neighbours of u contains the set $(A_{i-1} \cup A_{i+1}) \setminus S$. But as $u \in S$ and $|S| = 3$ (as $a \in N^5[z]$), this means that u has at least 4 new neighbours, a contradiction. \square

This completes the proof of (i) and hence the proof of Lemma 3.9. \square

2.3.4 Bounding the Number of ‘Atypical’ Vertices

We will now show that, for any $v \in V(G_{\max})$, all but a bounded number of vertices are typical in (G_{\max}, v) . Roughly speaking, we will show that if the graph contains too many atypical vertices, then the number of induced cycles containing v is small, contradicting Lemma 3.3. It will be helpful to introduce a definition, similar to Definition 2.10 to aid us when thinking about induced paths that can be extended to induced cycles containing v .

Definition 3.19. Let F be a graph and let $v \in V(F)$. Define $\mathcal{P}(v)$ to be the set of paths v, u_2, \dots, u_k in G such that v, u_2, \dots, u_{k-1} is an induced path and $u_k \notin N[\{u_2, \dots, u_{k-2}\}]$ (so u_k could be a neighbour of v). Say that a path $P := v, x_2, \dots, x_k \in \mathcal{P}(v)$ is *extendible* if there exists some $w \in V(G)$ such that $P \cup w := v, x_2, \dots, x_k, w \in \mathcal{P}(v)$. Define $C(P)$ to be the set of $w \in V(F)$ such that $P \cup w \in \mathcal{P}(v)$. Let $L(P)$ be the number of maximal paths in $\mathcal{P}(v)$ whose initial segment is P . In particular, when P is not extendible, $L(P) = 1$.

Remark 3.20. Let $v \in V(F)$ and let $P := v, u_2, \dots, u_k \in \mathcal{P}(v)$. If $u_k \in N(v)$, then $C(P) = \emptyset$. If $u_k \notin N^2[v]$, then $C(P)$ is precisely the set of neighbours of u_k that are not in $N[\{v, u_2, \dots, u_{k-1}\}]$. Recalling Definition 3.7, this is the set of new neighbours of u_k with respect to v, u_2, \dots, u_{k-1} . If $u_k \in N^2[v]$, then $C(P)$ is not necessarily the set of new neighbours of u_k . It may also include neighbours of v .

The next lemma is an easy consequence of Definition 3.19.

Lemma 3.21. *Let F be a graph and let $v \in V(F)$. We have*

$$f_v(F) \leq \frac{1}{2}L(v).$$

Proof. For every induced cycle of F starting at v , $\mathcal{P}(v)$ contains two non-extendible paths. So $L(v) \geq 2f_v(F)$, and the result follows. \square

Note that a non-extendible path of $\mathcal{P}(v)$ may not correspond to an induced cycle containing v . A path could be non-extendible because each neighbour of its endpoint is in the neighbourhood of some vertex other than v , earlier on the path.

Let F be any n -vertex graph such that $L(v) \geq c \cdot 3^{n/3}$ for some $c > 0$, $v \in V(F)$ and $\Delta(F) \leq \Delta$, for some constant Δ . (G_{\max} satisfies these conditions by Lemma 3.3, Lemma 3.5 and Lemma 3.21.) Our next aim is to prove that, for any vertex $v \in V(F)$, only a bounded number of vertices are atypical in (F, v) . Using this fact with Lemma 3.9 implies that the majority of the structure of G_{\max} is close to the structure of H_n . The remainder of the proof consists of ‘cleaning’ G_{\max} to show that it is in fact isomorphic to H_n .

We first outline how the proof will proceed before giving the details. We assume (in order to get a contradiction) that there is a large set $A \subseteq V(F) \setminus N^{10}[v]$ of vertices atypical in

(F, v) , such that for each $a, a' \in A$ we have $N^5[a] \cap N^5[a'] = \emptyset$ (for any set of atypical vertices there exists a subset of constant proportion with this property as $\Delta(F)$ is bounded). We will sequentially choose an induced path $P := u_1, u_2, \dots, u_{k-1}, u_k$ in $\mathcal{P}(v)$, where $u_1 := v$.

For each $a \in A$, there exists a winning strategy τ_a for Adversary in the a -typical game on (F, v) . This means that whatever vertices outside of $N^5[a]$ have been chosen so far, for every $a \in A$ Adversary is able to make choices inside $N^5[a]$ to ensure that either:

- (i) we choose a bad vertex $u_i \in N^5[a]$, or
- (ii) there is some vertex in $N^5[a] \setminus N[P]$.

By hypothesis, at the start $L(u_1)$ is bounded below by $c \cdot 3^{n/3}$, for some constant c . As we sequentially pick the path we keep track of a lower bound on $L(u_1, \dots, u_i)$ at each stage i . If a good vertex $u_i \in N^5[a]$ is chosen, then as $A \cap N^{10}[v] = \emptyset$, by Remark 3.20 $C(u_1, \dots, u_i)$ is the number of new neighbours of u_i and we have $|C(u_1, \dots, u_i)| = 3$. So we can pick u_{i+1} such that $L(u_1, \dots, u_{i+1}) \geq \frac{1}{3}L(u_1, \dots, u_i)$. If we are at a vertex u_i that is not in $N^5[a]$, for any $a \in A$, we choose the member of $C(u_1, \dots, u_i)$ that maximises $L(u_1, \dots, u_{i+1})$.

When u_i is in $N^5[a]$, for some $a \in A$, we play the winning strategy τ_a to move towards the outcomes (i) or (ii) (unless there is some member u_i of $C(u_1, \dots, u_i)$ such that $L(u_1, \dots, u_{i+1})$ is sufficiently large). That is, we will simultaneously play the strategy τ_a for all $a \in A$. This is possible as for each $a, a' \in A$ we have $N^5[a] \cap N^5[a'] = \emptyset$, so the strategies are played on disjoint sets of vertices. The outcomes (i) and (ii) mean that $\mathcal{P}(v)$ is ‘unbalanced’ in some way. For each $a \in A$, the strategy τ_a will pick some vertex u_{i+1} such that $L(u_1, \dots, u_{i+1}) > \frac{(1+\epsilon)}{3}L(u_1, \dots, u_i)$, for some $\epsilon > 0$. So, for each $a \in A$, we will *win an extra factor* in our lower bound. If $|A|$ is too large, when the process ends we find that $L(u_1, \dots, u_k) > 1$, a contradiction.

Lemma 3.22. *Fix $c > 0$. Let F be an n -vertex graph with $\Delta := \Delta(F) > 1$ and let $\epsilon := 2^{-\Delta^{100}}$. Let $v \in V(F)$. Let $A \subseteq V(F) \setminus N^{10}[v]$ be such that every $a \in A$ is atypical for (F, v) and for all $a, a' \in A$, we have $N^5[a] \cap N^5[a'] = \emptyset$. If $L(v) \geq c \cdot 3^{n/3}$, then $|A| < M$, where M is the smallest integer such that $c \cdot 3^{1/3}(1+\epsilon)^M > \Delta$.*

Proof. Suppose, in order to obtain a contradiction, that $|A| \geq M$. For each $a \in A$, Adversary has a winning strategy τ_a played on vertices of $N^5[a]$ in the a -typical game in (F, v) . As for all $a, a' \in A$, we have $N^5[a] \cap N^5[a'] = \emptyset$, these strategies are played on disjoint sets of vertices.

Let $u_1 := v$. We will sequentially choose a path $u_1, \dots, u_k \in \mathcal{P}(v)$. For ease of notation, define $P_i := u_1, \dots, u_i$. At each stage $i \geq 2$, we choose a vertex u_i and define A_i (the subset of A that we still care about). Once we have won an extra factor for a , we no longer

care about playing a special strategy in $N^5[a]$. We also define $t_i(a)$ for each $a \in A$. The $t_i(a)$ will be a set of constants that keep track of the factors we gain in the lower bound on $L(u_1, \dots, u_i)$. For each $a \in A$ define $t_2(a) := 1$. Also for $i \geq 2$ define

$$t_i := \frac{c \cdot 3^{1/3}}{\Delta} \prod_{a \in A} t_i(a) \quad \text{and} \quad q_i := t_i 3^{\frac{n-m_i-1}{3}}, \quad (3.23)$$

where m_i is the number of vertices in $N[\{u_2, \dots, u_{i-1}\}] \setminus \{v\}$.

So $n - m_i - 1$ vertices of $V(F) \setminus \{v\}$ are not in $N[\{u_2, \dots, u_{i-1}\}]$. These vertices are ones that could still be added to the path at a later stage. By definition, $q_2 = \frac{c \cdot 3^{1/3}}{\Delta} 3^{(n-1)/3}$. Throughout the process we maintain the property that, for each i ,

$$L(u_1, \dots, u_i) \geq q_i. \quad (3.24)$$

That is, q_i will be our lower bound on $L(u_1, \dots, u_i)$.

We now describe an algorithm that determines our choice of vertices. For $r \geq 1$, let $\epsilon_r = 2^{2(r-1)}\epsilon$.

Vertex Choice Algorithm

Pick $u_2 \in N(v)$ such that $L(P_2)$ is maximised, and define $A_2 := A$. Suppose the most recently chosen vertex is u_i .

If P_i is extendible; there are two cases:

Case 1: $u_i \in N^5[a]$ for some $a \in A_i$.

Suppose it is the r -th time we have chosen a vertex y such that $y \in N^5[a]$. We have two subcases:

Subcase 1: u_i is good.

In this case, as $u_i \notin N^5[v]$, by Remark 3.20 $C(P_i)$ is the set of new neighbours of u_i and contains exactly 3 vertices y_1, y_2, y_3 .

- (i) If there exists j such that $L(P_i \cup y_j) \geq \frac{1}{3}(1 + \epsilon_r)q_i$ (where q_i is defined in (3.23)), then choose $u_{i+1} := y_j$.
 - Set $t_{i+1}(a) := (1 + \epsilon_r)t_i(a)$ and $t_{i+1}(y) := t_i(y)$, for all $y \in A \setminus \{a\}$.
 - Set $A_{i+1} := A_i \setminus \{a\}$.

(ii) Else, every y_j satisfies $L(P_i \cup y_j) > \frac{1}{3}(1 - 2\epsilon_r)q_i$ (as q_i satisfies (3.24)). In this case, choose u_{i+1} according to strategy τ_a .

- Set $t_{i+1}(a) := (1 - 2\epsilon_r)t_i(a)$ and $t_{i+1}(y) := t_i(y)$, for all $y \in A \setminus \{a\}$.
- Set $A_{i+1} := A_i$.

Subcase 2: u_i is bad.

In this case, (again by Remark 3.20) $C(P_i)$ is the set of new neighbours of u_i , so $|C(P_i)| \neq 3$. Suppose $C(P_i) = \{y_1, \dots, y_\ell\}$ for some $\ell \neq 3$. Pick j such that $L(P_i \cup y_j)$ is maximised and set $u_{i+1} := y_j$.

- Set $t_{i+1}(a) = (1 + \epsilon_r)t_i(a)$ and $t_{i+1}(y) := t_i(y)$, for all $y \in A \setminus \{a\}$.
- Set $A_{i+1} := A_i \setminus \{a\}$.

Case 2: $u_i \notin N^5[a]$ for any $a \in A_i$:

Then $C(P_i) = \{y_1, \dots, y_\ell\}$ for some $\ell \geq 1$. Pick j such that $L(P_i \cup y_j)$ is maximised and set $u_{i+1} = y_j$.

- Set $t_{i+1}(y) := t_i(y)$, for all $y \in A$.
- Set $A_{i+1} := A_i$.

The process terminates when P_i is not extendible.

We now analyse the consequences of choosing vertices in this manner.

Claim 3.25. *For each vertex u_i chosen during the Vertex Choice Algorithm, we have $L(P_i) \geq q_i$.*

Proof. We first show that when $u_i \notin N(v)$, we have $m_{i+1} - m_i = |C(P_i)|$. This fact will then be used throughout the rest of the claim.

By definition,

$$m_{i+1} - m_i = |N[\{u_2, \dots, u_i\}] \setminus \{v\}| - |N[\{u_2, \dots, u_{i-1}\}] \setminus \{v\}|.$$

So in particular (as $u_i \notin N(v)$), $m_{i+1} - m_i$ is the number of vertices in $S := N(u_i) \setminus N[\{u_2, \dots, u_{i-1}\}]$. Recalling Definition 3.19, we see that $S = C(P_i)$ and so $m_{i+1} - m_i = |C(P_i)|$ as required.

We now show that $L(P_i) \geq q_i$ for each u_i , as required for the statement of the claim. We argue by induction on i ; the case $i = 2$ holds as $u_2 \in N(v)$ is chosen to maximise $L(P_2)$. Suppose $L(P_i) \geq q_i = t_i 3^{(n-m_i-1)/3}$. Now for the inductive step we consider each case of

the algorithm separately, and prove that the statement holds there. As in each case P_i is extendible, $u_i \notin N(v)$ and $m_{i+1} - m_i = |C(P_i)|$.

In Subcase 1(i), as u_i is good and $u_i \notin N^5[v]$, by Remark 3.20 we have $m_{i+1} - m_i = |C(P_i)| = 3$ and so:

$$L(P_{i+1}) \geq \frac{1}{3}(1 + \epsilon_r)t_i 3^{\frac{n-m_i-1}{3}} = t_{i+1} 3^{\frac{n-m_{i+1}-1}{3}} = q_{i+1}.$$

In Subcase 1(ii), again $m_{i+1} - m_i = 3$ and we have:

$$L(P_{i+1}) > \frac{1}{3}(1 - 2\epsilon_r)t_i 3^{\frac{n-m_i-1}{3}} = t_{i+1} 3^{\frac{n-m_{i+1}-1}{3}} = q_{i+1}.$$

In Subcase 2, recall that $C(P_i) = \{y_1, \dots, y_\ell\}$ for $\ell \neq 3$ (by Remark 3.20 as $u_i \notin N^5[v]$) and we pick u_{i+1} to be the y_j which maximises $L(P_i \cup y_j)$. Thus we have:

$$L(P_{i+1}) \geq \frac{t_i}{\ell} 3^{\frac{n-m_i-1}{3}} = t_i \frac{3^{\ell/3}}{\ell} 3^{\frac{n-m_i-\ell-1}{3}}.$$

The value of $\frac{3^{\ell/3}}{\ell}$ is minimised for $\ell \neq 3$ at $\ell = 2$. We have

$$L(P_{i+1}) \geq t_i \frac{3^{2/3}}{2} 3^{\frac{n-m_i-\ell-1}{3}} \geq (1 + \epsilon_r)t_i 3^{\frac{n-m_i-\ell-1}{3}} = t_{i+1} 3^{\frac{n-m_{i+1}-1}{3}} = q_{i+1}.$$

In Case 2, recall that $C(P_i) = \{y_1, \dots, y_\ell\}$ and we pick u_{i+1} to be the y_j which maximises $L(P_i \cup y_j)$. So we have

$$L(P_{i+1}) \geq \frac{t_i}{\ell} 3^{\frac{n-m_i-1}{3}} = t_i \frac{3^{\ell/3}}{\ell} 3^{\frac{n-m_i-\ell-1}{3}} \geq t_i 3^{\frac{n-m_i-\ell-1}{3}} = t_{i+1} 3^{\frac{n-m_{i+1}-1}{3}} = q_{i+1},$$

where the last inequality is strict unless $\ell = 3$. □

We are now equipped to analyse what remains once the algorithm terminates. Focussing on one $a \in A$, we may consider all choices outside $N^5[a]$ as made by Builder in the a -typical game on (F, v) . If we don't go through Subcase 1(i) or Subcase 2, then all choices inside $N^5[a]$ are made according to τ_a , which is a winning strategy for Adversary in this game. In this case Adversary wins and either we chose a bad vertex (which was included in Subcase 2) or there is some $w \in N^5[a]$ that is not contained in $N[P_k]$ upon termination of the algorithm.

So to summarise, at least one of the following outcomes occurs upon termination of the algorithm.

- (O1) During the algorithm, at a vertex in $N^5[a]$, we either went through Subcase 1(i) or we chose a bad vertex via Subcase 2.

(O2) There is some vertex $w \in N^5[a]$ that is not contained in $N[P_k]$ upon termination of the algorithm.

First observe that for all $s \leq \Delta^6 + 1$,

$$(1 - 2\epsilon_1)(1 - 2\epsilon_2) \dots (1 - 2\epsilon_s)(1 + \epsilon_{s+1}) > 1 + \epsilon. \quad (3.26)$$

Our choice of ϵ ensures that each factor on the left hand side is greater than zero.

Suppose $a \notin A_k$. Then there exists some j such that $a \in A_j$ but $a \notin A_{j+1}$. Thus at the j th stage of the algorithm $u_j \in N^5[a]$ and we either went through Subcase 1(i) or we chose a bad vertex via Subcase 2. Let $\lambda := |P_j \cap N^5[a]|$ be the number of vertices in $N^5[a]$ chosen up to the j th stage. From the algorithm we see

$$t_k(a) = t_{j+1}(a) = (1 - 2\epsilon_1)(1 - 2\epsilon_2) \dots (1 - 2\epsilon_{\lambda-1})(1 + \epsilon_\lambda). \quad (3.27)$$

As $|N^5[a]| \leq \Delta^6$, we have $\lambda \leq \Delta^6$, and so by (3.26) we have for $a \notin A_k$

$$t_k(a) > 1 + \epsilon. \quad (3.28)$$

Now suppose $a \in A_k$. Let $\lambda := |P_{k-1} \cap N^5[a]|$. By following the algorithm we see that whenever we are at a vertex $u \in N^5[a]$, we do not pass through Subcase 1(i) or Subcase 2, as this would imply $a \notin A_k$. Thus $t_k(a) = \prod_{i=1}^{\lambda} (1 - 2\epsilon_i)$. By definition of ϵ , for all $s \leq \Delta^6$ we have

$$3^{1/3} > 1 + 2^{2s}\epsilon,$$

so by (3.26) and the observation that $\lambda \leq \Delta^5$,

$$3^{1/3} \cdot t_k(a) > 1 + \epsilon. \quad (3.29)$$

For each $a \in A_k$, the set $N[P_k]$ does not contain $N^5[a]$, as we either achieve outcome (O1) or (O2) for a , and if we achieved (O1), then a would not be in A_k . So at termination we have $n - m_k - 1 \geq |A_k|$ and so by the definition (3.23) of q_k we have:

$$q_k \geq t_k 3^{\frac{|A_k|}{3}}. \quad (3.30)$$

By (3.28) and (3.29), we have:

$$\prod_{a \in A} t_k(a) \geq (1 + \epsilon)^{|A \setminus A_k|} \prod_{a \in A_k} t_k(a) \geq (1 + \epsilon)^{|A|} 3^{-\frac{|A_k|}{3}}.$$

So by substituting for t_k in (3.30) and recalling the assumption that $|A| \geq M$ we have:

$$q_k \geq \frac{c \cdot 3^{1/3}}{\Delta} 3^{\frac{|A_k|}{3}} \prod_{a \in A} t_k(a) \geq \frac{c \cdot 3^{1/3}}{\Delta} (1 + \epsilon)^{|A|} \geq \frac{c \cdot 3^{1/3}}{\Delta} (1 + \epsilon)^M > 1,$$

by definition of M . This contradicts Claim 3.25, as P_k is not extendible and thus $L(P_k) = 1$. Thus $|A| < M$, concluding the proof of Lemma 3.22. \square

Corollary 3.31. *Let F be an n -vertex graph with $f(F) \geq 2c \cdot m(n)$, for some $c > 0$, and $\Delta := \Delta(F) > 1$. There exists some $v \in V(F)$ and a constant $C = C(\Delta, c)$ such that at most C vertices are atypical in (F, v) .*

Proof. Let A be the set of vertices of $F \setminus N^{10}[v]$ that are atypical in (F, v) . Let U be the largest subset of A such that for all $a, a' \in U$, we have $N^5[a] \cap N^5[a'] = \emptyset$. As $|N^5[x]| \leq \Delta^6$ for every $x \in V(F)$, we have

$$|U| \geq \frac{|A|}{\Delta^6}. \quad (3.32)$$

We wish to apply Lemma 3.22 to F , v and U .

As by hypothesis $f(F) \geq 2c \cdot m(n)$, by Lemma 3.3 there exists $v \in V(F)$ such that $f_v(F) \geq \frac{c}{10} \cdot m(n)$ for all $v \in F$. Combining this with Lemma 3.21 gives that for some constant $\beta > 0$,

$$L(v) \geq 2f_v(F) \geq \frac{c}{5}m(n) \geq \beta \cdot 3^{n/3},$$

where the last inequality follows from (3.6).

Applying Lemma 3.22 shows that $|U| < M$, where M is the smallest integer such that $\beta \cdot 3^{1/3}(1 + 2^{-\Delta^{100}})^M > \Delta$. Combining this with (3.32) and using Lemma 3.5 to show that $|N^{10}[v]| = O(1)$ gives the required result. □

2.3.5 Concluding the Proof

Our next aim is to show that G_{\max} is a cyclic braid. Let $B = (B_1, \dots, B_k)$ be a braid in a graph G . If $|B_i| = 3$ for all i , we call B a *3-braid*. If there exists some $A \subseteq V(G)$ such that either (B_1, \dots, B_k, A) or (A, B_1, \dots, B_k) is a 3-braid in G , then we say that B is *extendible*. The following proposition will be used throughout the section.

Proposition 3.33. *Let $B = (B_1, \dots, B_k)$ and $A = (A_1, \dots, A_\ell)$ be braids in a graph F , where $k, \ell \geq 4$. Let $v \in V(F)$ and $i, j \in \{1, 2\}$. If $u \in B_i \cap A_j$ is typical in (F, v) , then $B_i = A_j$.*

Proof. By Lemma 3.9, as u is typical, $|N(u)| \leq 8$ and there exist vertices u_2, u_3 such that $|N(u) \cap N(u_2) \cap N(u_3)| = 6$. As $B_{i+1} \subseteq N(u)$ and $|N(u)| \leq 8$, we have $N(u_2) \cap B_{i+1} \neq \emptyset$. As B is a braid, this implies that $u_2 \in B_i \cup B_{i+1} \cup B_{i+2}$. However, no vertex in $B_{i+1} \cup B_{i+2}$ shares 6 neighbours with u , so $u_2 \in B_i$. Similarly, $u_3 \in B_i$ and so $B_i = \{u, u_2, u_3\}$. Analogously, $A_j = \{u, u_2, u_3\}$ and so $B_i = A_j$, as required. □

We now show that if a 3-braid with 4 clusters has an end cluster consisting of typical vertices, then it is extendible.

Lemma 3.34. *Let $B = (B_1, \dots, B_4)$ be a 3-braid in a graph F and let $v \in V(F)$. Suppose that every vertex in B_1 is typical in (F, v) . Then there exists a set C disjoint from B such that (C, B_1, B_2, B_3) is a 3-braid in F .*

Proof. Let $u \in B_1$. As u is typical in (F, v) , by Lemma 3.9, there exists a 3-braid $A = (A_1, A_2, A_3, A_4)$ in F with $u \in A_3$. By applying Proposition 3.33 to B and A , we have $B_1 = A_3$.

Let $x \neq y \in B_2$. As $B_2 \subseteq N(u)$ and $B_2 \cap A_3 = \emptyset$, $\{x, y\} \subseteq A_2 \cup A_4$. As y shares at least 6 neighbours with x and A is a braid, x and y are in the same cluster. So either $B_2 = A_2$ and we set $C := A_4$, or $B_2 = A_4$ and we set $C := A_2$. Hence every edge between C and B_1 is present.

As B is a braid, no vertex in B_1 has a neighbour in $B \setminus (B_1 \cup B_2)$. In particular $N[B_1] \cap (B_3 \cup B_4) = \emptyset$. As $C \subseteq N[B_1]$, this gives that $C \cap (B_3 \cup B_4) = \emptyset$ and hence C is disjoint from B (by definition, C is disjoint from B_1 and B_2). As both (C, B_1, B_2) and B are braids, to show (C, B_1, B_2, B_3) is a braid it remains to show that there are no edges between C and B_3 . However, as B is a braid, $N(B_3) \subseteq B_2 \cup B_3 \cup B_4$ and so $N(B_3)$ is disjoint from C . Therefore (C, B_1, B_2, B_3) is a braid, as required. \square

Definition 3.35. For a braid $B = (B_1, \dots, B_k)$, for $i \in \{3, \dots, k-2\}$ we say that B_i is an *inner cluster* of B and for $i \in \{1, 2, k-1, k\}$ we say that B_i is an *outer cluster* of B .

Recall from Definition 2.1 that B_1 and B_k are also called end clusters and B_2, \dots, B_{k-1} are also called central clusters. So the set of central clusters contains the set of inner clusters. We now define a particular set of braids which we will show covers all but $O(1)$ vertices of our extremal graph.

Definition 3.36. For a graph F and $v \in V(F)$, define $\mathcal{B}(F, v)$ to be a set of braids obtained via the following procedure.

- Begin with $\mathcal{B} = \emptyset$.
- Let Y be the set of vertices in inner clusters of all the braids added so far to \mathcal{B} . If there exists some x in $V(F) \setminus (N^5[v] \cup \bigcup_{B \in \mathcal{B}} B)$ that is typical in (F, v) , let B_x be a 12-vertex 3 braid containing x in a central cluster (given by Lemma 3.9). Let $B(x)$ be a non-extendible 3-braid in $F \setminus Y$ containing B_x (the existence of $B(x)$ will be proved directly below). Say that x is the *root* of $B(x)$. Add $B(x)$ to \mathcal{B} .

Claim 3.37. *Such a $B(x)$ exists.*

Proof. It suffices to show that $B_x \cap Y = \emptyset$. Let $B_x = (B_1, B_2, B_3, B_4)$, where $x \in B_3$. Let $A = (A_1, \dots, A_k)$ be any braid in $\mathcal{B} \setminus \{B_x\}$. We will first show that no vertex of $B_2 \cup B_3 \cup B_4$ is in a central cluster of A .

By definition, $x \notin A$. Let u be a vertex in a central cluster of A . As A is a braid in F , u has no neighbours in $F \setminus A$ and hence is not in $N(x)$. Therefore no vertex of $B_2 \cup B_4$ is in a central cluster of A . Now suppose, for a contradiction, that $y \in B_3 \setminus \{x\}$ is in a central cluster A_i of A . Given B_x , we have $|N(y) \cap N(x)| \geq 6$. As A is a braid, the only vertices that share at least 6 neighbours with y are in A_i . So $x \in A_i$, a contradiction. So no vertex of $B_2 \cup B_3 \cup B_4$ is in a central cluster of A , and hence no vertex of $B_2 \cup B_3 \cup B_4$ is in an inner cluster of A .

It remains to show that no vertex of B_1 is in an inner cluster of A . Suppose, for a contradiction, that $u \in B_1$ is in an inner cluster of A . Let $w \in B_2 \subseteq N(u)$. As A is a braid, $N(u)$ is contained within the central clusters of A and so w is in a central cluster of A , a contradiction as $w \in B_2$. \square

It follows easily from the definition of \mathcal{B} that braids in \mathcal{B} only intersect on outer clusters.

Proposition 3.38. *Let F be a graph and $v \in V(F)$. Then distinct braids in $\mathcal{B} := \mathcal{B}(F, v)$ intersect only on outer clusters.*

Proof. Let $B := (B_1, \dots, B_k)$ and $A := (A_1, \dots, A_\ell)$ be braids in \mathcal{B} . Suppose, without loss of generality, that A was added to \mathcal{B} before B . We will show that A and B intersect only on outer clusters.

By definition of \mathcal{B} , no vertex of B is in an inner cluster of A . So we must show that no vertex in an outer cluster of A is in an inner cluster of B . Suppose, in order to obtain a contradiction, that $u \in A_i \cap B_j$, where $j \in [k] \setminus \{1, 2, k-1, k\}$ and, without loss of generality, $i \in \{1, 2\}$. For this to be possible, we have $k \geq 5$ (else B has no inner clusters).

As B is a braid and u is in an inner cluster of B , we have $N^2[u] \subseteq B$. If $\ell \geq 5$ then $N^2[u] \cap A$ contains a vertex x of an inner cluster of A . But then $x \in B$, a contradiction. So $\ell = 4$ and if $i = 1$, then $A_1 \cup A_2 \cup A_3 \subseteq B$, whereas if $i = 2$, then $A \subseteq B$.

We show that, for either value of i , A could have been extended when it was added to \mathcal{B} . As $A_{i+1} \subseteq N(u)$ and B is a braid, we have $A_{i+1} \subseteq B_{j-1} \cup B_j \cup B_{j+1}$. As A_{i+1} is a central cluster of A , for $w, w' \in A_{i+1}$, we have $|N(w) \cap N(w')| \geq 6$. Therefore w and w' must be in the same cluster of B . As $A_i \cap A_{i+1} = \emptyset$ and $B_j \cap A_i \neq \emptyset$, without loss of generality $A_{i+1} = B_{j+1}$.

Let $y \in A_{i+1}$. We have $A_i \cup A_{i+2} \subseteq N(y)$ and so, as B is a braid, $A_i \cup A_{i+2} = B_j \cup B_{j+2}$. Let $w \in A_i$ and $w' \in A_{i+2}$. As A is a braid we have $|N(w) \cap N(w')| = 3$. So B_j cannot

intersect both A_i and A_{i+2} (as vertices in B_j share at least 6 common neighbours). As $B_j \cap A_i \neq \emptyset$, this means that $B_j = A_i$ and $B_{j+2} = A_{i+2}$.

So if $i = 1$, then $A_1 = B_j, A_2 = B_{j+1}$ and $A_3 = B_{j+2}$. If $k \geq j + 3$, then $A_4 = B_{j+3}$. Otherwise $k = j + 2$ and, as B is a braid and $A_3 = B_{j+2}$, we have $A_4 \cap B = \emptyset$. In particular, this implies that A_4 has no neighbours in B_{j-1} . If $i = 2$, then A is the braid $(B_{j-1}, B_j, B_{j+1}, B_{j+2})$. As $k \geq 5$, when A was added to \mathcal{B} it could have been extended with the addition of B_{j-i} (this cluster is disjoint from A and cannot intersect inner clusters of braids added before A to \mathcal{B} as otherwise it would not be in B). This contradicts the construction of \mathcal{B} . So A and B intersect only on outer clusters. \square

Our next aim is to show that there are only $O(1)$ braids in \mathcal{B} . We will require the following simple deduction for the proof.

Proposition 3.39. *Let F be a graph and let $v \in V(F)$. Let $B_1, B_2, B_3 \subseteq V(F)$ be disjoint sets of cardinality 3 such that for all $u \in B_2$, we have $B_1 \cup B_3 \subseteq N(u)$. If some $z \in B_2$ is typical then $N(B_2) \subseteq B_1 \cup B_2 \cup B_3$.*

Proof. As z is typical, by Lemma 3.9 there exists a 3-braid $C = (C_1, C_2, C_3, C_4)$ in F such that $z \in C_3$. Let $w \in B_2 \setminus \{z\}$. We have $|N(w) \cap N(z)| \geq 6$. As $N(z) \subseteq C_2 \cup C_3 \cup C_4$, there must be 6 neighbours of w in $C_2 \cup C_3 \cup C_4$. As C is a braid, this is only possible if $w \in C_3$. Hence $B_2 = C_3$. Let $y \in B_2 = C_3$. As C is a braid, y has exactly 6 neighbours outside of B_2 . So as $N(y) \supseteq B_1 \cup B_3$, we have $N(y) \setminus B_2 = B_1 \cup B_3$ as required. \square

We now use this to determine some information about the braids in \mathcal{B} . As in the proof of Lemma 3.9, we have broken the next argument into smaller claims to help the reader.

Lemma 3.40. *Let F be an n -vertex graph such that $f(F) = m(n)$ and let $v \in V(F)$. Let $B = (B_1, \dots, B_k)$ be a braid in $\mathcal{B} := \mathcal{B}(F, v)$. Then either \mathcal{B} contains a braid with $|V(F)| - 3$ vertices, or some vertex in $N[B_1 \cup B_k]$ is atypical in (F, v) .*

Proof. Let $B = (B_1, \dots, B_k) \in \mathcal{B}$. By construction of \mathcal{B} , either:

- (a) B intersects another braid $A = (A_1, \dots, A_\ell)$ in \mathcal{B} ; or
- (b) B is disjoint from all other braids in \mathcal{B} and is non-extendible in F .

If some vertex in $N[B_1 \cup B_k]$ is atypical, we are done. So throughout we assume that every vertex in $N[B_1 \cup B_k]$ is typical. By construction of \mathcal{B} , we have that $k, \ell \geq 4$. First suppose (a) occurs. As, by Proposition 3.38 distinct braids in \mathcal{B} intersect only on outer clusters, without loss of generality we are in one of the following cases.

(A1) A central cluster of A intersects a central cluster of B .

(A2) No central cluster of A intersects a central cluster of B , but a central cluster of A intersects an end cluster of B .

(A3) A and B intersect on end clusters but no central cluster of A intersects B .

We will resolve each case separately.

Claim 3.41. *No central cluster of A intersects a central cluster of B .*

Proof. Without loss of generality, let $u \in A_2 \cap B_2$. As $u \in N[B_1]$, by assumption it is typical. By applying Proposition 3.33 to A and B , we have $A_2 = B_2$. As A is a braid, $N(u) \subseteq A_1 \cup A_2 \cup A_3$. As $B_3 \subseteq N(u)$, this implies that $B_3 \subseteq A$. This is a contradiction unless $k = 4$ as, by Proposition 3.38, A and B intersect only on outer clusters. An analogous argument shows that $A_3 \subseteq B$ and $\ell = 4$.

So we have $B_2 \cup B_3 \subseteq A$ and $A_2 \cup A_3 \subseteq B$. By construction of \mathcal{B} , the root of A is in $A_2 \cup A_3$ and the root of B is in $B_2 \cup B_3$. We obtain a contradiction as, without loss of generality, A was added first to \mathcal{B} but the root of B (which is contained in $B_2 \cup B_3$) was chosen to be disjoint from A . \square

Therefore (A1) cannot happen. Now we consider (A2). Without loss of generality there exists a typical vertex $u \in B_2 \cap A_1$ (if u is atypical we are done).

Claim 3.42. *If there exists a typical vertex $u \in B_2 \cap A_1$, then \mathcal{B} contains a braid on $|V(F)| - 3$ vertices.*

Proof. Let $u \in B_2 \cap A_1$ be typical. For $C \in \mathcal{B}$, let Y_C be the set of vertices in inner clusters of the braids added to \mathcal{B} before C .

Applying Proposition 3.33 to B and A gives that $B_2 = A_1$. By Claim 3.41, $A_2 \cap B_3 = \emptyset$. As A_2 is disjoint from $A_1 = B_2$, $A_2 \subseteq N(u) \setminus (B_2 \cup B_3)$ and so $B_1 = A_2$. As A is a braid, no vertex in A_1 has a neighbour in $A \setminus (A_1 \cup A_2)$. As $B_3 \subseteq N(B_2) = N(A_1)$, B_3 is disjoint from $A \setminus (A_1 \cup A_2)$. However, as $A_1 \cup A_2 = B_1 \cup B_2$ and B is a braid, B_3 is disjoint from A . Similarly, A_3 is disjoint from B .

Suppose there are no edges from B_3 to A_ℓ or from A_3 to B_k . We will show that, in this case, the braid in $\{A, B\}$ which was added first to \mathcal{B} could have been extended in $F \setminus Y_A$ or $F \setminus Y_B$, respectively. This will contradict the definition of \mathcal{B} . As there are no edges from B_3 to A_ℓ , we have $N(B_3) \cap A = A_1$. Suppose also that A was added to \mathcal{B} before B and so $B_3 \cap Y_A = \emptyset$. As B_3 is disjoint from A and $N(B_3) \cap A = A_1$, when A was added to \mathcal{B} it could have been extended in $F \setminus Y_A$ with the addition of B_3 . This is a contradiction, as A was chosen to be non-extendible in $F \setminus Y_A$. An analogous argument shows we obtain a

contradiction if B was added to \mathcal{B} first (B would be extendible in $F \setminus Y_B$ with the addition of A_3). So there are either edges from B_3 to A_ℓ or from A_3 to B_k .

Now suppose that $y \in A_\ell$ has a neighbour in B_3 . As $B_3 \cap A = \emptyset$ and $A_1 = B_2$, we have $y \in B_4$. By Proposition 3.38, A and B do not intersect on vertices of inner clusters. So y is in an outer cluster of B and so $k \leq 5$. As $y \in N[B_k]$, by assumption y is typical. Applying Proposition 3.33 to B and A gives that $B_4 = A_\ell$.

Applying Proposition 3.39 shows that no vertex in A_ℓ has a neighbour outside of $A_{\ell-1} \cup A_\ell \cup B_3$. Given this, as A and B are braids, no vertex in $A \cup (B_1, \dots, B_4)$ has a neighbour outside of $A \cup (B_1, \dots, B_4)$. As $f(F) = m(n)$, F is connected and hence $V(F) = V(A \cup (B_1, \dots, B_4))$ and $|V(A)| = |V(F)| - 3$ as required.

Now suppose $y \in B_k$ has a neighbour in A_3 . As A is a braid $N(A_3) \subseteq A_2 \cup A_3 \cup A_4$. As A_3 is disjoint from B and $A_2 = B_1$, we have $y \in A_4$. By assumption y is typical. So applying Proposition 3.33 gives that $B_k = A_4$. Now applying Proposition 3.39 shows that $N(B_k) \subseteq B_{k-1} \cup B_k \cup A_3$. As in the previous paragraph, no vertex in $(A_1, A_2, A_3, A_4) \cup B$ has a neighbour outside of $(A_1, A_2, A_3, A_4) \cup B$ and so $|V(B)| = |V(F)| - 3$ as required. \square

This concludes the case (A2). For (a), it remains to resolve case (A3).

Claim 3.43. *If A and B intersect then we are done.*

Proof. By Claims 3.41 and 3.42, if $A \cap B \neq \emptyset$ then we may assume they intersect only on end clusters and, without loss of generality, there exists $u \in A_1 \cap B_1$. By assumption u is typical, so applying Proposition 3.33 gives $A_1 = B_1$. Without loss of generality, suppose B was added before A to \mathcal{B} . By Claims 3.41 and 3.42, $A_2 \cup A_3$ is disjoint from B . As $N(A_2) \subseteq A_1 \cup A_2 \cup A_3$, we have $N(A_2) \cap B = B_1$. Therefore when B was added it was extendible in $F \setminus Y_B$ (the cluster A_2 could have been added), a contradiction. So u is atypical as required. \square

This completes the proof in case (a).

Claim 3.44. *If B is disjoint from all other braids in \mathcal{B} and is non-extendible in F , then $|V(B)| = |V(F)| - 3$.*

Proof. Using Lemma 3.34, there exists a set C of cardinality 3 disjoint from $\bigcup_{j=1}^4 B_j$ such that (C, B_1, B_2, B_3) is a braid in F . As B_1 has no neighbours in B_3, \dots, B_k , C is disjoint from B . Analogously, there exists some set D disjoint from B such that $(B_{k-2}, B_{k-1}, B_k, D)$ is a braid in F .

As B is non-extendible in F , some vertex in C has a neighbour x in B_k (otherwise (C, B_1, \dots, B_k) is a braid in F). Hence there exists some $v \in C \cap D$, as C has no neighbour in B_{k-1} and $N(B_k) \subseteq B_{k-1} \cup B_k \cup D$, as $(B_{k-2}, B_{k-1}, B_k, D)$ is a braid in F . By assumption, v is typical. So we may apply Proposition 3.33 to the braids (C, B_1, \dots, B_k) and $(B_{k-2}, B_{k-1}, B_k, D)$ to obtain that $C = D$.

As B is a braid and v is typical, using Proposition 3.39 gives that (B_1, C, B_k) is a braid in F (and so $N(C)$ is contained within this braid). Therefore no vertex in $B \cup C$ has a neighbour outside of $B \cup C$. As $f(F) = m(n)$, F is connected and hence $V(F) = V(B \cup C)$ and $|V(B)| = |V(F)| - 3$ as required. \square

This completes the proof of Lemma 3.40. \square

We say that an induced cycle in G *passes through* B if it contains a vertex from every cluster of B . The following lemma will be useful.

Lemma 3.45. *Fix $0 < r < 1$. Let F be an n vertex graph with $f(F) = m(n)$, where n is taken to be sufficiently large with respect to r .*

- (i) *The vertices of F can be covered by a collection \mathcal{B} of $O(1)$ 3-braids such that, for any $B, B' \in \mathcal{B}$, B and B' intersect only on outer clusters, and a set U of cardinality $O(1)$.*
- (ii) *F contains a 3-braid B such that $|V(B)| = \Omega(n)$.*
- (iii) *For any 3-braid B on rn vertices, at least $f(H_n) (1 - 3^{-rn/6})$ induced cycles in F pass through B .*

Proof. Set $\mathcal{B} := \mathcal{B}(F, v)$. By construction of \mathcal{B} , every vertex in $U := V(F) \setminus \bigcup_{B \in \mathcal{B}} B$ is atypical. Therefore by Corollary 3.31 and Lemma 3.5, $|U| = O(1)$. Also by Proposition 3.38, any distinct $B, B' \in \mathcal{B}$, B and B' intersect only on outer clusters. Therefore, to prove (i) it remains to show that $|\mathcal{B}| = O(1)$.

By Lemma 3.40 and definition of \mathcal{B} , either \mathcal{B} contains a braid on $|V(F)| - 3$ vertices (from which (i) and (ii) immediately follow) or, for each $B = (B_1, \dots, B_k) \in \mathcal{B}$, we have that $N[B_1 \cup B_k]$ contains an atypical vertex a_B . Say that a_B is an *internal anchor* for B if $a_B \in B_1 \cup B_k$, else say that a_B is an *external anchor* for B . Let $S := \{a_B : B \in \mathcal{B}\}$.

As $\Delta(F) < 30$ (by Lemma 3.5) and as the braids in \mathcal{B} only intersect on outer clusters, each vertex in F is in an end cluster of at most $O(1)$ braids in \mathcal{B} . Therefore each vertex $s \in S$ is an internal anchor for at most $O(1)$ 3-braids in \mathcal{B} . Similarly, an external anchor is adjacent to $O(1)$ vertices in F . Each of these vertices is in the end cluster of at most $O(1)$ braids, so s can be an external anchor for at most $O(1)$ braids in \mathcal{B} . Putting this together gives $|\mathcal{B}| = O(|S|)$. As every vertex in S is atypical, this gives that $|\mathcal{B}| = O(1)$ as required.

This proves (i). It immediately follows that some braid in \mathcal{B} contains $\Omega(n)$ vertices, proving (ii).

To prove (iii), observe that if an induced cycle does not pass through a 3-braid $B = (B_1, \dots, B_k)$ on rn vertices, then it is either a C_4 contained in B (there are at most $O(n^4)$ of these), or it is contained in $V(F) \setminus \bigcup_{i=3}^{k-2} B_i$ (by Lemma 3.2, there are at most $((1-r)n + 12) \binom{30}{2} 3^{((1-r)n+9)/3}$ of these). Therefore at most

$$((1-r)n + 12) \binom{30}{2} 3^{((1-r)n+9)/3} + O(n^4)$$

induced cycles of F do not pass through B . So for n_0 sufficiently large, at least

$$f(H_n) \left(1 - 3^{-rn/6}\right)$$

induced cycles pass through B . This proves (iii). □

The next lemma shows that G_{\max} is a cyclic braid. Given this, it will remain to determine the cluster sizes and whether there are edges within the clusters of G_{\max} .

Lemma 3.46. *G_{\max} is a cyclic braid.*

Proof. By Lemma 3.45 (ii) G_{\max} contains a 3-braid B such that $|V(B)| = \Omega(n)$. Let $B := (B_1, \dots, B_{Cn/3})$ be the longest 3-braid in G_{\max} . Let Q be the number of induced cycles in G_{\max} that pass through B . By Lemma 3.45 (iii),

$$Q \geq f(H_n) \left(1 - 3^{-Cn/6}\right). \quad (3.47)$$

Now let $G' := G \setminus \bigcup_{i=2}^{Cn/3-1} B_i$. Let x and y be two new vertices and define H to be the graph with vertex set $V(H) := V(G') \cup \{x, y\}$, and edge set

$$E(H) := E(G') \cup \{xb : b \in B_1\} \cup \{yb : b \in B_{Cn/3}\}.$$

See Figure 2.11 for an illustration of how we obtain H from G_{\max} . We have

$$Q = 3^{(Cn-6)/3} p_2(H; x, y). \quad (3.48)$$

For a subgraph $G' \subseteq G_{\max}$, let $p_2(G'; x, y)$ be the number of induced paths between x and y contained within G' . Combining (3.47) and (3.48) gives

$$p_2(H; x, y) \geq 3^{-(Cn-6)/3} \cdot f(H_n) \left(1 - 3^{-Cn/6}\right). \quad (3.49)$$

We now focus on the structure of H . Let us call a central cluster W of a 3-braid B *supercentral* if for any $w \in W$ and v in an end cluster of B , $d(w, v) \geq 5$. By Lemma 3.45 (ii), the vertices of G_{\max} can be covered by a collection \mathcal{B} of $O(1)$ 3-braids (that intersect only on outer clusters) and a set U of cardinality $O(1)$. Define a new graph H' via the following process.

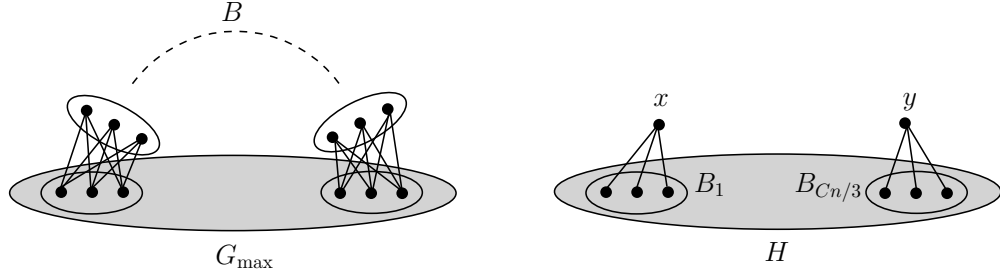


Figure 2.11: The shaded area is the part of the graph whose structure we wish to determine. The dashed line represents the rest of the braid B .

- Set $F_1 := H$ and $\mathcal{B}_1 = \mathcal{B}$.
- Let i be maximal such that we have defined F_i . Suppose there exists a vertex $v_i \in F_i$, contained in a supercentral cluster C_i of a (unique) 3-braid $M_i \in \mathcal{B}_i$, where C_i is adjacent to clusters C_1^i and C_2^i . Then define F_{i+1} to be the graph obtained from F_i by deleting C_i and adding every edge $\{uw : u \in C_1^i, w \in C_2^i\}$. Define M_i' to be the 3-braid obtained from M_i by deleting C_i and adding every edge $\{uw : u \in C_1^i, w \in C_2^i\}$. Set \mathcal{B}_{i+1} to be the collection obtained from \mathcal{B}_i by replacing each braid M containing C_i with M' .
- If there exists no such vertex v_i , define $H' := F_i$.

The process will terminate as H has a finite number of vertices. Observe that F_{i+1} is a braid if and only if F_i is a braid. In addition, when H' is a braid this process can be reversed to find H . We now show that H' is a braid.

When this process terminates, the only vertices remaining in H' are the vertices of U and at most $O(1)$ vertices of each 3-braid in \mathcal{B} . Therefore, $|V(H')| = O(1)$. At stage i of the process, F_{i+1} contains all induced cycles of F_i that do not pass through M_i and a third of the number of cycles in F_i that do pass through M_i . Thus we have

$$p_2(F_{i+1}; x, y) \geq 3^{-1} \cdot p_2(F_i; x, y), \quad (3.50)$$

and so

$$p_2(H'; x, y) \geq 3^{-(|V(H)|-|V(H')|)/3} \cdot p_2(H; x, y), \quad (3.51)$$

Combining (3.49) and (3.51) and observing that $|V(H)| = (1 - C)n + 8$ gives

$$p_2(H'; x, y) \geq 3^{(-n-2+|V(H')|)/3} \cdot f(H_n) \left(1 - 3^{-Cn/6}\right). \quad (3.52)$$

As

$$3^{(-n-2+|V(H')|)/3} f(H_n) = f_2(|V(H')|) + o(1),$$

when n_0 is sufficiently large we have

$$3^{(-n-2+|V(H')|)/3} \cdot f(H_n) \left(1 - 3^{-Cn/6}\right) > f_2(|V(H')|) - 1. \quad (3.53)$$

As $p_2(H'; x, y)$ is an integer, by taking n_0 to be sufficiently large, (3.52) and (3.53) give

$$p_2(H'; x, y) \geq f_2(|V(H')|).$$

Therefore, by Theorem 2.6, H' is isomorphic to a graph in $\mathcal{F}_{|V(H')|}$. Thus H' is a braid. By reversing the process applied above (adding back in the supercentral clusters) to recreate H from H' , we see that H is a graph in $\mathcal{F}_{|V(H)|}$, and hence G_{\max} is a cyclic braid, a contradiction. This completes the proof of the lemma. \square

In fact this proof immediately gives something stronger.

Corollary 3.54. *We have the following:*

- when $n \equiv 0$ modulo 3, G_{\max} has exactly $n/3$ clusters of size 3;
- when $n \equiv 1$ modulo 3, G_{\max} has either one cluster of size 4 and $(n-4)/3$ of size 3, or two of size 2 and $(n-4)/3$ of size 3;
- when $n \equiv 2$ modulo 3, G_{\max} has exactly one cluster of size 2 and $(n-2)/3$ of size 3.

We are now in a position to complete the proof of Theorem 1.3. The next lemma shows that the clusters in G_{\max} do not contain any edges, and thus we will prove the required result for $n \equiv 0$ or 2 modulo 3. In the remaining case we will need a short argument to decide whether the graph contains two clusters of size two, or one of size four. In both cases, the arguments are essentially routine checks.

Lemma 3.55. *No cluster of G_{\max} contains edges.*

Proof. First observe, that if e is an edge within a cluster, the only induced cycles containing e can be triangles, either contained within the cluster, or containing exactly one vertex from a neighbouring cluster; or induced copies of C_4 within the cluster (in the case that the cluster contains 4 vertices).

Let B be a cluster adjacent to clusters A and C . Suppose there exists an edge $e = uv$ where $u, v \in V(B)$. The edge e is contained in at most $|A| + |C| + (|B| - 2)$ induced cycles within G_{\max} . The graph $G' = G_{\max} \setminus \{e\}$ will contain at least $|A||C|$ induced copies of C_4 (for any $x \in A, y \in C$, the set $\{x, y, u, v\}$ induces a C_4) that are not induced cycles in G_{\max} .

As G_{\max} does not contain both a cluster of size 2 and a cluster of size 4, we have

$$|A||C| > |A| + |C| + (|B| - 2),$$

unless $|B| = 3$ and at least one of $|A|$ or $|C|$ is equal to 2. Except for this case, the number of induced cycles in $G' = G_{\max} \setminus \{e\}$ is greater than the number of induced cycles in G_{\max} , a contradiction.

Now suppose $|B| = 3$ and suppose without loss of generality that $A = \{a_1, a_2\}$. First consider the case where $|C| = 3$. Suppose B contains an edge $e = uv$. This edge is contained in at most 6 triangles in G_{\max} . By the argument of the previous paragraph, A does not contain an edge. The graph $G' = G_{\max} \setminus \{e\}$ will contain at least 7 induced copies of C_4 that are not induced cycles in G_{\max} (for any $x \in A$, $y \in C$, the sets $\{x, y, u, v\}$ and $\{a_1, a_2, u, v\}$ induce copies of C_4). Thus $f(G') > f(G_{\max})$, a contradiction.

The remaining case to consider is when $C = \{c_1, c_2\}$. If B contains an edge e , this edge is contained in at most 5 triangles in G_{\max} . The graph $G' = G_{\max} \setminus \{e\}$ contains at least 6 induced copies of C_4 that are not induced cycles in G_{\max} . Thus $f(G') > f(G_{\max})$, a contradiction. So no cluster in G_{\max} contains an edge. \square

We have proved Theorem 1.3 in the cases where $n \equiv 0$ or 2 modulo 3. It remains to prove the result in the case $n \equiv 1$ modulo 3.

Lemma 3.56. *When $n \equiv 1$ modulo 3, G_{\max} is isomorphic to H_n .*

Proof. By Corollary 3.54 and Lemma 3.55, we know that G_{\max} is one of two empty cyclic braids. One possibility is that it is isomorphic to H_n . The other possibility is that G_{\max} is an empty cyclic braid G_2 with exactly two clusters of size 2, and the rest of size 3. An induced cycle in H_n or G_2 either contains exactly one vertex from each cluster, or is an induced copy of C_4 . In both H_n and G_2 , the number of induced cycles containing exactly one vertex from each cluster is $4 \cdot 3^{(n-4)/3}$. Thus the only difference in $f(H_n)$ and $f(G_2)$ comes from the number of induced copies of C_4 .

There are two types of C_4 . Type 1 contains vertices from exactly two clusters. Type 2 contains vertices from three clusters. The graph H_n contains $3(n+5)$ induced type 1 cycles; G_2 contains at most $3n-14$ of this form (fewer if the two clusters of size 2 are not adjacent). The graph H_n contains $9(n+4)$ induced type 2 cycles; G_2 contains at most $9n-42$ of this form. Thus H_n contains more induced cycles than G_2 and therefore G_{\max} is isomorphic to H_n . \square

2.4 Proof of Theorem 1.6

The proof of Theorem 1.6 follows the same lines as the proof of Theorem 1.3. Before proceeding with the details of the proof, we first give an outline of what is to come. Let $0 < \alpha < 1$ be any constant and let F be an n -vertex graph containing at least $\alpha \cdot m(n)$

induced cycles. We will show that it is possible to delete a constant number of vertices from F to give a graph F' with maximum degree bounded by a constant. Applying Lemma 3.22 to F' then shows that the number of atypical vertices in F' is bounded by a constant. The result will immediately follow. We cannot simply apply Lemma 3.22 to F , as F may contain vertices of arbitrarily large degree.

Lemma 4.1. *Fix $0 < \alpha < 1$ and define $\Delta^* = \Delta^*(\alpha)$ to be the smallest integer such that $6\binom{\Delta^*}{2}3^{(1-\Delta^*)/3} < \alpha \cdot 4 \cdot 3^{-4/3}$. For n sufficiently large, let F be an n -vertex graph with $f(F) \geq \alpha \cdot m(n)$. Then there exists a constant $C = C(\alpha)$ such that deleting edges incident to C vertices of F gives a graph H with $\Delta(H) \leq \Delta^*$. Moreover, $f(H) \geq \frac{\alpha}{2} \cdot m(n)$.*

Proof. Suppose that $\Delta(F) > \Delta^*$ (else we are trivially done). We create a new graph H , with $\Delta(H) \leq \Delta^*$, in the following manner. Define $F_1 := F$. Let i be maximal such that F_i has been defined. If there exists a vertex $v_i \in V(F_i)$ with $d(v_i) > \Delta^*$ then define $F_{i+1} := F_i \setminus \{v_i\}$. This process will terminate as F has a finite number of vertices. Suppose the process terminates at a graph F_j . We have $\Delta(F_j) \leq \Delta^*$. Define $H := F_j$.

To prove the first statement of the lemma it suffices to show that there exists some constant $C := C(\alpha)$ such that $j \leq C$. To prove this, we will bound the size of $f(F) - f(H)$ and use this to show that when $j > C$, we have $f(F) < \alpha \cdot m(n)$, a contradiction.

By Lemma 3.2,

$$f_{v_i}(F_i) \leq \binom{d_{F_i}(v_i)}{2} 3^{(n-i-d_{F_i}(v_i)-1)/3} \leq \binom{\Delta^*}{2} 3^{(n-i-\Delta^*-1)/3},$$

where the second inequality follows as the function $\binom{x}{2}3^{-x/3}$ is decreasing for $x \geq 6$. So

$$\begin{aligned} f(H) &= f(F) - \sum_{i=1}^{j-1} f_{v_i}(F_i) \\ &\geq f(F) - \binom{\Delta^*}{2} 3^{(n-\Delta^*+1)/3} \sum_{i=1}^{j-1} 3^{-i/3} \\ &\geq f(F) - 3 \binom{\Delta^*}{2} 3^{(n-\Delta^*+1)/3}. \end{aligned} \tag{4.2}$$

As $|V(H)| = n - j + 1$, for some constant c we have $f(H) \leq c \cdot 3^{(n-j+1)/3}$ by (3.6). Combining this with (4.2) gives:

$$f(F) \leq c \cdot 3^{(n-j+1)/3} + 3 \binom{\Delta^*}{2} 3^{(n-\Delta^*+1)/3}. \tag{4.3}$$

There exists a constant C such that, whenever $j > C$ and n is sufficiently large:

$$c \cdot 3^{(n-j+1)/3} < \frac{1}{2} (\alpha \cdot 4 \cdot 3^{-4/3}) 3^{n/3}. \tag{4.4}$$

Suppose that $j \geq C$ and let n be sufficiently large. Using the definition of Δ^* and substituting (4.4) into (4.3), gives

$$f(F) < \alpha \cdot 4 \cdot 3^{-4/3} 3^{n/3} < \alpha \cdot m(n),$$

where the final inequality is implied by (3.1). This contradicts the hypothesis that $f(F) \geq \alpha \cdot m(n)$. Therefore $j < C$, completing the proof of the first statement of the lemma.

We now prove the second statement. By (4.2) we have

$$f(H) \geq f(F) - 3 \binom{\Delta^*}{2} 3^{(n-\Delta^*+1)/3}.$$

Given the definition of Δ^* , we have $f(H) \geq \frac{\alpha}{2} m(n)$. □

Proof of Theorem 1.6. Let F be an n -vertex graph containing at least $\alpha \cdot m(n)$ induced cycles. By Lemma 4.1, there exist constants $c = c(\alpha)$ and $\Delta^* = \Delta^*(\alpha)$ such that deleting c vertices from F gives a graph F' with $\Delta(F') \leq \Delta^*$ and $f(F') \geq \frac{\alpha}{2} m(n)$. By applying Corollary 3.31 we deduce that there is a constant M such that at most M vertices are atypical for (F', v) . By Lemma 3.9, every vertex that is typical in (F', v) is contained in a central cluster of a 3-braid. We obtain H from F' by adding and deleting edges incident to vertices that are atypical for (F', v) . The result follows with $C(\alpha) = c + M$. □

2.5 Conclusion

For sufficiently large n , we have determined precisely which graphs on n vertices contain the maximum number of induced cycles. There are a number of interesting related questions.

In our proofs above we make no attempts to optimise the value of n_0 . We know that in some small cases, H_n does not contain the maximum number of induced cycles [99, 118]. We believe Theorem 1.3 ought to be true for $n_0 = 30$, but our proof gives a much larger number. There are several places where we could improve the bound, most notably by choosing a more careful strategy in Lemma 3.22. However we omit the details as the bound would still be extremely large.

It is natural to consider induced cycles of some length that depends on n . Let $c(n, l)$ be the maximum number of length l induced cycles that can be contained in a graph on n vertices. Let $C(n, l)$ be the set of graphs containing $c(n, l)$ induced cycles of length l .

Question 5.1. For $l = l(n)$, what is $C(n, l)$?

When l is linear we believe the following should hold.

Conjecture 5.2. Fix $c \in (0, 1)$. If $l(n) = \lceil cn \rceil$, then for sufficiently large n the only graphs in $C(n, l)$ are cyclic braids of length l .

Perhaps a similar result holds down to cycles of length $\Omega(\sqrt{n})$.

Question 5.3. *Suppose $l(n) > \sqrt{n}$. For sufficiently large n , are all graphs in $C(n, l)$ cyclic braids?*

Another related question is to ask about induced subgraphs which are subdivisions of some fixed graph H .

Question 5.4. *Given a fixed finite graph H , what is the maximum number of induced subdivisions of H that can be contained in a graph on n vertices (and which graphs realise this maximum)?*

Theorem 1.3 answers this question for $H = C_3$, but what happens for other graphs? For instance, which graphs maximise the number of induced subdivisions of $K_{1,3}$? For large n , are the extremal graphs always blow-ups of some subdivision of H ? The rooted version of the question is also interesting, where we consider induced subdivisions of H where the branch vertices are fixed (for instance Theorem 2.6 is a result of this form for $H = K_2$).

Chapter 3

Minimum Saturation and Weak Saturation in the Hypercube

Joint work with Jonathan A. Noel and Alex Scott

3.1 Introduction

Recall the following definitions from Chapter 1. Given a graph F and a family of graphs \mathcal{H} , a spanning subgraph G of F is said to be (F, \mathcal{H}) -saturated if it does not contain any $H \in \mathcal{H}$ as a subgraph, but for every edge $e \in E(F) \setminus E(G)$, the graph $G + e$ contains a copy of some $H \in \mathcal{H}$. The *saturation number* $\text{sat}(F, \mathcal{H})$ is the minimum number of edges in an (F, \mathcal{H}) -saturated graph. When \mathcal{H} consists of a single graph H , we write $\text{sat}(F, H)$ for $\text{sat}(F, \mathcal{H})$.

A spanning subgraph G of F is said to be *weakly* (F, \mathcal{H}) -saturated if the edges of $E(F) \setminus E(G)$ can be added to G , one edge at a time, in such a way that every added edge creates a new copy of some $H \in \mathcal{H}$. Note that we do not require here that G contains no copy of H . The minimum number of edges in a weakly (F, \mathcal{H}) -saturated graph is known as the *weak saturation number* and denoted $\text{wsat}(F, \mathcal{H})$. As above, when \mathcal{H} consists of a single graph H , we write $\text{wsat}(F, H)$ for $\text{wsat}(F, \mathcal{H})$ and say that a graph is weakly (F, H) -saturated.

Recall in addition that we have the following relationship between the saturation number and the weak saturation number.

$$\text{wsat}(F, \mathcal{H}) \leq \text{sat}(F, \mathcal{H}). \tag{1.1}$$

The d -dimensional hypercube, which we denote by Q_d , is the graph on vertex set $\{0, 1\}^d$ where two vertices are joined by an edge if and only if they differ in exactly one coordinate. In this chapter, we are interested in saturation and weak saturation problems in hypercubes and grids (which will be defined later). As discussed in Chapter 1, the value of $\text{sat}(Q_d, Q_m)$

has been previously studied by Choi and Guan [41] and Santolupo (see [53]). However, the best previously known general result on $\text{sat}(Q_d, Q_m)$ is the following Theorem of Johnson and Pinto [80].

Theorem 1.2 (Johnson and Pinto [80]). *For every fixed $m \geq 2$, there exists $0 < \varepsilon_m < 1$ such that*

$$\text{sat}(Q_d, Q_m) = O\left(d^{(1-\varepsilon_m)2^d}\right) = O(|E(Q_d)|/d^{\varepsilon_m}).$$

Moreover, when $m = 2$, Johnson and Pinto [80] proved that for every d , there exists a (Q_d, Q_2) -saturated graph with bounded average degree. They gave a construction which implied that

$$\text{sat}(Q_d, Q_2) < 10 \cdot 2^d. \quad (1.3)$$

This result inspired them to ask: For which fixed values of m is $\text{sat}(Q_d, Q_m) = O(2^d)$? Our main result on saturation shows that this is the case for every $m \geq 2$.

Theorem 1.4. *For every fixed $m \geq 2$, $\text{sat}(Q_d, Q_m) \leq (1 + o(1))72m^22^d$.*

Johnson and Pinto [80] also asked: For $d \geq m \geq 1$, what is the value of $\text{wsat}(Q_d, Q_m)$? They proved that $\text{wsat}(Q_d, Q_2) = 2^d - 1$ by constructing a weakly (Q_d, Q_m) -saturated spanning tree. We resolve their question with the following theorem.

Theorem 1.5. *For $d \geq m \geq 1$,*

$$\text{wsat}(Q_d, Q_m) = (m-1)2^d - \sum_{j=0}^{m-2} (m-1-j) \binom{d}{j}.$$

In fact, this is a particular case of a more general result (Theorem 3.1) about weak saturation in multidimensional grids. By combining Theorem 1.5 with (1.1) we get a lower bound for $\text{sat}(Q_d, Q_m)$ of $(m-1-o(1))2^d$. This shows that Theorem 1.4 is tight up to a (constant) factor of $O(m)$. This lower bound improves on the bound of $(m+1-o(1))2^{d-1}$ obtained by Johnson and Pinto [80].

The chapter is structured as follows. At the beginning of Section 3.2, we give a brief background to saturation numbers to place the problem discussed here in a wider context. Our goal is then to prove Theorem 1.4. To do this, we first review some fundamental properties of hypercubes which will be used in the proof. In Subsection 3.2.2, we briefly discuss the construction of Johnson and Pinto and its relation to our own. Then, we prove Theorem 1.4 by giving an explicit construction of a (Q_d, Q_m) -saturated graph with bounded average degree. In Section 3.3, we turn our attention to weak saturation. We give some history and then prove Theorem 1.5. We end the chapter in Section 3.4 by mentioning a number of open problems.

3.2 Minimum Saturation

The problem of determining $\text{sat}(F, H)$ was introduced in 1964 by Zykov [138] and independently by Erdős, Hajnal and Moon [49]. Motivated by a conjecture of Erdős and Gallai [48], they proved that $\text{sat}(K_n, K_t) = (t-2)(n - \frac{1}{2}(t-1))$, for $n \geq t$. In 1986, Kászonyi and Tuza [84] began a systematic survey of saturation numbers. For a graph H they proved a general upper bound on $\text{sat}(K_n, H)$. To state their theorem we first need a few definitions. As usual, we denote the independence number of H by $\alpha(H)$. We also define

$$u = u(H) = |V(H)| - \alpha(H) - 1,$$

and

$$d = d(H) = \min\{|E(H')| : H' \text{ is induced by } S \cup x\},$$

where S is a maximal independent set in H and $x \in V(H) \setminus S$. So for every vertex $x \in H$ and maximal independent set S not containing x , we have that x has at least d neighbours in S . They proved the following theorem.

Theorem 2.1 (Kászonyi and Tuza [84]). *For n sufficiently large,*

$$\text{sat}(K_n, H) \leq un + (d-1)(n-u)/2 - \binom{u+1}{2}.$$

They also proved an extension of this result to families (see [84]). Theorem 2.1 has recently been improved by Faudree and Gould [54].

Theorem 2.2 (Faudree and Gould [54]).

$$\text{sat}(K_n, H) \leq un + (d-1)(n-u)/2 - \binom{u+1}{2} - \frac{1}{2} \left\lfloor \frac{d^2}{4} \right\rfloor.$$

The bound in Theorem 2.2 is tight whenever H contains a clique, however it is not very good for sparse graphs. For example, when $H = P_4$, the 4 vertex path, the bound in Theorem 2.2 gives $\text{sat}(K_n, P_4) \leq n-2$. However, we can observe that when $n \geq 4$ is even, a matching of $n/2$ edges is a P_4 saturated graph and hence $\text{sat}(K_n, P_4) \leq n/2$. In fact, $\text{sat}(K_n, P_4) = n/2$ when n is even, as a (K_n, P_4) -saturated graph cannot contain two isolated vertices (adding the edge between these vertices would not create a copy of P_4). Similar deductions yield that for odd $n \geq 5$, we have $\text{sat}(K_n, P_4) = (n+3)/2$.

Observe that, in a (K_n, H) -saturated graph, every vertex must have degree at least $\delta(H) - 1$ (else we would not be able to add an edge to that vertex to create a copy of H). Thus we get the following trivial lower bound on $\text{sat}(K_n, \mathcal{H})$.

$$\text{sat}(K_n, \mathcal{H}) \geq \frac{\delta(\mathcal{H}) - 1}{2} n,$$

where $\delta(\mathcal{H}) := \min\{\delta(H) : H \in \mathcal{H}\}$.

An interesting property to note about $\text{sat}(K_n, H)$ is that, unlike $\text{ex}(K_n, H)$, it is not non-decreasing in n (the above analysis of $\text{sat}(K_n, P_4)$ gives an example of this). This behaviour motivated the following conjecture of Tuza.

Conjecture 2.3 ([125, 126]). *For every graph H , the limit $\lim_{n \rightarrow \infty} \frac{\text{sat}(K_n, H)}{n}$ exists.*

Although some progress has been made towards this conjecture in both directions (see [110, 112, 123]), it remains open.

The area of minimum saturation is a rich source of open problems; although $\text{sat}(K_n, H)$ is known for H a clique, path or a short cycle, most saturation numbers remain to be discovered. For $\text{sat}(F, H)$ where F is not K_n our knowledge is even sparser, despite a body of work addressing this. For additional background on known results and open questions, see [53].

3.2.1 Preliminaries

Let e_j denote the j th standard basis vector in \mathbb{F}_2^d ; ie., the vector in which the j th coordinate is equal to one and every other coordinate is zero. Given a vertex v of Q_d , let the *weight* of the vertex, denoted $|v|$, be the sum of its coordinates. A basic fact about hypercubes is that, for $d \geq m$, every copy Q of Q_m in Q_d is induced by a set of vertices of the form $\{v + \sum_{j \in J'} e_j : J' \subseteq J\}$ where v is a fixed vertex and $J \subseteq [d]$ has cardinality m . We say that the coordinates of J and $[d] \setminus J$ are *variable* and *fixed* under Q , respectively.

We will need a standard result from Coding Theory, due to Hamming [71] (for another reference, see [95, 96]). This result was also used by Johnson and Pinto [80] (and we use some ideas from [80] in our proof).

Theorem 2.4 (Hamming [71]). *There is an independent set $C \subseteq V(Q_{2^t-1})$ such that every vertex of $V(Q_{2^t-1}) \setminus C$ has a unique neighbour in C .*

A set C as in Theorem 2.4 is often referred to as a *Hamming code*. We remark that, since every vertex of Q_{2^t-1} has exactly $2^t - 1$ neighbours,

$$|C| = \frac{2^{2^t-1}}{(2^t - 1) + 1} = 2^{2^t-t-1}. \quad (2.5)$$

Another ingredient of our proof of Theorem 1.4 is the following result of Conder [43].

Theorem 2.6 (Conder [43]). *For every $s \geq 1$, there is a 3-colouring of the edges of Q_s in which there is no monochromatic cycle of length 4 or 6.*

In Subsection 3.2.3 we will give an overview of the construction and draw attention to where these results will be applied.

3.2.2 The Johnson and Pinto Construction

As mentioned above, Johnson and Pinto [80] obtained the following bound for the case $m = 2$:

$$\text{sat}(Q_d, Q_2) < 10 \cdot 2^d.$$

Before introducing our construction, we briefly sketch their construction. We hope this will provide the reader with a light introduction to some of the concepts we will be using later and will help draw attention to the similarities and differences between our construction and theirs. At the heart of the Johnson and Pinto construction lies the following claim.

Claim 2.7 (Johnson and Pinto [80]). *There exists a Q_2 -free spanning subgraph H of Q_{2^t-1} that has two disjoint independent dominating sets $C, D \subseteq V(H)$, where $|C| = 2^{2^t-t-1}$ and $|D| = 3 \cdot 2^{2^t-t-1}$. Further, H only contains edges incident to $C \cup D$ and $|E(H)| \leq 2^{2^t}$.*

In particular, C is taken to be a Hamming code (see Theorem 2.4) and D is taken to be the union of three disjoint translates of C .

Their construction lies in Q_d , where $d := 2(2^t - 1) + s$ and $0 \leq s < 2^{t+1}$. They view the set $[d]$ as the union of three intervals indexed by the elements of $[3]$, the first two of length $2^t - 1$ and the third of length s . Given $v \in V(Q_d)$, for $i \in \{1, 2\}$ let $v(i)$ denote the vertex of Q_{2^t-1} obtained by restricting v to the coordinates of the interval corresponding to i , and let $v(3)$ be the vertex of Q_s obtained by restricting v to its last s coordinates.

The vertices of Q_d are naturally partitioned into three sets, A_0, A_1, A_2 , where A_j is defined to be the set of all $v \in V(Q_d)$ for which exactly j members of $\{v(1), v(2)\}$ are in $C \cup D$. The idea is to create a graph G satisfying the following three properties:

- G contains no copy of Q_2 .
- G contains no edge uv with $u, v \in A_0$.
- For any edge e with both endpoints in A_0 , the graph $G + e$ contains a copy of Q_2 .

The final graph G' is obtained from G by adding a maximal set of edges that do not create a copy of Q_2 . Note that by the third property, no edge of A_0 will get added in the step and hence every edge of G' has an endpoint in $V(Q_d) \setminus A_0$.

For the purposes of our sketch, we briefly show why such a construction will give the bound $|E(G')| = O(2^d)$. We remark that Johnson and Pinto [80] count the edges more carefully to give a bound on $|E(G')|$ with a precise constant. We have

$$|A_0| = (|V(Q_{2^t-1})| - |C \cup D|)^2 \cdot 2^s = \left(2^{2^t-1} \cdot (1 - 4 \cdot 2^{-t})\right)^2 \cdot 2^s = 2^d (1 - 4 \cdot 2^{-t})^2.$$

As $d = \Theta(2^t)$, we have

$$|V(Q_d) \setminus A_0| = 2^d - 2^d(1 - 4 \cdot 2^{-t})^2 \leq 8 \cdot 2^d \cdot 2^{-t} = O(2^d/d).$$

As every edge of G' has an endpoint in $V(Q_d) \setminus A_0$, this implies $\text{sat}(Q_d, Q_2) \leq |E(G')| = O(2^d)$.

So now we have heuristically described the aim of the construction, we give the precise details. The graph G is created from $V(Q_d)$ in 4 steps.

- (1) For $j \in \{1, 2\}$, add all edges uv with $u(j) = v(j) \in C$ such that the fixed coordinates in the union of the intervals in $\{v(1), v(2), v(3)\} \setminus \{v(j)\}$ contain an even number of ones.
- (2) For $j \in \{1, 2\}$, add all edges uv with $u(j) = v(j) \in D$ such that the fixed coordinates in the union of the intervals in $\{v(1), v(2), v(3)\} \setminus \{v(j)\}$ contain an odd number of ones.
- (3) Add all edges uv where either $u(1)v(1) \in H$ or $u(2)v(2) \in H$.
- (4) Delete any edge with an endpoint in A_2 .

So by construction, every edge of G has an endpoint in A_1 and so G contains no edge uv with $u, v \in A_0$.

Let us first determine why G contains no copy of Q_2 . Suppose, for a contradiction, that G contains a copy Q of Q_2 . As every edge in G has an endpoint in A_1 , without loss of generality we may assume that $v \in Q$ with $v(1) \in C \cup D$. By (4), there is no edge of G with an endpoint in A_2 , so $V(Q) \subseteq A_0 \cup A_1$. Let J be the set of variable coordinates of Q , and for each $i \in [3]$, let $J(i)$ be the set of coordinates of J which are in the interval of $[d]$ corresponding to i .

Consider $J(1)$. If $|J(1)| = 2$, then every edge in Q was added in step (3), which is a contradiction as H is Q_2 -free. If $|J(1)| = 1$, then let w be the neighbour of v obtained by changing the coordinate not in $J(1)$. In particular, $w(1) = v(1)$ and $w(2) \notin C \cup D$ (by (4)). So both v and w are in A_1 . As G contains no edge in A_0 or A_2 , $V(Q) \subseteq A_1$ and in particular the neighbour x of v obtained by changing the coordinate of $J(1)$ satisfies $x(1) \in C \cup D$. As C and D are disjoint independent sets, one of $x(1)$ and $v(1)$ is in C and the other is in D . As $Q \subseteq A_1$, for all $u \in Q$ we have $u(2) \notin C \cup D$. Therefore the pair of edges e_1, e_2 obtained by changing the variable coordinate not in $J(1)$ were added in (1) or (2) (as every edge of H has an endpoint in $C \cup D$). However, as one of $x(1)$ and $v(1)$ is in C and the other is in D , because of the parity condition it is impossible for both e_1 and e_2 to be added via (1) or (2). This completes the case $|J(1)| = 1$.

The final case to consider is $|J(1)| = 0$. So in particular, for all $u \in Q$ we have $u(1) = v(1) \in C \cup D$. As our graph has no edge with an endpoint in A_2 and every edge in H has an endpoint in $C \cup D$, no edge of Q was added in step (3). Therefore all the edges in Q were added either in (1) (if $v(1) \in C$) or in (2) (if $v(1) \in D$). Again, this is a contradiction to the parity conditions imposed by (1) and (2).

It remains to show that for any edge uv with $u, v \in A_0$, the graph $G + e$ contains a copy of Q_2 . As $u, v \in A_0$, we have $u(1), v(1), u(2), v(2) \notin C \cup D$. Suppose $u(1) \neq v(1)$ (the other cases follow from similar deductions, in the case $u(3) \neq v(3)$ the properties of H are used). Suppose that the number of ones contained in the combination of $u(1)$ and $u(3)$ is even and that $v(1)$ has higher weight than $u(1)$. Let $c \in C$ be the unique neighbour of $u(2)$ in Q_{2^t-1} (which exists as C is a Hamming code). Let x be such that $x(1) = u(1)$, $x(2) = c$ and $x(3) = u(3)$. Let y be such that $y(1) = v(1)$, $y(2) = c$ and $y(3) = u(3)$. The edge xy is an edge added in (1). It follows from the construction of H (we omit the details here as this is a sketch of the main ideas behind their construction, but see [80] for details) that the edges ux and yv are added to G in step (3). The case where the number of ones in the combination of $v(1)$ and $v(3)$ is odd follows from an analogous argument (using D instead of C).

As we will see in the next section, our construction uses some of these ideas of Johnson and Pinto, however it is more than a straightforward generalisation of their ideas.

3.2.3 Definitions and Proof Outline

Our construction will use Theorems 2.4 and 2.6 to build a spanning subgraph G of Q_d which contains no copy of Q_m and which can be extended to a graph G' which is (Q_d, Q_m) -saturated and satisfies the bound in Theorem 1.4.

Throughout the proof, let $m \geq 2$ be fixed and let d be an integer which we may choose to be sufficiently large. Let t and s be the unique integers such that

$$d = 6m(2^t - 1) + s \text{ and } 0 \leq s < 6m2^t. \quad (2.8)$$

Similarly to Johnson and Pinto, we view the set $[d]$ as a union of intervals. However, we view $[d]$ as a union of m intervals of length $6(2^t - 1)$ indexed by the elements of $[m]$ followed by one interval of length s . Given $v \in V(Q_d)$ and $1 \leq i \leq m$, let $v(i)$ denote the vertex of $Q_{6(2^t-1)}$ obtained by restricting v to the coordinates of the interval corresponding to i , and let $v(m+1)$ be the vertex of Q_s obtained by restricting v to its last s coordinates. In our proof, unlike the proof of Johnson and Pinto, each of the first m intervals of $[d]$ is further divided into 6 subintervals of length $2^t - 1$ indexed by the pairs (r, γ) where $r \in \{0, 1\}$

and $\gamma \in \{0, 1, 2\}$. For $(i, r, \gamma) \in [m] \times \{0, 1\} \times \{0, 1, 2\}$, let $v(i, r, \gamma)$ be the vertex of Q_{2^t-1} obtained by restricting $v(i)$ to the subinterval corresponding to (r, γ) .

Let C be a subset of $V(Q_{2^t-1})$ as in Theorem 2.4. We will treat the vertices $v \in V(Q_d)$ differently depending on which of the triples $(i, r, \gamma) \in [m] \times \{0, 1\} \times \{0, 1, 2\}$ satisfy $v(i, r, \gamma) \in C$. For starters, let X denote the set of all vertices v for which there exists some i and $(r, \gamma) \neq (r', \gamma')$ such that both $v(i, r, \gamma)$ and $v(i, r', \gamma')$ are in C . The vertices of X will be isolated in G and will not play a large role in the construction. Now, divide the vertices of $Q_d \setminus X$ into sets A_0, \dots, A_m where A_j is defined to be the set of all vertices of $Q_d \setminus X$ for which there are exactly j triples (i, r, γ) such that $v(i, r, \gamma) \in C$. By (2.5) and (2.8), we have

$$|A_0| = (|V(Q_{2^t-1})| - |C|)^{6m} \cdot 2^s = \left(2^{2^t-1} \cdot (1 - 2^{-t})\right)^{6m} \cdot 2^s = 2^d(1 - 2^{-t})^{6m}.$$

Using (2.8), we have $d = \Theta(2^t)$ and so

$$|V(Q_d) \setminus A_0| = 2^d - 2^d(1 - 2^{-t})^{6m} = O(2^d \cdot 2^{-t}) = O(2^d/d). \quad (2.9)$$

Our goal, as in the construction of Johnson and Pinto, is to construct a graph $G \subseteq Q^d$ satisfying the following three properties.

(P1) G does not contain Q_m as a subgraph.

(P2) G does not contain any edge uv , where $u, v \in A_0$.

(P3) For any edge uv , where $u, v \in A_0$, the graph $G + uv$ contains a copy of Q_m .

Given such a graph G , let G' be a graph obtained by adding a maximal set of edges to G without creating a copy of Q_m . Then G' is (Q_d, Q_m) -saturated and every edge of G' has at least one endpoint in $V(Q_d) \setminus A_0$. So by (2.9), the total number of edges incident to $V(Q_d) \setminus A_0$ is at most $O(2^d)$ and so we obtain $\text{sat}(Q_d, Q_m) \leq |E(G')| = O(2^d)$, as desired.

Before giving the precise details of our construction, let us give a brief overview. As mentioned above, we wish to create a spanning subgraph $G \subseteq Q_d$ satisfying (P1)-(P3). Now not only will our construction satisfy (3), but in fact, for each edge $e = uv$ with $u, v \in A_0$, the addition of e to G will make a very particular copy of Q_m . It will make a copy $Q \subseteq Q_d$ such that Q has no more than one variable coordinate in each interval $J(i)$, for $i \in [m+1]$.

We define $\phi_1 : E(Q_{6(2^t-1)}) \rightarrow \{0, 1, 2\}$ and $\phi_2 : E(Q_s) \rightarrow \{0, 1, 2\}$ be colourings as in Theorem 2.6, and let $\phi : E(Q_{6(2^t-1)}) \sqcup E(Q_s) \rightarrow \{0, 1, 2\}$ be the mapping which is equal to ϕ_1 on $E(Q_{6(2^t-1)})$ and equal to ϕ_2 on $E(Q_s)$. This definition of ϕ will also be used throughout the proof in the next subsection. Let k be the unique element of $[m+1]$ such that $u(k) \neq v(k)$ and let $\gamma = \phi(u(k)v(k))$.

As both endpoints of e are in A_0 , there is no triple (j, r, γ) such that $u(j, r, \gamma) \in C$ or $v(j, r, \gamma) \in C$. For $j \neq k$ and every edge xy of Q whose variable coordinate is in $J(j)$, there will exist r_j such that either $x(j, r_j, \gamma) \in C$ or $y(j, r_j, \gamma) \in C$. So roughly speaking, the additional directions of the copy of Q_m containing e are all created by changing a particular subinterval to its unique neighbour in Q_{2^t-1} contained in C (such a neighbour exists because C is a dominating set). This is similar to the way the Johnson and Pinto construction works to create a copy of Q_2 (but they are changing some interval to a member of one of four disjoint Hamming codes).

Johnson and Pinto chose a particular subset of edges to add between A_0 and A_1 (and amalgamated their parity condition here). Our construction differs as, beginning with $V(Q_d)$, in the first step of our construction we will add every edge which joins a vertex of A_j with A_{j+1} for $0 \leq j \leq m-2$. So given that we have all these edges, the tricky part of the construction lies with choosing which edges to add within the sets A_j , in order to ensure that G simultaneously satisfies (P1) and that adding an edge uv with $u, v \in A_0$ creates a copy of Q_m of the type described in the previous paragraph. This is where our parity condition comes into play, and where we apply the result of Conder (Theorem 2.6). These are used to control the edges added within the A_j and to ‘break’ any potential copies of Q_m .

3.2.4 Proof of Theorem 1.4

We now give the details of our construction. Begin with $G = V(Q^d)$ (and $E(G) = \emptyset$). We add the following edges to G in two steps.

Step 1. Add to G every edge of Q_d which joins a vertex of A_j to a vertex of A_{j+1} for $0 \leq j \leq m-2$.

Step 2. Suppose that uv is an edge of Q_d such that $u, v \in A_j$ for some $1 \leq j \leq m-1$. Let k be the unique element of $[m+1]$ such that $u(k) \neq v(k)$. We add the edge uv to G if

- (i) either $k = m+1$ or there does not exist $(r', \gamma') \in \{0, 1\} \times \{0, 1, 2\}$ such that $u(k, r', \gamma') = v(k, r', \gamma') \in C$,

and for every $(i, r, \gamma) \in [m] \times \{0, 1\} \times \{0, 1, 2\}$ such that $i \neq k$ and $v(i, r, \gamma) \in C$ we have

- (ii) $\gamma = \phi(u(k)v(k))$, and
- (iii) $|v| + |v(k)| + |v(i)| + j - 1 \equiv r \pmod{2}$.

Note that (iii) is well defined since u and v differ only on a coordinate of the interval corresponding to k and so for every $i \neq k$ we have

$$|v| + |v(k)| + |v(i)| \equiv |u| + |u(k)| + |u(i)| \pmod{2}.$$

Before moving on we make a few observations, of which all but two follow immediately from Steps 1 and 2.

Observation 2.10. $A_m \cup X$ is a set of isolated vertices in G .

Observation 2.11. A_0 is an independent set in G .

The next observation follows from Step 2 (i).

Observation 2.12. If uv is an edge of G such that $u(i, r, \gamma) = v(i, r, \gamma) \in C$, then $u(i, r', \gamma') = v(i, r', \gamma')$ for every pair $(r', \gamma') \in \{0, 1\} \times \{0, 1, 2\}$.

Observation 2.13. If uv is an edge of G such that $u(m+1) \neq v(m+1)$, then $u, v \in A_j$ for some $1 \leq j \leq m-1$.

The next observation follows from the fact that C is an independent set in Q_{2^t-1} .

Observation 2.14. If uv is an edge of G such that $u(i, r, \gamma) \neq v(i, r, \gamma)$, then at most one of $u(i, r, \gamma)$ or $v(i, r, \gamma)$ is in C .

To complete the proof of Theorem 1.4, it suffices to establish the following two lemmas. We state these now and show that they imply Theorem 1.4 before proving the lemmas themselves.

Lemma 2.15. *For every edge uv of Q_d such that $u, v \in A_0$, the graph $G + uv$ contains a copy of Q_m .*

Lemma 2.16. *G does not contain Q_m as a subgraph.*

Proof of Theorem 1.4. If G is not (Q_d, Q_m) -saturated, then by Lemma 2.16 we can extend G to a (Q_d, Q_m) -saturated graph G' by adding a maximal set of edges which do not create a copy of Q_m . By Lemma 2.15 none of these additional edges are between vertices in A_0 and so, by Observation 2.11, A_0 is an independent set in G' . Thus, every edge of G' has at least one endpoint in $(\bigcup_{k=1}^m A_k) \cup X$. The total number of edges incident to vertices of this set is at most

$$d \left| \left(\bigcup_{k=1}^m A_k \right) \cup X \right| = d|A_1| + d \left| \left(\bigcup_{k=2}^m A_k \right) \cup X \right|$$

Note that $|(\bigcup_{k=2}^m A_k) \cup X| = O(|C|^2 2^{d-2(2^t-1)})$, which is $O(2^d/d^2)$ by (2.5) and (2.8). Therefore the second term of the above expression is $o(2^d)$ and it suffices to bound $d|A_1|$. We have

$$\begin{aligned} d|A_1| &\leq d \left(6m|C|2^{d-(2^t-1)} \right) = d \left(6m2^{2^t-t-1}2^{d-(2^t-1)} \right) = 6md2^{d-t} \\ &= 6m(6m(2^t-1) + s)2^{d-t} < 72m^22^d \end{aligned}$$

by (2.8). The result follows. \square

Remark 2.17. Note that if d is of the form $6m(2^t-1)$ for some t , then we obtain a better bound: $\text{sat}(Q_d, Q_m) \leq (1+o(1))36m^22^d$.

Thus, it suffices to prove Lemmas 2.15 and 2.16. See Figure 3.1 for an illustration of the proof of Lemma 2.15 in a particular case.

Proof of Lemma 2.15. Let uv be an edge of Q_d where $u, v \in A_0$ and let k be the unique element of $[m+1]$ for which $u(k) \neq v(k)$. Define $\gamma := \phi(u(k)v(k))$ and let $I \subseteq [m] \setminus \{k\}$ be a set of size $m-1$. For $i \in I$, pick $r_i \in \{0, 1\}$ so that

$$|v| + |v(k)| + |v(i)| \equiv r_i \pmod{2}. \quad (2.18)$$

For each $i \in I$, let c_i be the unique neighbour of $v(i, r_i, \gamma)$ in Q_{2^t-1} contained in C (which exists as C is a dominating set). Given $I' \subseteq I$, we let $v_{I'}$ and $u_{I'}$ be the vertices of Q_d such that $u_{I'}(i, r_i, \gamma) = v_{I'}(i, r_i, \gamma) = c_i$ for all $i \in I'$ and, on all other coordinates, $u_{I'}$ and $v_{I'}$ agree with u and v , respectively. In particular, $v_\emptyset = v$ and $u_\emptyset = u$. We will show that $\{v_{I'} : I' \subseteq I\} \cup \{u_{I'} : I' \subseteq I\}$ induces a copy of Q_m in $G + uv$, which will complete the proof.

If $I' \subsetneq I$, then for any $i' \in I \setminus I'$ we see that $v_{I'}$ is adjacent to $v_{I' \cup \{i'\}}$ and $u_{I'}$ is adjacent to $u_{I' \cup \{i'\}}$ in G via edges added in Step 1. We now wish to show that for $\emptyset \neq I' \subseteq I$, $u_{I'}$ is adjacent to $v_{I'}$ via an edge added in Step 2. As $u_{I'}, v_{I'} \in A_{|I'|}$, it suffices to show that for each $\emptyset \neq I' \subseteq I$ and each $i \in I'$, we have

$$|v_{I'}| + |v_{I'}(k)| + |v_{I'}(i)| \equiv r_i + |I'| - 1 \pmod{2}. \quad (2.19)$$

For each $i \in I'$, as C is a dominating set, $v(i, r_i, \gamma)$ differs from c_i in exactly one coordinate. Consider the process of transforming v to $v_{I'}$ by changing $|I'|$ coordinates (one by one). Each time a coordinate is changed during this process the parity also changes. At no point is a coordinate in the k th interval changed and a coordinate in the i th interval is changed in exactly one step. Using this analysis and (2.18) gives (2.19). This completes the proof. \square

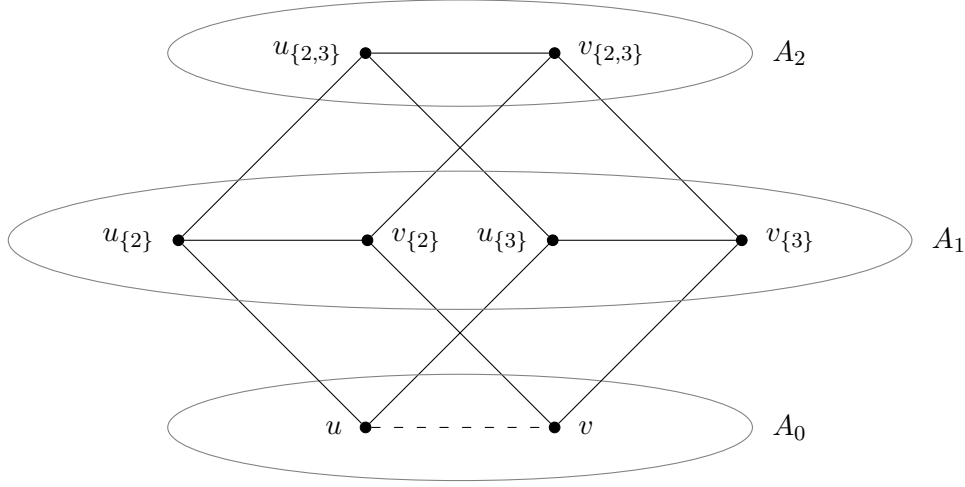


Figure 3.1: An illustration of the proof of Lemma 2.15 in the case that $m = 3$ and $k = 1$.

Proof of Lemma 2.16. Suppose, to the contrary, that G contains a copy Q of Q_m . Let $J \subseteq [d]$ be the set of coordinates which are variable under Q . For each $i \in [m + 1]$, let $J(i)$ be the set of coordinates of J which are in the interval of $[d]$ corresponding to i . Moreover, given $(i, r, \gamma) \in [m] \times \{0, 1\} \times \{0, 1, 2\}$, let $J(i, r, \gamma)$ be the set of coordinates of J which are in the interval of $[d]$ corresponding to (i, r, γ) .

Claim 2.20. *For each $i \in [m]$ there is a pair $(r_i, \gamma_i) \in \{0, 1\} \times \{0, 1, 2\}$ such that $J(i) = J(i, r_i, \gamma_i)$.*

Proof. Suppose to the contrary that there exists $k \in [m]$ and $(r, \gamma) \neq (r', \gamma')$ such that both $J(k, r, \gamma)$ and $J(k, r', \gamma')$ are non-empty.

First, suppose that there exists a vertex $v \in Q$ with $v(k, r, \gamma) \in C$. Let u be a neighbour of v in Q obtained by changing a coordinate in $J(k, r', \gamma')$. By Observation 2.12, the edge uv is not contained in G , which is a contradiction. Thus, $v(k, r, \gamma), v(k, r', \gamma') \notin C$ for all $v \in V(Q)$.

Now, let Q' be a copy of Q_2 in Q obtained by starting with an arbitrary vertex of Q and varying one coordinate of $J(k, r, \gamma)$ and one coordinate of $J(k, r', \gamma')$. By the result of the previous paragraph, we see that $V(Q') \subseteq A_j$ for some j and so every edge of Q' was added in Step 2 (therefore $j \geq 1$). Let uv be any edge in Q' . As $V(Q') \subseteq A_j$, by definition of A_j , for every $u \in V(Q')$ there exists some triple $(\ell, \hat{r}, \hat{\gamma})$, such that $v(\ell, \hat{r}, \hat{\gamma}) \in C$. By the conclusion of the previous paragraph and the condition in Step 2 (i), we have $\ell \neq k$.

By Step 2 (ii), $\hat{\gamma} = \phi(u(k)v(k))$. By taking the edges of Q' and restricting them to the interval of $[d]$ corresponding to k , we see that all of the resulting edges must receive the same colour under ϕ , namely $\hat{\gamma}$. This contradicts the fact that there is no copy of Q_2 which

is monochromatic under ϕ (as by definition and Theorem 2.6 we have chosen ϕ to be a colouring of $E(Q_{6(2^t-1)})$ with no monochromatic C_4 or C_6). This completes the proof of the claim. \square

Claim 2.21. *Suppose that $|J(i)| \geq 2$ for some $i \in [m]$ and let Q' be a copy of Q_2 in Q obtained by starting with an arbitrary vertex and varying two coordinates of $J(i)$. Then there is a unique vertex $v \in Q'$ with $v(i, r_i, \gamma_i) \in C$.*

Proof. By Claim 2.20 we have $J(i) = J(i, r_i, \gamma_i)$. By Observation 2.14, if uv is an edge of Q' , then we cannot have $u(i, r_i, \gamma_i), v(i, r_i, \gamma_i) \in C$. Thus, if there are two vertices u, v of Q' for which $u(i, r_i, \gamma_i), v(i, r_i, \gamma_i) \in C$, then they must be non-adjacent. So, in this case, the vertices of Q' alternate between A_j and A_{j+1} for some j and every edge of Q' was added in Step 1. However, this implies that for $w \in V(Q')$ such that $w(i, r_i, \gamma_i) \notin C$, the vertex $w(i, r_i, \gamma_i)$ of Q_{2^t-1} must have two distinct neighbours in C , which contradicts Theorem 2.4.

Now we assume that every vertex v of Q' satisfies $v(i, r_i, \gamma_i) \notin C$. However, in this case, we see that every edge of Q' was added in Step 2. As in the proof of Claim 2.20, we obtain a copy of Q_2 in $Q_{6(2^t-1)}$ which is monochromatic under ϕ , a contradiction. \square

Claim 2.22. $|J(i)| \leq 2$ for every $i \in [m]$.

Proof. Suppose not. Let Q' be a copy of Q_3 in Q obtained by starting with an arbitrary vertex and varying three coordinates of $J(i)$. If we vary any pair of these coordinates, leaving the third one fixed, we obtain a copy of Q_2 in Q' which must obey Claim 2.21. This implies that there are precisely two vertices $x, y \in Q'$ such that $x(i, r_i, \gamma_i), y(i, r_i, \gamma_i) \in C$ and they are at distance 3 in Q' (by Theorem 2.4 they cannot be nearer). However, now we get that the vertices of $Q' \setminus \{x, y\}$ induce a copy F of C_6 in Q , where every edge of F was added in Step 2. As in the proof of Claim 2.20, taking the edges of F and restricting them to the interval of $[d]$ corresponding to i , we see that all such edges must receive the same colour under ϕ . This contradicts the fact that there is no copy of C_6 which is monochromatic under ϕ and completes the proof of the claim. \square

Claim 2.23. $|J(m+1)| \leq 1$.

Proof. If not, then let Q' be a copy of Q_2 in Q obtained by starting at an arbitrary vertex and varying two coordinates of $J(m+1)$. Then, as in the proof of Claim 2.20, by restricting the edges of Q' to the last s coordinates we obtain a copy of Q_2 in Q_s which is monochromatic under ϕ , a contradiction. \square

Claim 2.24. *Suppose that uv and vw are distinct edges of Q where $u, v, w \in A_j$ for some j . Then there is some $k \in [m]$ such that $u(k) \neq v(k)$ and $v(k) \neq w(k)$.*

Proof. We show that there cannot exist distinct integers $k, k' \in [m+1]$ such that $u(k) \neq v(k)$ and $v(k') \neq w(k')$. This is sufficient to prove the claim since, by Claim 2.23, we cannot have $u(m+1) \neq v(m+1)$ and $v(m+1) \neq w(m+1)$.

Suppose that such integers k, k' exist, and let x be the unique vertex of Q_d distinct from v which is joined to both u and w in Q_d . Note that both of the edges xu and wx are present in Q . Also, $x \in A_j$ and so all of the edges uv, vw, wx, xu were added to G in Step 2. This implies that $j \geq 1$ and that there is a triple (i, r, γ) with $i \notin \{k, k'\}$ such that

$$u(i, r, \gamma) = v(i, r, \gamma) = w(i, r, \gamma) = x(i, r, \gamma) \in C.$$

By the definition of x and the assumption that such k, k' exist we have $|v| \equiv |x|$ modulo 2, $|v(i)| \equiv |x(i)|$ modulo 2 and $|v(k)| \not\equiv |x(k)|$ modulo 2. So the parity of $|v| + |v(k)| + |v(i)|$ is different from the parity of $|x| + |x(k)| + |x(i)|$ modulo 2 and so only one can be equivalent to r . By definition of Step 2, only one of the edges uv and xw can exist in G . This contradiction completes the proof. \square

Claim 2.25. *If $|J(m+1)| = 1$, then $|J(i)| \leq 1$ for all $i \in [m]$.*

Proof. If not, let Q' be a copy of Q_3 in Q obtained by starting at an arbitrary vertex and varying one coordinate of $J(m+1)$ and two of $J(i)$. Then, by Claim 2.21, there is an edge $uv \in E(Q')$ such that $u(i) \neq v(i)$ and $u, v \in A_j$ for some j (in fact, there are many such edges). Now, let w be the neighbour of v in Q' obtained by changing the coordinate of $J(m+1)$. By Observation 2.13, we have that $w \in A_j$ as well. However, this contradicts Claim 2.24. \square

Claim 2.26. *There is at most one $i \in [m]$ for which $|J(i)| = 2$.*

Proof. If not, let $i, i' \in [m]$ such that $|J(i)| = |J(i')| = 2$ and let Q' be a copy of Q_4 obtained by starting at an arbitrary vertex and varying two coordinates of $J(i)$ and two coordinates of $J(i')$. By Claim 2.21, we see that there must exist edges $uv, vw \in E(Q')$ such that $u(i) \neq v(i)$, $v(i') \neq w(i')$ and $u, v, w \in A_j$ for some j . This contradicts Claim 2.24 and completes the proof. \square

Now, let us complete the proof of the claim. Throughout, for each $i \in [m]$, we let (r_i, γ_i) be a pair such that $J(i) = J(i, r_i, \gamma_i)$, which exists by Claim 2.20. In what follows, we let j^* be the minimum integer j such that $V(Q) \cap A_j \neq \emptyset$ and let $v^* \in V(Q) \cap A_{j^*}$. Note that for every $i \in [m]$ such that $J(i) \neq \emptyset$ we must have $v^*(i, r_i, \gamma_i) \notin C$. If not, then starting with v^* and changing a coordinate of $J(i)$ yields a vertex of $Q \cap A_{j^*-1}$ by Observations 2.12

and 2.14, contradicting our choice of j^* . By Claims 2.25 and 2.26, there is at most one $i \in [m]$ such that $J(i) = \emptyset$. This implies that

$$j^* \in \{0, 1\}, \quad (2.27)$$

where j^* must equal zero if no such i exists. We divide the proof into cases.

Case 1: $|J(i)| = 1$ for every $i \in [m]$.

In this case, we must have $j^* = 0$ and so $v^* \in A_0$. However, if we start with v and change the coordinate of $J(i)$ for any i , then we obtain a vertex u with $u(i, r_i, \gamma_i) \in C$ by Observation 2.11. Thus, changing all m of these coordinates yields a vertex w such that $w(i, r_i, \gamma_i) \in C$ for all $i \in [m]$; that is, $w \in A_m$. This is a contradiction since, by Observation 2.10, w is an isolated vertex and therefore cannot belong to Q .

Case 2: $|J(m+1)| = 1$.

Let u be the neighbour of v^* in Q obtained by changing the coordinate in $J(m+1)$. We must have $u, v^* \in A_1$ by (2.27) and Observation 2.13. This implies that there is some $i \in [m]$ such that $J(i) = \emptyset$ and $v^*(i, r_i, \gamma_i) \in C$. Also, by Claim 2.25 and the pigeonhole principle, we must have $|J(i')| = 1$ for every $i' \in [m] \setminus \{i\}$.

Now, for each $i' \in [m] \setminus \{i\}$, let $w_{i'}$ be the neighbour of v^* in Q obtained by changing the coordinate in $J(i')$. By Claim 2.24 and the fact that $u, v^* \in A_1$, we cannot have $w_{i'} \in A_1$. So, by our choice of j^* , we must have $w_{i'}(i', r_{i'}, \gamma_{i'}) \in C$. As in the proof of Case 1, if we start with v^* and change the coordinate of $J(i')$ for every $i' \in [m] \setminus \{i\}$, we obtain a vertex of A_m contained in Q , contradicting Observation 2.10. This completes the proof in this case.

Case 3: $|J(i)| = 2$ for some $i \in [m]$.

By Claim 2.21, there is a neighbour u of v^* in Q obtained by changing a coordinate in $J(i)$ such that $u \in A_{j^*}$. This immediately implies that $j^* = 1$ by (2.27) and Observation 2.11.

Now, for each $i' \in [m]$ for which $|J(i')| = 1$, let $w_{i'}$ be the neighbour of v^* in Q obtained by changing the coordinate in $J(i')$. By Claim 2.24 and the fact that $u, v^* \in A_1$, we must have $w_{i'}(i', r_{i'}, \gamma_{i'}) \in C$. Thus, if we let x be the vertex obtained from v^* by changing the coordinate of $J(i')$ for every such i' , then $x \in A_{m-1}$.

Let Q' be the copy of Q_2 in Q obtained by starting at x and varying the two coordinates of $J(i)$. Then, by Claim 2.21, there must be some vertex y of Q' such that $y(i, r_i, \gamma_i) \in C$. However, this implies that $y \in A_m$, contradicting Observation 2.10. This completes the proof of Lemma 2.16 and of Theorem 1.4. \square

3.3 Weak Saturation

We now turn our attention to weak saturation. The study of $\text{wsat}(F, H)$ was initiated by Bollobás [29] in 1968. Motivated by the problem of determining the saturation number of k -uniform hypergraphs for $k \geq 3$, he proved that $\text{sat}(K_n, K_m) = \text{wsat}(K_n, K_m)$ for $3 \leq m < 7$. He later conjectured in [30] that $\text{sat}(K_n, K_m) = \text{wsat}(K_n, K_m)$ for all $m \geq 3$.

This conjecture was originally proved by Lovász [94] using a very beautiful generalisation of the Bollobás Two Families Theorem [28]. The ‘skew’ Two Families Theorem of Lovász has a lovely proof using exterior algebras. Kalai [81] and Frankl [58] independently discovered the ‘skew’ version, also proving it via using linear algebraic techniques in combinatorics. The conjecture has since been proved using a variety of different methods (see [82, 137]). More recently, other generalisations of the Two Families Theorem have been proved and applied to determine weak-saturation numbers in a hypergraph setting (see [3, 106]). For more background and references about results on the weak saturation number, see Section 10 of [53].

We now introduce the theorem we will prove in this section that will imply Theorem 1.5. Given $k \geq 2$ and $d \geq 1$, the d -dimensional k -grid denoted $[k]^d$ is the graph with vertex set $[k]^d$ where two vertices are adjacent if they differ by one in some coordinate and are equal in the other $d - 1$ coordinates. In particular, $[2]^d$ is isomorphic to the hypercube Q_d defined previously.

Given $k \geq r \geq 2$ and $d \geq m \geq 1$, say that a copy of $[r]^m$ in $[k]^d$ is *axis aligned* if it is induced by a set of vertices of the form $I_1 \times \cdots \times I_d$ where exactly m of the sets I_i are intervals of length r in $[k]$ and the rest are singletons. Let $\text{wsat}^*([k]^d, [r]^m)$ be the minimum number of edges in a spanning subgraph G of $[k]^d$ such that the edges of $E([k]^d) \setminus E(G)$ can be added to G , one edge at a time, such that every added edge creates an axis aligned copy of $[r]^m$. We prove the following.

Theorem 3.1. *For $k \geq r \geq 2$ and $d \geq m \geq 1$,*

$$\begin{aligned} \text{wsat}^*([k]^d, [r]^m) &= \sum_{j=0}^d \sum_{i=0}^{d-j} (m-1+i) \binom{d}{j} \binom{d-j}{i} (k-r+1)^j (r-2)^i \\ &\quad - \sum_{j=0}^{m-2} \sum_{i=0}^{d-j} (m-1-j) \binom{d}{j} \binom{d-j}{i} (k-r+1)^j (r-2)^i, \end{aligned}$$

where, by convention, $0^0 = 1$.

We remark that if $r = 2$ or $d = m$, then every copy of $[r]^m$ in $[k]^d$ is axis aligned. Thus, in these cases, $\text{wsat}^*([k]^d, [m]^r) = \text{wsat}([k]^d, [m]^r)$. In particular, Theorem 1.5 is implied by the case $k = r = 2$ of Theorem 3.1.

3.3.1 Outline of the Proof

Before embroiling ourselves in the details of the proof, let us briefly summarise how it will proceed in the case where $k = r = 2$. The multidimensional grid case will be slightly more technical (as now coordinates can take on more than two values and, for $r \geq 3$, $[r]^m$ is not vertex transitive), but the key ideas will emerge from sketching the hypercube case. In this case, we have $\text{wsat}^*([2]^d, [2]^m) = \text{wsat}(Q_d, Q_m)$ as every subhypercube of Q_d is axis-aligned.

To prove the bound on $\text{wsat}(Q_d, Q_m)$, we will provide both an upper bound and a matching lower bound. To give an upper bound, it suffices to provide a construction of a spanning subgraph $G \subseteq Q_d$ with the correct number of edges that is weakly- (Q_d, Q_m) -saturated. Let G be a graph obtained from $V(Q_d)$, by adding precisely $\min\{|v|, m - 1\}$ edges from every vertex v to vertices of lower weight. Observe that G is not unique. There may be many choices of edges to add.

In particular, every edge between vertices of weight at most $m - 1$ is contained in G . Now let us sketch why G is weakly- (Q_d, Q_m) -saturated. We must show that there exists an ordering of the edges of $E(Q_d) \setminus G$, such that the addition of each edge creates a new copy of Q_m . This will proceed by induction on $|v|$.

Let v be the vertex of lowest weight that has a neighbour u in Q_d of lower weight such that uv has not yet been added. Let G' be the union of $E(G)$ and the edges that have been added so far. By construction, there are (at least) $m - 1$ distinct edges f_1, \dots, f_{m-1} in G from v to vertices of lower weight. Let J be the set of coordinates that vary in the edges uv and f_1, \dots, f_{m-1} . So J is a set of m distinct coordinates.

The set of vertices $\left\{v - \sum_{j \in J'} e_j : J' \subseteq J\right\}$ induces a copy Q of Q_m in Q_d . As v is the vertex of highest weight in Q , by the inductive hypothesis and choice of J , every edge of Q apart from uv is contained in G' . So the addition of uv creates a new copy of Q_m , as required. This completes the discussion of the upper bound.

Finding a lower bound is more tricky, as we must show that *no* graph with fewer than $|E(G)|$ edges can be weakly- (Q_d, Q_m) -saturated. In order to do this, we apply a lovely linear algebraic lemma from [16]. Roughly speaking, this lemma tells us that to find a lower bound on $\text{wsat}(Q_d, Q_m)$ it suffices to find a vector space W and a family of vectors $\{f_e : e \in E(Q_d)\} \subseteq W$ satisfying some specific conditions. We find such a family and apply the lemma to obtain our matching lower bound. In particular, the vectors corresponding to the edges of G will be linearly independent. The precise construction of the vectors will be given in the proof of Theorem 3.1. This completes our overview of what is to come.

3.3.2 A Lovely Linear Algebraic Lemma

The lemma used to obtain the lower bound is stated as follows.

Lemma 3.2 (Balogh, Bollobás, Morris and Riordan [16]). *Let F be a graph, let \mathcal{H} be a collection of subgraphs of F , and let W be a vector space. Suppose that there exists a set $\{f_e : e \in E(F)\} \subseteq W$ such that for every $H \in \mathcal{H}$ there are non-zero scalars $\{c_e : e \in E(H)\}$ such that $\sum_{e \in E(H)} c_e f_e = 0$. Then*

$$\text{wsat}(F, \mathcal{H}) \geq \dim(\text{span}\{f_e : e \in E(F)\}).$$

We include the proof from [16] as it is short and contains a very beautiful idea. The essence of the proof is to track the dimension of the vector space associated to the edges in the graph at each step of the process, and to show that at no point does the dimension increase.

Proof. Let G be a weakly (F, \mathcal{H}) -saturated graph. Define $G_0 := G$ and label the edges of $E(F) \setminus E(G)$ by e_1, \dots, e_k such that for $1 \leq i \leq k$ there is a subgraph $H_i \in \mathcal{H}$ of $G_i := G_{i-1} + e_i$ such that H_i contains e_i . Thus, by hypothesis, f_{e_i} can be written as a linear combination of the vectors in $\{f_e : e \in E(H_i) \setminus \{e_i\}\}$. This implies that

$$\text{span}\{f_e : e \in E(G_0)\} = \text{span}\{f_e : e \in E(G_1)\} = \dots = \text{span}\{f_e : e \in E(G_k)\}.$$

Since $G = G_0$ and $F = G_k$, it must be the case that $|E(G)| \geq \dim(\text{span}\{f_e : e \in E(F)\})$. Since G was an arbitrary weakly (F, \mathcal{H}) -saturated graph, this completes the proof. \square

For other examples in which the notions of linear dependence and independence are applied to solve problems in weak saturation, see Alon [3], Kalai [82] and Pikhurko [110, 111].

Balogh, Bollobás, Morris and Riordan [16], used Lemma 3.2 to solve a problem about the \mathcal{H} -bootstrap process (defined in Section 1.4 of Chapter 1) that can be thought of as a ‘vertex’ version of a weak-saturation problem. For fixed graphs G and H , they considered the case where \mathcal{H} is a hypergraph with vertex set $V(\mathcal{H}) := V(G)$ and edge set $E(\mathcal{H}) := \{U \subseteq V(G) : G[U] \cong H\}$, (i.e. each edge is a set of vertices of G that induces a copy of H). They determined the minimum size of a percolating set in \mathcal{H} when $G = K_k^d$ and $H = K_r^m$, for all $k \geq r \geq 2$ and $d \geq m \geq 1$ (here K_k^d is the graph with vertex set $[k]^d$ and two vertices are adjacent if they differ (by any amount) in exactly one coordinate). They also determined the minimum size of a percolating set in \mathcal{H}' , where $V(\mathcal{H}') := V([k]^d)$ and a subset $U \subset V(\mathcal{H}')$ is in $E(\mathcal{H}')$ if and only if it induces an axis aligned copy of $[r]^m$ in $[k]^d$, for all $k \geq r \geq 2$ and $d \geq m \geq 1$. Somewhat surprisingly, the two quantities are the same. An analogous relationship between the \mathcal{H} -bootstrap process and weak-saturation will be discussed in Chapter 5.

To solve this problem, Balogh, Bollobás, Morris and Riordan gave a clever construction of a suitable vector space and then proved that it has the same dimension as a suitable

percolating set. The result then follows by Lemma 3.2. Our proof of Theorem 3.1 uses the same approach, but requires a different construction.

3.3.3 Proof of Theorem 3.1

In this subsection we prove Theorem 3.1, which immediately implies Theorem 1.5.

Given a vertex v of $[k]^d$, say that a coordinate of v is *large* if it is at least r and *small* otherwise. We let $L(v)$ denote the number of large coordinates of v and let $|v|$ denote the sum of the coordinates of v . Given $i \in [d]$, a *line* \mathcal{L} in direction i is a k -vertex path in $[k]^d$ in which any two vertices of \mathcal{L} differ on the i th coordinate.

Proof of Theorem 3.1. Let G be a spanning subgraph of $[k]^d$ in which, for each vertex v ,

- (1) we add every edge uv of $[k]^d$ such that $|u| = |v| - 1$ and u and v differ on a small coordinate of v , and
- (2) we add $\min\{L(v), m - 1\}$ edges uv of $[k]^d$ such that $|u| = |v| - 1$ and u and v differ on a large coordinate of v .

Let us now count the number of edges in G . We will first count the number of vertices v with i small coordinates strictly greater than one (note there exists no u with $|u| = |v| - 1$ that differs on a small coordinate equal to one). Each v contributes i edges via (1) and $\min\{L(v), m - 1\}$ edges via 2. For every v with i small coordinates, we will initially count $(m - 1 + i)$ edges, then we will subtract the number of edges which we have overcounted (for when $L(v) < m - 1$).

As there are d coordinates, each vertex can have $0 \leq j \leq d$ large coordinates. Given j , there $\binom{d}{j}$ choices for which set of coordinates are large and there are $k - r + 1$ possible values of each large coordinate. Now once the large coordinates have been fixed, a vertex can have $0 \leq i \leq d - j$ small coordinates strictly greater than one, and there are $r - 2$ choices for each such coordinate. So counting $(m - 1 + i)$ edges for every vertex with i small coordinates gives

$$\sum_{j=0}^d \sum_{i=0}^{d-j} (m - 1 + i) \binom{d}{j} \binom{d-j}{i} (k - r + 1)^j (r - 2)^i.$$

We now determine what is overcounted by this sum in order to subtract it. This concerns the vertices v where $L(v) < m - 1$. Each such vertex has $0 \leq j \leq m - 2$ large coordinates and we have counted $m - 1$ edges where we should have counted j . So we should subtract $m - 1 - j$ for each such vertex. Therefore we must subtract

$$\sum_{j=0}^{m-2} \sum_{i=0}^{d-j} (m - 1 - j) \binom{d}{j} \binom{d-j}{i} (k - r + 1)^j (r - 2)^i.$$

Putting this all together gives

$$|E(G)| = \sum_{j=0}^d \sum_{i=0}^{d-j} (m-1+i) \binom{d}{j} \binom{d-j}{i} (k-r+1)^j (r-2)^i \\ - \sum_{j=0}^{m-2} \sum_{i=0}^{d-j} (m-1-j) \binom{d}{j} \binom{d-j}{i} (k-r+1)^j (r-2)^i.$$

So, to prove the upper bound, we need only show that G is weakly $([k]^d, [r]^m)$ -saturated. We prove by induction on $|v|$ that adding a missing edge of $E([k]^d) \setminus E(G)$ from v to a vertex u with $|u| = |v| - 1$ creates a new copy of $[r]^m$. By construction, for every missing edge uv with $|u| = |v| - 1$, the coordinate on which u and v differ is large in v . For any vertex x , if $L(x) \leq m - 1$, then all edges xy with $|x| = |y| - 1$ are already present in G .

So let $v = (v_1, \dots, v_d)$ be a vertex with $L(v) \geq m$, such that for every vertex u with $|u| < |v|$, every edge obtained by decreasing a coordinate of u has been added. Let G' be the union of $E(G)$ and the edges that have been added so far. Let y be a neighbour of v such that $|y| = |v| - 1$ and y and v differ on a large coordinate of v . Without loss of generality, assume that v and y differ on the first coordinate. As $L(v) \geq m$, we have that G' contains $m - 1$ edges wv with $|w| = |v| - 1$ such that the coordinate on which w and v differ is large in v . Without loss of generality, assume that each of these edges are obtained by changing the 2nd to m th coordinates of v respectively. Now consider the m -dimensional r -grid in $[k]^d$ induced by the vertices

$$\{v_1 - (r - 1), \dots, v_1\} \times \dots \times \{v_m - (r - 1), \dots, v_m\} \times \{(v_{m+1}, \dots, v_d)\}.$$

By our inductive hypothesis, every edge of this grid is in G' apart from vy . Thus adding vy completes a new copy of $[r]^m$. This completes the proof of the upper bound.

For the lower bound, we apply Lemma 3.2 where \mathcal{H} consists of all axis aligned copies of $[r]^m$ in $[k]^d$. So, it suffices to show that there exists a vector space W and a set $\{f_e : e \in E([k]^d)\} \subseteq W$ which satisfy the hypotheses of Lemma 3.2 such that

$$\dim \left(\text{span} \left\{ f_e : e \in E([k]^d) \right\} \right) \geq |E(G)|. \quad (3.3)$$

The space W that we choose is the direct sum of k^d copies of \mathbb{R}^{m-1} , one for each vertex of $[k]^d$, and dk^{d-1} copies of \mathbb{R}^{r-2} , one for each line in $[k]^d$. That is,

$$W := \left(\bigoplus_x \mathbb{R}^{m-1} \right) \oplus \left(\bigoplus_{\mathcal{L}} \mathbb{R}^{r-2} \right)$$

where x ranges over all vertices of $[k]^d$ and \mathcal{L} ranges over all lines in $[k]^d$. Given a vector w of W , a vertex x of $[k]^d$ and a line \mathcal{L} of $[k]^d$, let $\pi_x(w)$ denote the projection of w onto the

copy of \mathbb{R}^{m-1} corresponding to x and let $\pi_{\mathcal{L}}(w)$ denote the projection of w onto the copy of \mathbb{R}^{r-2} corresponding to \mathcal{L} . Note that w is determined by its projections.

For $n \geq 1$, a collection $Z \subseteq \mathbb{R}^n$ is in *general position* in \mathbb{R}^n if any set of at most n vectors from Z is linearly independent. Let $Z := \{z_1, \dots, z_d\}$ be a collection of d vectors of \mathbb{R}^{m-1} in general position. Also, let $Y := \{y_1, \dots, y_{k-2}\}$ be a set of $k-2$ vectors of \mathbb{R}^{r-2} such that y_1, \dots, y_{r-2} are linearly independent,

$$y_{r-1} := - \sum_{j=1}^{r-2} y_j, \quad (3.4)$$

and for $i \geq r$, $y_i = y_{i-r+1}$. Thus any consecutive $r-2$ vectors y_i, \dots, y_{i+r-3} are linearly independent and any consecutive $r-1$ vectors y_i, \dots, y_{i+r-2} sum to zero. Suppose that $e = uv$ is an edge of $[k]^d$ such that u and v differ on coordinate $i \in [d]$ and let t be the maximum of the i th coordinates of u and v . Further, let \mathcal{L} be the unique line of $[k]^d$ containing e . We define f_e to be the vector of W such that

- $\pi_u(f_e) = \pi_v(f_e) = z_i$ and $\pi_x(f_e) = 0$ for every $x \in V([k]^d) \setminus \{u, v\}$, and
- $\pi_{\mathcal{L}} = y_t$ and $\pi_{\mathcal{L}'} = 0$ for every line $\mathcal{L}' \neq \mathcal{L}$.

In order to apply Lemma 3.2, we need to show that for every axis aligned copy P of $[r]^m$ in $[k]^d$ there are non-zero scalars $\{c_e : e \in E(P)\}$ such that $\sum_{e \in E(P)} c_e f_e = 0$. Let P be an axis aligned copy of $[r]^m$ in $[k]^d$ and let $I \subseteq [d]$ be the set of m coordinates which vary under P . As the vectors of Z are chosen to be in general position in \mathbb{R}^{m-1} , any subset of Z with cardinality m is minimally linearly dependent. Therefore there exists a set of non-zero scalars $\{c_i : i \in I\}$ such that $\sum_{i \in I} c_i z_i = 0$.

For each vertex v of P , let $M(v) \subseteq [d]$ be the set of indices j such that both $v - e_j$ and $v + e_j$ are contained in P . Define the *lines* of P to be the r -vertex paths in P obtained by taking the intersection of a line of $[k]^d$ with $V(P)$. Note that $|M(v)|$ is precisely the number of lines of P in which v has degree two. For each line \mathcal{L} of P , define $m(\mathcal{L}) := |M(v)|$ where v is an endpoint of \mathcal{L} (clearly the endpoints give the same value). For each $i \in I$, let E_i be the set of all edges of P for which i is the variable coordinate. For each edge $e \in E_i$ contained in a line \mathcal{L} of P , define $d_e := 2^{m(\mathcal{L})} c_i$. We claim that $\sum_{e \in E(P)} d_e f_e = 0$.

First, let \mathcal{L} be a line of P in direction i . Then, for some $t \geq r$, we have

$$\pi_{\mathcal{L}} \left(\sum_{e \in E(P)} d_e f_e \right) = \sum_{e \in E(P)} d_e \pi_{\mathcal{L}}(f_e) = \sum_{e \in E(\mathcal{L})} 2^{m(\mathcal{L})} c_i \pi_{\mathcal{L}}(f_e) = 2^{m(\mathcal{L})} c_i \sum_{j=t-r+2}^t y_j.$$

However, this sum is equal to zero by (3.4).

Now, fix a vertex v of P and for each $i \in I$ let \mathcal{L}_i be the line of P in direction i containing v . We have

$$\begin{aligned} \pi_v \left(\sum_{e \in E(P)} d_e f_e \right) &= \sum_{e \in E(P)} d_e \pi_v(f_e) \\ &= \sum_{i \in M(v)} \sum_{e \in E_i} 2^{m(\mathcal{L}_i)} c_i \pi_v(f_e) + \sum_{i \in I \setminus M(v)} \sum_{e \in E_i} 2^{m(\mathcal{L}_i)} c_i \pi_v(f_e) \end{aligned}$$

which is equal to

$$\sum_{i \in M(v)} 2^{m(\mathcal{L}_i)+1} c_i \pi_v(f_e) + \sum_{i \in I \setminus M(v)} 2^{m(\mathcal{L}_i)} c_i \pi_v(f_e)$$

by definition of $M(v)$. We observe that, for any $i \in M(v)$, we have $m(\mathcal{L}_i) = |M(v)| - 1$ and for any $i \in I \setminus M(v)$ we have $m(\mathcal{L}_i) = |M(v)|$. Therefore, the above sum is equal to $2^{|M(v)|} \sum_{i \in I} c_i z_i$ which is zero by definition of $\{c_i : i \in I\}$. Combining this with the result of the previous paragraph, we get $\sum_{e \in E(P)} d_e f_e = 0$, as desired.

To complete the proof, it suffices to prove (3.3). To do so, we let G be a graph as in the proof of the upper bound and show that the vectors of $\{f_e : e \in E(G)\}$ are linearly independent. Let $\{c_e : e \in E(G)\}$ be any set of scalars, not all of which are zero, and let F be the spanning subgraph of G containing all edges $e \in E(G)$ such that $c_e \neq 0$. Let v be a vertex of non-zero degree in F such that $|v|$ is maximum.

First, consider the case that v has degree at most $m - 1$ in F . Let J denote the set of $d_F(v)$ coordinates such that, for each $j \in J$, the edge from v to $v - e_j$ is present in F . Then

$$\sum_{e \in E(G)} c_e \pi_v(f_e) = \sum_{j \in J} c_j z_j,$$

for some set of non-zero scalars $\{c_j : j \in J\}$. Since $|J| \leq m - 1$ and the vectors of Z are in general position, this sum is non-zero and therefore $\sum_{e \in E(G)} c_e f_e \neq 0$.

Now, suppose that v has degree at least m in F . By construction of G , this implies that there is a coordinate j which is small for v such that the edge e from v to $v - e_j$ is present in F . Let \mathcal{L} be the unique line of $[k]^d$ containing e . Then, by maximality of $|v|$, every edge of \mathcal{L} contained in F joins two vertices for which j is a small coordinate. This implies that

$$\sum_{e \in E(F)} c_e \pi_{\mathcal{L}}(f_e) = \sum_{i=2}^{r-1} c_i y_i,$$

for some set of non-zero scalars $\{c_i : 2 \leq i \leq r - 1\}$. Thus, since the vectors y_2, \dots, y_{r-1} are linearly independent, this sum is non-zero and we obtain $\sum_{e \in E(G)} c_e f_e \neq 0$, which completes the proof. \square

Remark 3.5. The proof of Theorem 3.1 implies that, for most values of $k \geq r \geq 2$ and $d \geq m \geq 1$, there are many tight examples. For an example in two dimensions, see Figure 3.2.

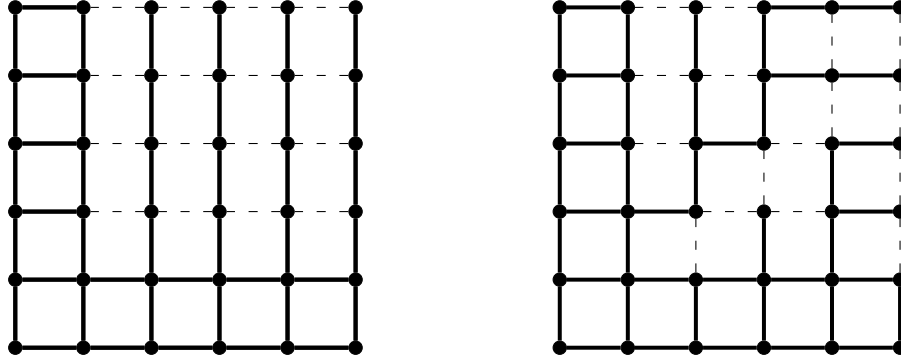


Figure 3.2: Two non-isomorphic weakly $([6]^2, [3]^2)$ -saturated graphs with $\text{wsat}([6]^2, [3]^2)$ edges.

3.4 Open Problems

Many interesting open problems remain, and we mention some of them here.

3.4.1 Minimum Saturation

Recall that we were able to prove a better upper bound on $\text{sat}(Q_d, Q_m)$ when d is of the form $6m(2^t - 1)$ than for general values of d (see Remark 2.17). It is not clear whether this is simply an artifact of our proof, or part of a more general phenomenon. We ask the following.

Question 4.1. For fixed $m \geq 2$, does $\lim_{d \rightarrow \infty} \frac{\text{sat}(Q_d, Q_m)}{2^d}$ exist?

An answer to this would resolve a particular case of the natural analogous conjecture to Conjecture 2.3, where K_n is replaced by Q_d . As it stands, the best known lower bound on $\text{sat}(Q_d, Q_m)$ is $(m - 1 - o(1))2^d$, which follows from Theorem 1.5 and (1.1). We doubt that this bound is tight. We ask the following.

Question 4.2. For fixed $m \geq 2$, let $c_m := \liminf_{d \rightarrow \infty} \text{sat}(Q_d, Q_m)/2^d$. Does $\frac{c_m}{m} \rightarrow \infty$?

It would also be interesting to study saturation numbers in the grid.

Question 4.3. Fix $m \geq 2$, $r \geq 2$ and $k \geq \max\{r, 3\}$, and let $d \rightarrow \infty$. Is it true that

$$\text{sat}^*([k]^d, [r]^m) = o(|E([k]^d)|)?$$

How about

$$\text{sat}^*([k]^d, [r]^m) = O(k^d)?$$

Here, the notation sat^* indicates that we are considering axis aligned copies of $[r]^m$ in $[k]^d$.

Recall that K_k^d is the graph with vertex set $[k]^d$ in which two vertices are adjacent if they differ (by any amount) in exactly one coordinate. As mentioned above, $\text{sat}(K_k, K_r)$ has been determined for all values of k and r . One could also consider the following multidimensional analogue.

Problem 4.4. For $k \geq r \geq 2$ and $d \geq m \geq 1$, determine $\text{sat}(K_k^d, K_r^m)$.

The problem of determining $\text{sat}(K_k, C_\ell)$ was raised by Bollobás [30] and has been extensively studied. The case $\ell = 3$ is trivial, and the cases $\ell = 4$ and $\ell = 5$ were solved by Ollmann [108] and Chen [39, 40], respectively. The first non-trivial bounds for general ℓ were given by Barefoot, Clark, Entringer, Porter, Székely, and Tuza [20], who proved that there are positive constants c_1 and c_2 such that, for $\ell \neq 8, 10$,

$$(1 + c_1/\ell)n \leq \text{sat}(K_k, C_\ell) \leq (1 + c_2/\ell)n.$$

The value of c_2 was improved by Gould, Łuczak and Schmitt [64]. Currently, the best general upper and lower bounds are due to Füredi and Kim [60]. It is natural to ask about cycles in the hypercube.

Problem 4.5. For $\ell \geq 2$ and $d \geq \log_2(2\ell)$, determine $\text{sat}(Q_d, C_{2\ell})$.

More generally, similar questions can be asked for cycles in $[k]^d$ or K_k^d . We remark that the related problem of determining $\text{ex}(Q_d, C_{2\ell})$ was proposed by Erdős [47] and is very well studied; see, e.g., [9, 17, 34, 42–44, 61, 62, 122].

3.4.2 Weak Saturation

As mentioned above, it has been proved that $\text{wsat}(K_k, K_r) = \text{sat}(K_k, K_r)$ for all k and r . A multidimensional version of this problem also seems interesting.

Problem 4.6. For general $1 \leq m \leq d$ and $2 \leq r \leq k$, determine $\text{wsat}(K_k^d, K_r^m)$.

As mentioned in Section 3.3, the ‘vertex version’ of Problem 4.6 was solved in [16].

3.4.3 Semi-saturation

Johnson and Pinto [80] also study another type of saturation problem, which is often referred to as semi-saturation [45, 60] (other terms have also been used: see [30, 81, 103, 110, 127]). Given graphs F and H , say that a spanning subgraph G of F is (F, H) -semi-saturated if for every edge $e \in E(F) \setminus E(G)$, the graph $G + e$ contains more copies of H than G does. Note that G may contain H as a subgraph. Let $\text{ssat}(F, H)$ denote the *semi-saturation number*, which is the minimum number of edges in a (F, H) -semi-saturated graph. Then it is easy to see that

$$\text{wsat}(F, H) \leq \text{ssat}(F, H) \leq \text{sat}(F, H).$$

Johnson and Pinto [80] proved that $\text{ssat}(Q_d, Q_m) \leq (m^2 + m/2) 2^d$. Based on this result and (1.3), they asked whether $\text{ssat}(Q_d, Q_2)$ is equal to $\text{sat}(Q_d, Q_2)$ for sufficiently large d and, if not, whether $\limsup_{d \rightarrow \infty} \frac{\text{sat}(Q_d, Q_2)}{2^d} > \limsup_{d \rightarrow \infty} \frac{\text{ssat}(Q_d, Q_2)}{2^d}$. Given Theorem 1.4, it now seems natural to pose these questions more generally.

Question 4.7. *For fixed $m \geq 2$, is $\text{ssat}(Q_d, Q_m)$ equal to $\text{sat}(Q_d, Q_m)$ for sufficiently large d ? If not, is $\limsup_{d \rightarrow \infty} \frac{\text{sat}(Q_d, Q_m)}{2^d} > \limsup_{d \rightarrow \infty} \frac{\text{ssat}(Q_d, Q_m)}{2^d}$?*

Of course, the open problems on saturation mentioned above also make sense for semi-saturation.

Chapter 4

Extremal Bounds for Bootstrap Percolation in the Hypercube

Joint work with Jonathan A. Noel

4.1 Introduction

A *cellular automaton* is a dynamic process on a discrete set of ‘sites’ in which each site is initially assigned to one of a finite number of ‘states’ and, at each time step, the states are updated according to a deterministic rule depending on the current states of nearby sites. The notion of a cellular automaton was formally introduced by von Neumann [131] but grew out of his discussions with Ulam during the 1940’s and 50’s (see [130] or [120]). The focus of this chapter is on a well studied example of a simple cellular automaton known as the *r-neighbour bootstrap process*, introduced by Chalupa, Leath and Reich [38].

Recall the following definition, given in Section 1.3 of Chapter 1. Given a positive integer r and a graph G , the *r-neighbour bootstrap process* begins with an initial set of *infected* vertices of G and, at each step of the process, a vertex becomes infected if it has at least r infected neighbours. More formally, if A_0 is the initial set of infected vertices, then the set of vertices that are infected after the j th step of the process for $j \geq 1$ is defined by

$$A_j := A_{j-1} \cup \{v \in V(G) : |N_G(v) \cap A_{j-1}| \geq r\},$$

where $N_G(v)$ denotes the neighbourhood of v in G . We say that A_0 *percolates* if $\bigcup_{j=0}^{\infty} A_j = V(G)$.

The *r-neighbour bootstrap process* was introduced as a mathematical simplification of existing dynamic models of ferromagnetism [38]. Since then, it has been used to model many other complex processes, such as the spread of opinion through a social network [65, 134] and the impact of defaults in banking systems [5]. It also been used in the study of crack

formation, computer storage arrays and hydrogen mixtures (see the survey [1] for more details on these and other applications). In addition, advances in bootstrap percolation have been highly influential in the study of more complicated processes including, for example, the Glauber dynamics of the Ising model [56, 101]. In recent years, there has been a surge of interest in the behaviour of the r -neighbour bootstrap process on various distributions of random graphs motivated by potential applications in understanding propagation phenomena in real-world networks; see, for example, [4, 6, 7, 19, 32, 57, 70, 78, 79, 83].

The most well studied problem about the r -neighbour bootstrap process is to estimate the *critical probability* at which a randomly generated set of vertices in a graph G becomes likely to percolate. To be more precise, for $p \in [0, 1]$, suppose that A_0 is a subset of $V(G)$ obtained by including each vertex randomly with probability p independently of all other vertices and define

$$p_c(G, r) := \inf \{p : \mathbb{P}(A_0 \text{ percolates}) \geq 1/2\}.$$

The problem of estimating $p_c([n]^d, r)$ for fixed d and r and $n \rightarrow \infty$ was first considered by Aizenman and Lebowitz [2] and subsequently studied in [12, 36, 37, 66, 76]. This rewarding line of research culminated in a paper of Balogh, Bollobás, Duminil-Copin and Morris [11] in which $p_c([n]^d, r)$ is determined asymptotically for all fixed values of d and $2 \leq r \leq d$.

Comparably, far less is known about the critical probability when d tends to infinity. In this regime, the main results are due to Balogh, Bollobás and Morris in the case $r = d$ [13] and $r = 2$ [14]. In the latter paper, the extremal bound (1.7) below was applied to obtain precise asymptotics for $p_c([n]^d, 2)$ whenever $d \gg \log(n) \geq 1$. In contrast, very little is known about the critical probability for fixed $r \geq 3$ and $d \rightarrow \infty$. For example, the logarithm of $p_c(Q_d, 3)$ is not even known to within a constant factor (see [14]). As was mentioned in [16], motivation for studying $m(Q_d, r)$ stems from its potential applications to obtaining good estimates for $p_c(Q_d, r)$ when $d \rightarrow \infty$. One stumbling block in determining this has been the lack of an asymptotically tight lower bound $m(Q_d, r)$ for $r \geq 3$.

In the case $r = 2$ the value of $m(Q_d, r)$ was determined completely.

Theorem 1.1 (Balogh, Bollobás [10]).

$$m(Q_d, 2) = \left\lceil \frac{d}{2} \right\rceil + 1.$$

They also made the following conjecture (see also [14, 16]): For fixed $r \geq 3$ and $d \rightarrow \infty$,

$$m(Q_d, r) = \frac{1 + o(1)}{r} \binom{d}{r-1}.$$

As described in equation (3.2) of Chapter 1, an upper bound of

$$m(Q_d, r) \leq \frac{d^{r-1}}{r!} + \frac{d^{r-2}(r+2)}{2r(r-2)!} + O(d^{r-3}). \quad (1.2)$$

follows from a construction based on Steiner systems. The previous best known lower bound, obtained using Theorem 1.1 was linear in n .

In this chapter, we prove the conjecture of Balogh and Bollobás.

Theorem 1.3. *For $d \geq r \geq 1$,*

$$m(Q_d, r) \geq 2^{r-1} + \sum_{j=1}^{r-1} \binom{d-j-1}{r-j} \frac{j2^{j-1}}{r}$$

where, by convention, $\binom{a}{b} = 0$ when $a < b$.

This theorem provides a lower bound that matches (1.2) in the first order term. In order to prove Theorem 1.3, we will exploit a relationship between bootstrap percolation and weak saturation. Our proof of Theorem 1.3 relies on the following bound:

$$m(G, r) \geq \frac{\text{wsat}(G, S_{r+1})}{r} \quad (1.4)$$

where S_{r+1} denotes the star with $r+1$ leaves. A slightly stronger version of (1.4) is stated and proved in the next section. We obtain an exact expression for the weak saturation number of S_{r+1} in the d -dimensional hypercube Q_d .

Theorem 1.5. *If $d \geq r \geq 0$, then*

$$\text{wsat}(Q_d, S_{r+1}) = r2^{r-1} + \sum_{j=1}^{r-1} \binom{d-j-1}{r-j} j2^{j-1}.$$

Note that when $d > r$, this is equivalent to writing

$$\text{wsat}(Q_d, S_{r+1}) = \sum_{j=1}^r \binom{d-j-1}{r-j} j2^{j-1}.$$

Theorem 1.3 follows directly from Theorem 1.5 and (1.4).

Thinking beyond the hypercube, for $d \geq r \geq 1$ we also determine an exact expression for the weak saturation number of S_{r+1} in the d -dimensional $a_1 \times \cdots \times a_d$ grid, denoted by $\prod_{i=1}^d [a_i]$. (The d -dimensional k -grid $[k]^d$ introduced in Chapter 3 is the specific case where $a_1 = a_2 = \cdots = a_d = k$.)

Theorem 1.6. *Let $G := \prod_{i=1}^d [a_i]$. For $d \geq r \geq 0$ and $a_1, \dots, a_d \geq 2$,*

$$\text{wsat}(G, S_{r+1}) = \sum_{\substack{S \subseteq [d] \\ |S| \leq r-1}} \left(\prod_{i \in S} (a_i - 2) \right) \left((r - |S|)2^{r-|S|-1} + \sum_{j=1}^{r-|S|-1} \binom{d-|S|-j-1}{r-|S|-j} j2^{j-1} \right).$$

Observe that a lower bound on $m\left(\prod_{i=1}^d [a_i], r\right)$ follows from Theorem 1.6 and (1.4). To our knowledge, the combination of Theorem 1.6 and (1.4) implies all of the known lower bounds on the cardinality of percolating sets in multidimensional grids. In particular, it implies the (tight) lower bounds

$$m\left([n]^d, d\right) \geq n^{d-1},$$

and

$$m\left(\prod_{i=1}^d [a_i], 2\right) \geq \left\lceil \frac{\sum_{i=1}^d (a_i - 1)}{2} \right\rceil + 1. \quad (1.7)$$

established in [109] and [10], respectively.

Concerning upper bounds, we give a recursive upper bound on $m(Q_d, r)$, which improves on the second order term of (1.2). In the case $r = 3$, we use this to determine the exact value of $m(Q_d, r)$.

Theorem 1.8. *For $d \geq 3$, we have $m(Q_d, 3) = \left\lceil \frac{d(d+3)}{6} \right\rceil + 1$.*

As well as considering the minimum cardinality of a percolating set, a number of other extremal results have been established about the 2-neighbour process. The maximum size of a minimal set that percolates under the 2-neighbour process has been studied in $[n]^2$ by Morris [100] and in the hypercube by Riedl [117]. The maximum time that a set takes to percolate under the 2-neighbour process has also been studied in the hypercube by Przykucki [115] and in $[n]^2$ by Benevides and Przykucki [21, 22]. Also, Shapiro and Stevens [121] studied the structure of the sets of minimum size that percolate under the 2-neighbour process on $[n]^2$.

The rest of the chapter is organised as follows. In the next section, we outline our approach to proving Theorems 1.3 and 1.6 and establish a preliminary lemma. In Section 4.3, we determine the value of $\text{wsat}(Q_d, S_{r+1})$, from which we derive Theorem 1.3 via (1.4). We then determine $\text{wsat}\left(\prod_{i=1}^d [a_i], S_{r+1}\right)$ in full generality in Section 4.4 using similar ideas (which become somewhat more cumbersome in the general setting). In Section 4.5, we provide constructions of small percolating sets in the hypercube and prove Theorem 1.8. In Section 4.6, for interest, we give an explicit construction that provides an alternate proof of Lemma 3.3. Finally, we conclude the chapter in Section 4.7 by stating some open problems related to our work.

4.2 Preliminaries and Proof Outline

As mentioned in Chapter 1, the previous best known lower bound on $m(Q_d, r)$ was only linear in n . Therefore, even in the case $r = 3$, there was a large gap between the quadratic

upper bound described in Chapter 1 and this lower bound. Here we describe how we prove our better lower bound.

The key idea behind the proof of Theorem 1.3 is to study a related process on the edges on the hypercube, rather than to directly analyse the vertex process. Indeed, in this section we will prove a lemma (Lemma 2.1) that implies (1.4), which gives a direct relationship between the r -neighbour bootstrap percolation process and a weak-saturation process. In particular, to find a lower bound on $m(Q_d, r)$, it suffices to find a lower bound for $\text{wsat}(Q_d, S_{r+1})$.

Given this relationship, the proof will proceed by finding a good lower bound for $\text{wsat}(Q_d, S_{r+1})$. In fact, we will determine $\text{wsat}(Q_d, S_{r+1})$ exactly (Theorem 1.5). To do this, as in Chapter 3 we must prove matching upper and lower bounds on $\text{wsat}(Q_d, S_{r+1})$. Recall that to prove an upper bound on a weak saturation number, one only needs to construct a *single* example of a weakly saturated graph of small size and this is what we will do. As in Chapter 3, our main tool for proving the lower bound is the linear algebraic lemma of Balogh, Bollobás, Morris and Riordan [16] of Chapter 3.

We now prove the following lemma, which improves on (1.4) for graphs with vertices of degree less than r (including, for example, the graph $\prod_{i=1}^d [a_i]$ for $d < r$).

Lemma 2.1. *Let G be a graph and let F be a spanning subgraph of G . Define*

$$A_F := \{v \in V(G) : d_F(v) \geq \min\{r, d_G(v)\}\}.$$

If A_F percolates under the r -neighbour bootstrap process on G , then F is weakly (G, S_{r+1}) -saturated.

Proof. By hypothesis, we can label the vertices of G by v_1, \dots, v_n in such a way that

- $\{v_1, \dots, v_{|A_F|}\} = A_F$, and
- for $|A_F| + 1 \leq i \leq n$, the vertex v_i has at least r neighbours in $\{v_1, \dots, v_{i-1}\}$.

Let us show that F is weakly (G, S_{r+1}) -saturated. We begin by adding to F every edge of $E(G) \setminus E(F)$ which is incident to a vertex in A_F (one edge at a time in an arbitrary order). For every vertex $v \in A_F$, we have that either

- there are at least r edges of F incident to v , or
- every edge of G incident with v is already present in F .

Therefore, every edge of $E(G) \setminus E(F)$ incident to a vertex in A_F completes a copy of S_{r+1} when it is added.

Now, for each $i = |A_F| + 1, \dots, n$ in turn, we add every edge incident to v_i which has not already been added (one edge at a time in an arbitrary order). Since v_i has at least r neighbours in $\{v_1, \dots, v_{i-1}\}$ and every edge incident to a vertex in $\{v_1, \dots, v_{i-1}\}$ is already present, we get that every such edge completes a copy of S_{r+1} when it is added. The result follows. \square

For completeness, we will now deduce (1.4) from Lemma 2.1.

Proof of (1.4). Let A_0 be a set of cardinality $m(G, r)$ which percolates with respect to the r -neighbour bootstrap process on G and let F be a spanning subgraph of G such that $d_F(v) \geq \min\{d_G(v), r\}$ for each $v \in A_0$. Note that this can be achieved by adding at most r edges per vertex of A_0 and so we can assume that $|E(F)| \leq r|A_0| = rm(G, r)$. By Lemma 2.1, F is weakly (G, S_{r+1}) -saturated and so

$$\text{wsat}(G, S_{r+1}) \leq |E(F)| \leq rm(G, r)$$

as required. \square

We turn our attention to determining the weak saturation number of stars in hypercubes and, more generally, in multidimensional rectangular grids.

4.3 The Hypercube Case

Our goal in this section is to prove Theorem 1.5. This proves the case $a_1 = \dots = a_d = 2$ of Theorem 1.6 and, as discussed earlier, implies Theorem 1.3 via (1.4). First, we require some definitions.

Definition 3.1. Given $k \geq 1$, an index $i \in [k]$ and $x \in \mathbb{R}^k$, let x_i denote the i th coordinate of x . The *support* of x is defined by $\text{supp}(x) := \{i \in [k] : x_i \neq 0\}$.

Definition 3.2. The *direction* of an edge $e = uv \in E(Q_d)$ is the unique index $i \in [d]$ such that $u_i \neq v_i$. Given a vertex $v \in V(Q_d)$, we define $e(v, i)$ to be the unique edge in direction i that is incident to v . We say that $e(v, i)$ is a *label* for e .

Note that each edge of Q_d receives two labels (one for each of its endpoints). Our approach will make use of the following simple linear algebraic fact.

Lemma 3.3. *Let $k \geq \ell \geq 0$ be fixed. Then there exists a subspace X of \mathbb{R}^k of dimension $k - \ell$ such that $|\text{supp}(x)| \geq \ell + 1$ for every $x \in X \setminus \{0\}$.*

Proof. Define X to be the span of a set $\{v_1, \dots, v_{k-\ell}\}$ of unit vectors of \mathbb{R}^k chosen independently and uniformly at random with respect to the standard Lebesgue measure on the unit sphere S^{k-1} . Given a fixed subspace W of \mathbb{R}^k of dimension at most ℓ and $1 \leq i \leq k-\ell$, the space

$$\text{span}(W \cup \{v_1, \dots, v_{i-1}\})$$

has dimension less than k . Thus, the unit sphere of this space has measure zero in S^{k-1} and so, with probability one, $v_i \notin \text{span}(W \cup \{v_1, \dots, v_{i-1}\})$. It follows that $\dim(X) = k-\ell$ and $X \cap W = \{0\}$ almost surely. In particular, if we let $T \subseteq [k]$ be a fixed set of cardinality ℓ and define

$$W_T := \left\{ x \in \mathbb{R}^k : \text{supp}(x) \subseteq T \right\},$$

then $X \cap W_T = \{0\}$ almost surely. Since there are only finitely many sets $T \subseteq [k]$ of cardinality ℓ , we can assume that X is chosen so that $X \cap W_T = \{0\}$ for every such set. This completes the proof. \square

In section 4.6, we provide an explicit (ie. non-probabilistic) example of a vector space X satisfying Lemma 3.3. We defer the construction to this section as it is of tangential interest to the main proof. The following lemma highlights an important property of the space X .

Lemma 3.4. *Let $k \geq \ell \geq 0$ and let X be a subspace of \mathbb{R}^k of dimension $k-\ell$ such that $|\text{supp}(x)| \geq \ell+1$ for every $x \in X \setminus \{0\}$. For every set $T \subseteq [k]$ of cardinality $\ell+1$, there exists $x \in X$ with $\text{supp}(x) = T$.*

Proof. Let $T \subseteq [k]$ with $|T| = \ell+1$. The space $\{x \in \mathbb{R}^k : \text{supp}(x) \subseteq T\}$ has dimension $\ell+1$. Therefore, since $\dim(X) = k-\ell$, there must be a non-zero vector $x \in X$ with $\text{supp}(x) \subseteq T$. However, this inclusion must be equality since $|\text{supp}(x)| \geq \ell+1$. \square

We are now in position to prove Theorem 1.5. For notational convenience, we write

$$w = w_{d,r} := r2^{r-1} + \sum_{j=1}^{r-1} \binom{d-j-1}{r-j} j2^{j-1}.$$

For $d > r \geq 1$, we have

$$w_{d,r} = w_{d-1,r} + w_{d-1,r-1}. \tag{3.5}$$

We deduce Theorem 1.5 from the following lemma, after which we will prove the lemma itself.

Lemma 3.6. *For $d \geq r \geq 0$, let X be a subspace of \mathbb{R}^d of dimension $d-r$ such that $|\text{supp}(x)| \geq r+1$ for every $x \in X \setminus \{0\}$. Then there is a spanning subgraph F of Q_d and a collection $\{f_e : e \in E(Q_d)\} \subseteq \mathbb{R}^w$ such that*

(P1) F is weakly (Q_d, S_{r+1}) -saturated and $|E(F)| = w$,

(P2) $\sum_{i=1}^d x_i f_{e(v,i)} = 0$ for every $v \in V(Q_d)$ and $x \in X$, and

(P3) $\text{span}\{f_e : e \in E(Q_d)\} = \mathbb{R}^w$.

We remark that such a space X must exist by Lemma 3.3.

Proof of Theorem 1.5. For each $d \geq r \geq 0$, we will show that $\text{wsat}(Q_d, S_{r+1}) = w$. Clearly, the existence of a graph F satisfying (P1) implies the upper bound $\text{wsat}(Q_d, S_{r+1}) \leq w$. We show that the lower bound follows from (P2), (P3) and Lemma 3.2. Note that the edge sets of copies of S_{r+1} in Q_d are precisely the sets of the form $\{e(v, i) : i \in T\}$ where v is a fixed vertex of Q_d and T is a subset of $[d]$ of cardinality $r + 1$. By Lemma 3.4 we know that there exists some $x \in X$ with $\text{supp}(x) = T$. By (P2) we have

$$\sum_{i=1}^d x_i f_{e(v,i)} = \sum_{i \in T} x_i f_{e(v,i)} = 0.$$

Therefore, by Lemma 3.2 we have

$$\text{wsat}(Q_d, S_{r+1}) \geq \dim(\text{span}\{f_e : e \in E(Q_d)\}),$$

which equals w by (P3). The result follows. \square

The proof of Lemma 3.6 goes by induction on d . We first settle some easy base cases constructively. In the inductive step, we partition Q_d into two copies of Q_{d-1} , Q_{d-1}^0 and Q_{d-1}^1 . On Q_{d-1}^0 we apply our inductive hypothesis with $d - 1$ and r , and on Q_{d-1}^1 we apply our inductive hypothesis with $d - 1$ and $r - 1$. Using the weakly saturated graphs and vectors obtained on the copies of Q_{d-1} by induction, we are able to construct a graph F and vectors satisfying the required hypotheses for the d case.

Proof of Lemma 3.6. We proceed by induction on d . We begin by settling some easy boundary cases before explaining the inductive step.

Case 1: $r = 0$.

In this case, $S_{r+1} \simeq K_2$. Also, $w = 0$ and $X = \mathbb{R}^d$. We let F be a spanning subgraph of Q_d with no edges and set $f_e := 0$ for every $e \in Q_d$. It is not difficult to see that (P1), (P2) and (P3) are satisfied.

Case 2: $d = r \geq 1$.

In this case, $w = d2^{d-1} = |E(Q_d)|$ and $X = \{0\}$. We define $F := Q_d$ and let $\{f_e : e \in E(Q_d)\}$ be a basis for \mathbb{R}^w . Clearly (P1), (P2) and (P3) are satisfied.

Case 3: $d > r \geq 1$.

Observe that our induction hypothesis combined with the proof of Theorem 1.5 gives

$$\text{wsat}(Q_{d-1}, S_{r+1}) = w_{d-1,r} \quad (3.7)$$

for all $d-1 \geq r \geq 0$. We begin by constructing F in such a way that (P1) is satisfied. For $i \in \{0, 1\}$, let Q_{d-1}^i denote the subgraph of Q_d induced by $\{0, 1\}^{d-1} \times \{i\}$. Note that both Q_{d-1}^0 and Q_{d-1}^1 are isomorphic to Q_{d-1} . By (3.7), we know that $\text{wsat}(Q_{d-1}, S_{r+1}) = w_{d-1,r}$ and $\text{wsat}(Q_{d-1}, S_r) = w_{d-1,r-1}$. Thus we choose F to be a spanning subgraph of Q_d such that

- the subgraph F_0 of F induced by $V(Q_{d-1}^0)$ is a weakly (Q_{d-1}, S_{r+1}) -saturated graph with $|E(F_0)| = w_{d-1,r}$,
- the subgraph F_1 of F induced by $V(Q_{d-1}^1)$ is a weakly (Q_{d-1}, S_r) -saturated graph with $|E(F_1)| = w_{d-1,r-1}$, and
- F contains no edge in direction d .

See Figure 4.1 for a specific example of this construction. By construction, we have

$$|E(F)| = w_{d-1,r} + w_{d-1,r-1},$$

which is equal to w by (3.5).

Let us verify that F is weakly (Q_d, S_{r+1}) -saturated. To see this, we add the edges of $E(Q_d) \setminus E(F)$ to F in three stages. By construction, we can begin by adding all edges of Q_{d-1}^0 which are not present in F_0 in such a way that each edge completes a copy of S_{r+1} in Q_{d-1}^0 when it is added. In the second stage, we add all edges of Q_d in direction d one by one in any order. Since every vertex of Q_d has degree $d \geq r+1$ and every edge of Q_{d-1}^0 has already been added, we get that every edge added in this stage completes a copy of S_{r+1} in Q_d . Finally, we add the edges of Q_{d-1}^1 which are not present in F_1 in such a way that each added edge completes a copy of S_r in Q_{d-1}^1 . Since the edges in direction d have already been added, we see that every such edge completes a copy of S_{r+1} in Q_d . Therefore, (P1) holds.

Thus, all that remains is to construct $\{f_e : e \in E(Q_d)\}$ in such a way that (P2) and (P3) are satisfied. For ease of notation, set $w_0 := w_{d-1,r}$ and $w_1 := w_{d-1,r-1}$. Let $\pi : X \rightarrow \mathbb{R}^{d-1}$ be the standard projection defined by $\pi : (x_1, \dots, x_d) \mapsto (x_1, \dots, x_{d-1})$. Let $z \in X$ be an arbitrary vector such that $d \in \text{supp}(z)$ (such a vector exists by Lemma 3.4) and let $T_z : X \rightarrow X$ be the linear map defined by

$$T_z(x) := x - \frac{x_d}{z_d} z$$

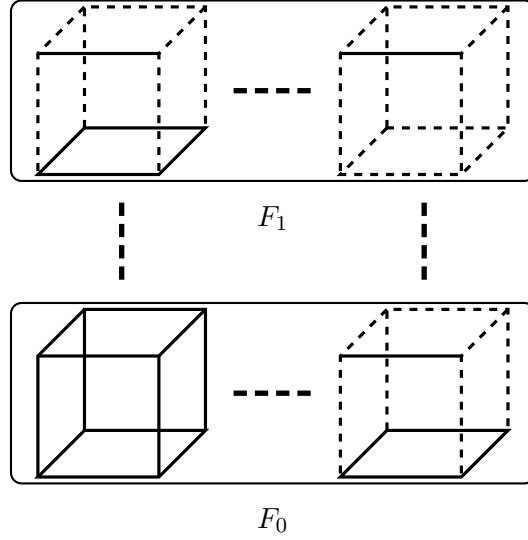


Figure 4.1: A weakly (Q_5, S_4) -saturated graph F constructed inductively from a weakly (Q_4, S_4) -saturated graph F_0 and a weakly (Q_4, S_3) -saturated graph F_1 , each of which is also constructed inductively.

for $x \in X$. Note that $T_z(x)$ has d th coordinate equal to zero. This property of T_z will be important later. Define

$$X_0 := \pi(T_z(X)) \quad \text{and}$$

$$X_1 := \pi(X).$$

Clearly, $\ker(T_z) = \text{span}\{z\}$ and, since every $x \in X \setminus \{0\}$ has $|\text{supp}(x)| \geq r+1 \geq 2$, we have $\ker(\pi) = \{0\}$. This implies that X_0 has dimension $d-r-1$ and that X_1 has dimension $d-r$. Also, by construction, we have that $|\text{supp}(x)| \geq r+1$ for every non-zero $x \in X_0$ and $|\text{supp}(x)| \geq r$ for every non-zero $x \in X_1$.

Therefore, by the inductive hypothesis, there exists $\{f_e^0 : e \in E(Q_{d-1}^0)\}$ in \mathbb{R}^{w_0} and $\{f_e^1 : e \in E(Q_{d-1}^1)\}$ in \mathbb{R}^{w_1} such that

$$(P2.0) \quad \sum_{i=1}^{d-1} x_i f_{e(v,i)}^0 = 0 \quad \text{for every } v \in V(Q_{d-1}^0) \text{ and } x \in X_0,$$

$$(P2.1) \quad \sum_{i=1}^{d-1} x_i f_{e(v,i)}^1 = 0 \quad \text{for every } v \in V(Q_{d-1}^1) \text{ and } x \in X_1,$$

$$(P3.0) \quad \text{span}\{f_e^0 : e \in E(Q_{d-1}^0)\} = \mathbb{R}^{w_0}, \text{ and}$$

$$(P3.1) \quad \text{span}\{f_e^1 : e \in E(Q_{d-1}^1)\} = \mathbb{R}^{w_1}.$$

We will now define the vectors $\{f_e : e \in E(Q_d)\} \subseteq \mathbb{R}^{w_0} \oplus \mathbb{R}^{w_1}$ (the space $\mathbb{R}^{w_0} \oplus \mathbb{R}^{w_1}$ is isomorphic to \mathbb{R}^w by (3.5)). We define these vectors in three stages, in such a way that they satisfy (P2) and (P3). First, if $e \in E(Q_{d-1}^0)$, then we set

$$f_e := f_e^0 \oplus 0. \tag{3.8}$$

Next, for each edge of the form $e = e(v, d)$ for $v \in V(Q_{d-1}^0)$, we let

$$f_e := -\frac{1}{z_d} \sum_{i=1}^{d-1} z_i f_{e(v,i)} \quad (3.9)$$

(recall the definition of z above). Finally, if $e = uv \in E(Q_{d-1}^1)$, then we let $e' = u'v'$ where u' and v' are the unique neighbours of u and v in $V(Q_{d-1}^0)$ and define

$$f_e := f_{e'}^0 \oplus f_{e'}^1. \quad (3.10)$$

We now show that $\dim(\text{span}\{f_e : e \in E(Q_d)\}) = w$. By (P3.0) and (3.8), we have

$$\text{span}(\{f_e : e \in E(Q_{d-1}^0)\}) = \mathbb{R}^{w_0} \oplus \{0\}. \quad (3.11)$$

For any $e = uv \in E(Q_{d-1}^1)$, let $e' = u'v'$ where u' and v' are the unique neighbours of u and v in $V(Q_{d-1}^0)$. By (3.10), $f_{e'}^0 \oplus f_{e'}^1 \in \text{span}(\{f_e : e \in E(Q_{d-1}^1)\})$. By (3.8), $f_{e'}^0 \oplus 0 \in \text{span}(\{f_e : e \in E(Q_{d-1}^0)\})$. Thus for any $e \in E(Q_{d-1}^1)$,

$$0 \oplus f_e^1 = f_{e'}^0 \oplus f_{e'}^1 - f_{e'}^0 \oplus 0 \in \text{span}(\{f_e : e \in E(Q_d)\}). \quad (3.12)$$

It follows from (3.12) and (P3.1) that

$$\{0\} \oplus \mathbb{R}^{w_1} \subseteq \text{span}\{f_e : e \in E(Q_d)\}. \quad (3.13)$$

From (3.11) and (3.13) we have $\dim(\text{span}\{f_e : e \in E(Q_d)\}) \geq w_0 + w_1$, which equals w by (3.5). As $\text{span}\{f_e : e \in E(Q_d)\} \subseteq \mathbb{R}^w$, we see that $\dim(\text{span}\{f_e : e \in E(Q_d)\}) = w$ and (P3) holds.

Finally, we prove that (P2) is satisfied. First, let $v \in V(Q_{d-1}^0)$ and let $x \in X$ be arbitrary. Define $x^\dagger := T_z(x)$ and note that $d \notin \text{supp}(x^\dagger)$. We have

$$\sum_{i=1}^d x_i f_{e(v,i)} = \sum_{i=1}^{d-1} x_i^\dagger f_{e(v,i)} + \frac{x_d}{z_d} \sum_{i=1}^d z_i f_{e(v,i)}$$

by definition of T_z . Both of the sums on the right side are zero by (P2.0) and (3.9). Now, suppose that $v \in V(Q_{d-1}^1)$ and let v' be the unique neighbour of v in $V(Q_{d-1}^0)$. By (3.8) and (3.10), when $i \in [d-1]$ we have

$$f_{e(v,i)} = f_{e(v',i)}^0 \oplus f_{e(v,i)}^1 = f_{e(v',i)} + 0 \oplus f_{e(v,i)}^1. \quad (3.14)$$

Using (3.14) and the fact that $e(v, d) = e(v', d)$, for $x \in X$ we have

$$\sum_{i=1}^d x_i f_{e(v,i)} = \sum_{i=1}^d x_i f_{e(v',i)} + \sum_{i=1}^{d-1} x_i (0 \oplus f_{e(v,i)}^1).$$

The first sum is zero by the fact that $\sum_{i=1}^d x_i f_{e(v',i)} = 0$, which was proved above (as $v' \in V(Q_{d-1}^0)$). The second sum is zero by (P2.1). Therefore, (P2) holds. This completes the proof of the lemma. \square

4.4 General Grids

Our objective in this section is to determine the weak saturation number of S_{r+1} in $\prod_{i=1}^d [a_i]$ when $0 \leq r \leq d$. For $2 \leq a_1, \dots, a_d$, define

$$w_r(a_1, \dots, a_d) := \sum_{\substack{S \subseteq [d] \\ |S| \leq r-1}} \left(\prod_{i \in S} (a_i - 2) \right) \left((r - |S|)2^{r-|S|-1} + \sum_{j=1}^{r-|S|-1} \binom{d-|S|-j-1}{r-|S|-j} j2^{j-1} \right).$$

We will prove the following, which directly implies Theorem 1.6.

Theorem 4.1. *For $0 \leq r \leq d$ and $a_1, \dots, a_d \geq 2$, we have*

$$\text{wsat} \left(\prod_{i=1}^d [a_i], S_{r+1} \right) = w_r(a_1, \dots, a_d).$$

The next proposition is an easy consequence of the definition of $w_r(a_1, \dots, a_d)$.

Proposition 4.2. *For $0 \leq r \leq d$, the value of $w_r(a_1, \dots, a_d)$ is equal to:*

(W1) 0, if $r = 0$;

(W2) $r2^{r-1} + \sum_{j=1}^{r-1} \binom{d-j-1}{r-j} j2^{j-1}$, if $a_1 = \dots = a_d = 2$ and $1 \leq r \leq d$; and

(W3) $w_r(a_1, \dots, a_{i-1}, a_i - 1, a_{i+1}, \dots, a_d) + w_{r-1}(a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_d)$, for any $1 \leq i \leq d$ such that $a_i \geq 3$, if $1 \leq r \leq d$.

Proof. (W1) is trivial. Now consider (W2). Here, only the term $S = \emptyset$ contributes to the expression in the definition of $w_r(a_1, \dots, a_d)$ and (W2) holds.

Now for (W3), without loss of generality assume that $i = d$. Define

$$f(r, |S|, d) := (r - |S|)2^{r-|S|-1} + \sum_{j=1}^{r-|S|-1} \binom{d-|S|-j-1}{r-|S|-j} j2^{j-1}$$

and observe that $f(r, |S|, d) = f(r-1, |S|-1, d-1)$. Thus we have

$$\begin{aligned} w_r(a_1, \dots, a_d) - w_r(a_1, \dots, a_{d-1}) &= \sum_{\substack{S \subseteq [d] \\ |S| \leq r-1 \\ d \in S}} \left(\prod_{i \in S \setminus \{d\}} (a_i - 2) \right) f(r, |S|, d) \\ &= \sum_{\substack{S \subseteq [d] \\ |S| \leq r-1 \\ d \in S}} \left(\prod_{i \in S \setminus \{d\}} (a_i - 2) \right) f(r-1, |S|-1, d-1) \\ &= \sum_{\substack{S \subseteq [d-1] \\ |S| \leq r-2}} \left(\prod_{i \in S} (a_i - 2) \right) f(r-1, |S|, d-1) \\ &= w_{r-1}(a_1, \dots, a_{d-1}), \end{aligned}$$

as required. \square

Observe that for any $\sigma : [d] \rightarrow [d]$, a permutation of the integers $\{1, \dots, d\}$, we have

$$w_r(a_1, \dots, a_d) = w_r(a_{\sigma(1)}, \dots, a_{\sigma(d)}).$$

So throughout the section we may assume that $2 \leq a_1 \leq \dots \leq a_d$.

Define $G := \prod_{i=1}^d [a_i]$. In the proof of Theorem 4.1, we employ an inductive approach similar to the one used in the proof of Theorem 1.5. The main difference is that a vertex v of G may be incident to either one or two edges in direction $i \in [d]$ depending on whether or not $v_i \in \{1, a_i\}$. With this in mind, we define a labelling of the edges of G .

Definition 4.3. Say that an edge $e = uv \in E(G)$ in direction $i \in [d]$ is *odd* if $\min\{u_i, v_i\}$ is odd and *even* otherwise. We label e by $e(v, 2i - 1)$ if e is odd and $e(v, 2i)$ if e is even.

Note that each edge of G receives two labels, one for each of its endpoints.

Definition 4.4. For $v \in V(G)$, define $I_v^G := \{j \in [2d] : e(v, j) \in E(G)\}$.

See Figure 4.2 for an explicit example of how we label the edges and define I_v^G .

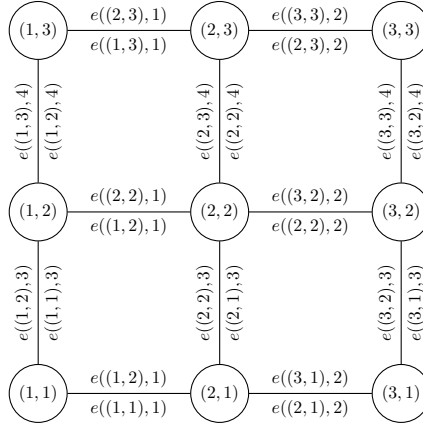


Figure 4.2: The edge labels where $G = \prod_{i=1}^2 [3]$. The lower label of each pair corresponds to the vertex on the left of the pair. We have $I_{(1,1)}^G = \{1, 3\}$, $I_{(2,2)}^G = \{1, 2, 3, 4\}$ and $I_{(3,3)}^G = \{2, 4\}$, to give three examples of I_v^G .

We are now in a position to prove Theorem 4.1. Define $w := w_r(a_1, \dots, a_d)$. As with the proof of Theorem 1.5, we state a lemma from which we deduce Theorem 4.1, and then we prove the lemma. The hypercube case of the previous section will be used as a base case via Lemma 3.6.

Lemma 4.5. Let $0 \leq r \leq d$ and let $G := \prod_{i=1}^d [a_i]$. Let X be a subspace of \mathbb{R}^{2d} of dimension $2d - r$ such that $|\text{supp}(x)| \geq r + 1$ for every $x \in X \setminus \{0\}$. Then there is a spanning subgraph F of G and a collection $\{f_e : e \in E(G)\} \subseteq \mathbb{R}^w$ such that

(G1) F is weakly (G, S_{r+1}) -saturated and $|E(F)| = w$,

(G2) $\sum_{i=1}^{2d} x_i f_{e(v,i)} = 0$ for every $v \in V(G)$ and $x \in X$ such that $|\text{supp}(x)| \subseteq I_v^G$, and

(G3) $\text{span}\{f_e : e \in E(G)\} = \mathbb{R}^w$.

Observe that such a space X exists by Lemma 3.3.

Proof of Theorem 4.1. We will show $\text{wsat}(G, S_{r+1}) = w$, for each $0 \leq r \leq d$. First observe that the existence of a graph F satisfying (G1) implies $\text{wsat}(G, S_{r+1}) \leq w$. To obtain a matching lower bound, we apply Lemma 3.2 as we did in the hypercube case. The edge sets of copies of S_{r+1} in G are the sets of the form $\{e(v, i) : i \in T\}$, where $v \in V(G)$ and T is a subset of I_v^G of cardinality $r + 1$. By applying Lemma 3.4 together with (G2), we see that the conditions of Lemma 3.2 are satisfied. Thus by (G3), $\text{wsat}(G, S_{r+1}) \geq w$. \square

Proof of Lemma 4.5. We proceed by induction on $|V(G)|$. We begin with the boundary cases.

Case 1: $r = 0$.

In this case, $S_{r+1} \simeq K_2$. Also, $w = 0$ and $X = \mathbb{R}^{2d}$. We let F be a spanning subgraph of G with no edges and set $f_e := 0$ for every $e \in Q_d$. Properties (G1), (G2) and (G3) are satisfied trivially.

Case 2: $a_1 = \dots = a_d = 2$ and $1 \leq r \leq d$.

In this case, G is isomorphic to Q_d and every edge of G is odd. Let X' be the subspace of X consisting of all vectors x of X such that every element of $\text{supp}(x)$ is odd. As $X' \subseteq X$, every vector $x \in X'$ has $|\text{supp}(x)| \geq r + 1$. We now show that X' has dimension $d - r$. As X is a subspace of \mathbb{R}^d , by Lemma 3.4, if X' had dimension $d - r + \ell$, for some $\ell > 0$, then there would exist some $x \in X'$ with $|\text{supp}(x)| = r - \ell + 1$, a contradiction. Thus $\dim(X') \leq d - r$. Now, again using Lemma 3.4, for $i = r + 1, \dots, d$ let $y_i \in X$ be a vector with support $\{2k - 1 : k = 1, \dots, r\} \cup \{2i - 1\}$. The set $\{y_i : i = r + 1, \dots, d\}$ is a set of $d - r$ linearly independent vectors in X' . Thus $\dim(X') = d - r$ as required.

As $|\text{supp}(x)| \geq r + 1$ for all $x \in X'$ and $\dim(X') = d - r$, we are done by applying Lemma 3.6.

Case 3: $a_d \geq 3$ and $1 \leq r \leq d$.

Using our induction hypothesis and the proof of Theorem 4.1, for any d -dimensional rectangular grid $\prod_{i=1}^d [a_i]$, with $|V(G')| < |V(G)|$ and $0 \leq r \leq d$, we have

$$\text{wsat} \left(\prod_{i=1}^d [a_i], S_{r+1} \right) = w_r(a_1, \dots, a_d). \quad (4.6)$$

Define

$$G_1 := \prod_{i=1}^{d-1} [a_i] \times [a_d - 1], \text{ and}$$

G_2 to be the graph induced by $V(G \setminus G_1)$.

Observe that every vertex of G_2 has a unique neighbour in $V(G_1)$. The edges with one endpoint in G_1 and the other in G_2 will play a particular role in the proof. We define

$$\tau := \begin{cases} 2d - 1 & a_d - 1 \text{ is odd} \\ 2d & a_d - 1 \text{ is even,} \end{cases}$$

and we write $\bar{\tau}$ for the unique element of $\{2d - 1, 2d\} \setminus \{\tau\}$. Observe that for $v \in V(G_2)$, we have that $\bar{\tau} \notin I_v^G$, and that $I_v^{G_2} = I_v^G \setminus \{\tau\}$. On the other hand, if $v \in V(G_1)$, then

$$I_v^{G_1} = \begin{cases} I_v^G \setminus \{\tau\} & \text{if } v_d = a_d - 1, \\ I_v^G & \text{otherwise.} \end{cases}$$

For brevity we write

$$w_1 := \text{wsat}(G_1, S_{r+1}),$$

and

$$w_2 := \text{wsat}(G_2, S_r).$$

We begin by constructing a graph F satisfying (G1). By (4.6) we have $\text{wsat}(G_1, S_{r+1}) = w_1$ and $\text{wsat}(G_2, S_r) = w_2$. Thus we can choose F to be a spanning subgraph of G such that

- the subgraph F_1 of F induced by $V(G_1)$ is a weakly (G_1, S_{r+1}) -saturated graph with $|E(F_1)| = w_1$,
- the subgraph F_2 of F induced by $V(G_2)$ is a weakly (G_2, S_r) -saturated graph with $|E(F_2)| = w_2$, and
- no edge from $V(G_1)$ to $V(G_2)$ is contained in F .

By construction and the $i = d$ case of (W3), we have

$$|E(F)| = w_1 + w_2 = w,$$

as required. To see that F is weakly (G, S_{r+1}) -saturated, we add the edges of $E(G) \setminus E(F)$ to F in three stages. First, by definition of F_1 , we can add the edges that are not present in $E(F_1)$ in such a way that every added edge completes a copy of S_{r+1} in G_1 . Next, we can add the edges of the form $e(v, \tau)$ where $v_d = a_d - 1$, in any order. As $r \leq d$, every such v has at least r neighbours in G_1 . As every edge in $E(G_1)$ has already been added, the addition of $e(v, \tau)$ completes a copy of S_{r+1} in G . Finally, we add the edges of G_2 that are not present in F_2 in such a way that each added edge completes a copy of S_r in G_2 . Every such edge completes a copy of S_{r+1} in G since every vertex in G_2 has a neighbour in G_1 and every edge between G_1 and G_2 is already present. Thus, (G1) holds.

It remains to find a collection $\{f_e : e \in E(G)\}$ satisfying (G2) and (G3). Let $\pi : X \rightarrow \mathbb{R}^{2d-2}$ be the projection defined by $\pi : (x_1, \dots, x_{2d}) \mapsto (x_1, \dots, x_{2d-2})$. Let z be a fixed vector of X such that $\bar{\tau} \in \text{supp}(z)$ and define $T_z : X \rightarrow X$ by

$$T_z(x) := x - \frac{x_{\bar{\tau}}}{z_{\bar{\tau}}} z.$$

Define $X_1 := X$ and $X_2 := \pi(T_z(X))$. Since $\ker(T_z) = \text{span}\{z\}$ and $\ker(\pi)$ has empty intersection with the image of T_z , we see that X_2 has dimension $2d - r - 1 = 2(d-1) - (r-1)$. Also, by construction, we have $|\text{supp}(x)| \geq r$ for every non-zero $x \in X_2$. By applying the inductive hypothesis to both G_1 and G_2 , we can find collections $\{f_e^1 : e \in E(G_1)\}$ in \mathbb{R}^{w_1} and $\{f_e^2 : e \in E(G_2)\}$ in \mathbb{R}^{w_2} such that

$$(G2.1) \quad \sum_{i=1}^{2d} x_i f_{e(v,i)}^1 = 0 \text{ for every } v \in V(G_1) \text{ and } x \in X_1 \text{ with } \text{supp}(x) \subseteq I_v^{G_1},$$

$$(G2.2) \quad \sum_{i=1}^{2d-2} x_i f_{e(v,i)}^2 = 0 \text{ for every } v \in V(G_2) \text{ and } x \in X_2 \text{ with } \text{supp}(x) \subseteq I_v^{G_2},$$

$$(G3.1) \quad \text{span}\{f_e^1 : e \in E(G_1)\} = \mathbb{R}^{w_1}, \text{ and}$$

$$(G3.2) \quad \text{span}\{f_e^2 : e \in E(G_2)\} = \mathbb{R}^{w_2}.$$

Using this, we will now construct a collection $\{f_e : e \in E(G)\} \subseteq \mathbb{R}^{w_1} \oplus \mathbb{R}^{w_2} \simeq \mathbb{R}^w$ in three steps. First, for $e \in E(G_1)$, we define

$$f_e := f_e^1 \oplus 0. \tag{4.7}$$

Now consider edges uv , where $v \in V(G_1)$ and $u \in V(G_2)$. As $d \geq r$, we have $|I_v^{G_1}| \geq r$. Thus we can choose $z^v \in X$ to be a vector such that $\text{supp}(z^v) \subseteq I_v^G$ and $\tau \in \text{supp}(z^v)$, which exists by Lemma 3.4. Define

$$f_e := -\frac{1}{z_{\bar{\tau}}^v} \sum_{i \in [2d] \setminus \{\tau\}} z_i^v f_{e(v,i)}. \tag{4.8}$$

Finally if $e = uv \in E(G_2)$, then let $e' = u'v'$ where $u'v'$ are the unique neighbours of u and v in $V(G_1)$ and define

$$f_e := f_{e'}^1 \oplus f_e^2. \quad (4.9)$$

Using identical arguments to those where we determine the span of our chosen vectors in the hypercube case, from (G3.1), (G3.2), (4.7) and (4.9) we see that

$$\dim(\text{span}\{f_e : e \in G_1 \cup G_2\}) = w_1 + w_2 = w,$$

by the $i = d$ case of (W3). Thus (G3) is satisfied.

It remains to show that (G2) holds. Firstly, suppose $v \in V(G_1)$ and let $x \in X$ be such that $\text{supp}(x) \subseteq I_v^G$. If $v_d < a_d - 1$, then $\sum_{i=1}^{2d} x_i f_{e(v,i)} = 0$ by (G2.1). Now suppose that $v_d = a_d - 1$. Define

$$x^\dagger := x - \frac{x_\tau}{z_\tau^v} z^v.$$

We have,

$$\sum_{i=1}^{2d} x_i f_{e(v,i)} = \sum_{i=1}^{2d} x_i^\dagger f_{e(v,i)} + \frac{x_\tau}{z_\tau^v} \sum_{i=1}^{2d} z_i^v f_{e(v,i)}. \quad (4.10)$$

Note that $\tau \notin \text{supp}(x^\dagger)$ and thus $\text{supp}(x^\dagger) \subseteq I_v^{G_1}$. Therefore, both of the sums on the right side of (4.10) are zero by (G2.1) and (4.8).

Finally, consider $v \in V(G_2)$. Let v' be the unique neighbour of v in $V(G_2)$. Given $x \in X$ with $\text{supp}(x) \subseteq I_v^G$ we have

$$\sum_{i=1}^{2d} x_i f_{e(v,i)} = \sum_{i=1}^{2d} x_i f_{e(v',i)} + \sum_{i=1}^{2d-2} x_i (0 \oplus f_{e(v,i)}^2).$$

We have that $\sum_{i=1}^{2d} x_i f_{e(v',i)} = 0$ for $v' \in V(G_1)$, as proved above. The second sum on the right side is zero by (G2.2), which is applicable as $\bar{\tau} \notin I_v^G \supseteq \text{supp}(x)$, and so $x = T_z(x) \in T_z(X)$. This completes the proof of the lemma. \square

4.5 Upper Bound Constructions

In this section, we prove a recursive upper bound on $m(Q_d, r)$ for general $d \geq r \geq 1$ and then apply it to obtain an exact expression for $m(Q_d, 3)$. The main idea behind the proof is to consider Q_d as $Q_r \square Q_{d-r}$, the Cartesian graph product of Q_r and Q_{d-r} . We think of this as having a Q_r where each vertex is replaced by a copy of Q_{d-r} . We call these *primary* Q_{d-r} s. Each edge of Q_r is replaced by a set of edges between corresponding vertices of the primary Q_{d-r} s at the endpoints.

Given $x \in V(Q_d)$, let $[x]_r$ and $[x]_{d-r}$ denote the vectors obtained by restricting x to its first r coordinates and last $d - r$ coordinates, respectively. The coordinates in $[x]_r$ play

the role of telling us which primary Q_{d-r} contains x . The coordinates $[x]_{d-r}$ identify which vertex x is, in the primary Q_{d-r} . In each primary Q_{d-r} , say Q , we choose an optimal t -neighbour percolating set T_Q , where t ranges between 0 and r . We choose these such that the union of the T_Q form an r -neighbour percolating set in Q_d .

Lemma 5.1. For $d \geq r \geq 1$,

$$m(Q_d, r) \leq m(Q_{d-r}, r) + (r-1)m(Q_{d-r}, r-1) + \sum_{j=1}^{\lceil r/2 \rceil - 1} \binom{r}{2j+1} m(Q_{d-r}, r-2j).$$

Proof. For $1 \leq t \leq r$, let B_t be a fixed subset of $V(Q_{d-r})$ of cardinality $m(Q_{d-r}, t)$ which percolates with respect to the t -neighbour bootstrap process in Q_{d-r} .

We partition $\{0, 1\}^r$ into $r+1$ layers L_0, \dots, L_r such that L_i consists of the vectors whose coordinate sum is equal to i . Using this partition, we construct a percolating set A_0 in Q_d . Given $x \in V(Q_d)$, we include x in A_0 if one of the following holds:

- $[x]_r \in L_1$ and either
 - $[x]_r = (1, 0, \dots, 0)$ and $[x]_{d-r} \in B_r$.
 - $[x]_r \neq (1, 0, \dots, 0)$ and $[x]_{d-r} \in B_{r-1}$.
- $[x]_r \in L_{2j+1}$ for some $1 \leq j \leq \lceil r/2 \rceil - 1$ and $[x]_{d-r} \in B_{r-2j}$.

It is clear that

$$|A_0| = m(Q_{d-r}, r) + (r-1)m(Q_{d-r}, r-1) + \sum_{j=1}^{\lceil r/2 \rceil - 1} \binom{r}{2j+1} m(Q_{d-r}, r-2j)$$

by construction. See Figure 4.3 for an example in the case $r = 3$.

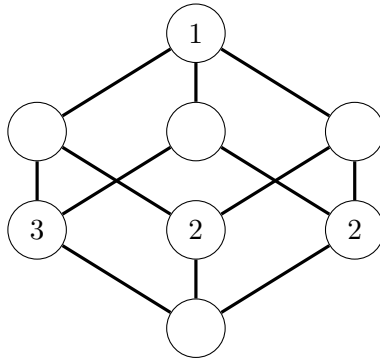


Figure 4.3: An illustration of the set A_0 constructed in the proof of Theorem 5.1 in the case $r = 3$. Each node represents a copy of Q_{d-3} . The set A_0 consists of a copy of B_i on each node labelled $i \in \{1, 2, 3\}$.

We will be done if we can show that A_0 percolates with respect to the r -neighbour bootstrap process.

We begin by showing that every vertex x with $[x]_r \in L_0 \cup L_1$ is eventually infected. First, we can infect every vertex x such that $[x]_r = (1, 0, \dots, 0)$, one by one in some order, by definition of B_r . Next, consider a vertex x such that $[x]_r \in L_0$ and $[x]_{d-r} \in B_{r-1}$. We will see that x has r infected neighbours. For $i \in [r]$, define $e_i \in \mathbb{R}^r$ to be the vector taking the value 1 on coordinate i , and 0 on all other coordinates. For $i = 2, \dots, r$, the vector $e_i \oplus [x]_{d-r}$ is a neighbour of x in A_0 . The final neighbour of x is $e_1 \oplus [x]_{d-r}$, which was either in A_0 or is infected by the previous argument.

Thus, every such x becomes infected. Now, by definition of B_{r-1} , the remaining vertices x with $[x]_r \in L_0$ can be infected since every such vertex has an infected neighbour y with $[y]_r = (1, 0, \dots, 0)$. Finally, each vertex x such that $x \neq (1, 0, \dots, 0)$ and $[x]_r \in L_1$ becomes infected using the definition of B_{r-1} and the fact that every vertex y with $[y]_r \in L_0$ is already infected.

Now, suppose that, for some $1 \leq j \leq \lceil r/2 \rceil - 1$ every vertex x such that $[x]_r \in L_0 \cup \dots \cup L_{2j-1}$ is already infected. We show that every vertex x with $[x]_r \in L_{2j} \cup L_{2j+1}$ is eventually infected. First, consider a vertex x with $[x]_r \in L_{2j}$ and $[x]_{d-r} \in B_{r-2j}$. Such a vertex has $2j$ infected neighbours y such that $[y]_r \in L_{2j-1}$ and $r - 2j$ neighbours $z \in A_0$ such that $[z]_r \in L_{2j+1}$. Therefore, every such x becomes infected. Now, by definition of B_{r-2j} , the remaining vertices x such that $[x]_r \in L_{2j}$ can be infected since every such vertex has $2j$ infected neighbours y such that $[y]_r \in L_{2j-1}$. Finally, each vertex x such that $[x]_r \in L_{2j+1}$ becomes infected using the definition of B_{r-2j} and the fact that every vertex y with $[y]_r \in L_{2j-1}$ is already infected.

Finally, if r is even, then we need to show that every vertex of L_r becomes infected. Every such vertex has precisely r neighbours in L_{r-1} . Thus, given that every vertex of L_{r-1} is infected, x becomes infected as well. This completes the proof. \square

We remark that the recursion in Lemma 5.1 gives a bound of the form $m(Q_d, r) \leq \frac{1+o(1)d^{r-1}}{r!}$ where the second order term is better than the one in (1.2). Next, we prove Theorem 1.8.

Proof of Theorem 1.8. The lower bound follows from Theorem 1.3. We prove the upper bound by induction on d . First, we settle the cases $d \in \{3, \dots, 8\}$. For notational convenience, we associate each element v of $\{0, 1\}^d$ with the of subset of $[d]$ for which v is the characteristic vector. Moreover, we identify each non-empty subset of $[d]$ with the concatenation of its elements (e.g. $\{1, 3, 7\}$ is written 137). One can verify (by hand or

by computer) that the set A_0^d , defined below, percolates with respect to the 3-neighbour bootstrap process in Q_d and that it has cardinality $\left\lceil \frac{d(d+3)}{6} \right\rceil + 1$.

$$A_0^3 := \{1, 2, 3, 123\},$$

$$A_0^4 := (A_0^3 \setminus \{3\}) \cup \{134, 4, 234\},$$

$$A_0^5 := (A_0^4 \setminus \{134\}) \cup \{135, 245, 12345\},$$

$$A_0^6 := (A_0^5 \setminus \{135, 245\}) \cup \{346, 12356, 456, 23456\},$$

$$A_0^7 := (A_0^6 \setminus \{346\}) \cup \{13457, 24567, 12367, 1234567\},$$

$$A_0^8 := (A_0^7 \setminus \{13457, 24567\}) \cup \{34568, 1234578, 34678, 25678, 2345678\}.$$

Now, suppose $d \geq 9$ and that the theorem holds for smaller values of d . If d is odd, then we apply Lemma 5.1 to obtain

$$m(Q_d, 3) \leq m(Q_{d-3}, 3) + 2m(Q_{d-3}, 2) + m(Q_{d-3}, 1).$$

Clearly, $m(Q_{d-3}, 1) = 1$ and, by Theorem 1.1, $m(Q_{d-3}, 2) \leq \frac{d-3}{2} + 1$ (since $d-3$ is even). Therefore, by the inductive hypothesis,

$$m(Q_d, 3) \leq \left\lceil \frac{(d-3)d}{6} \right\rceil + 1 + 2 \left(\frac{d-3}{2} + 1 \right) + 1 = \left\lceil \frac{d(d+3)}{6} \right\rceil + 1.$$

Now, suppose that $d \geq 10$ is even. For $t \in \{1, 2, 3\}$, let B_t be a subset of $V(Q_{d-6})$ of cardinality $m(Q_{d-6}, t)$ which percolates with respect to the t -neighbour bootstrap process on Q_{d-6} and let A_0^6 be as above. Given a vector $x \in V(Q_d)$, let $[x]_6$ be the restriction of x to its first six coordinates and $[x]_{d-6}$ be the restriction of x to its last $d-6$ coordinates. We define a subset A_0 of $V(Q_d)$. We include a vertex $x \in V(Q_d)$ in A_0 if $[x]_6 \in A_0^6$ and one of the following holds:

- $[x]_6 = (0, 0, 1, 1, 0, 1)$ and $[x]_{d-6} \in B_3$.
- $[x]_6 \neq (0, 0, 1, 1, 0, 1)$ and we have $x_5 = 1$ and $[x]_{d-6} \in B_2$.
- $x_5 = x_6 = 0$ and $[x]_{d-6} \in B_1$.

We now describe a set of rules, from which it is not difficult to check that A_0 percolates. We require some preliminary definitions. For $w \in \{0, 1\}^r$, we define

$$Y_w^0 := \{x \in V(Q_d) : [x_r] = w\}$$

and for $i = 1, 2, 3$ we define

$$Y_w^i := \{x \in Y_w^0 : [x]_{d-r} \in B_i\}.$$

If every vertex in some set X is infected, we say X is *infected*. We now state the rules.

- For $0 \leq i \leq 3$, Y_w^i becomes infected if there exist distinct $w_1, w_2, w_3 \in N_{Q_r}(w)$ such that $Y_{w_j}^i$ is infected for all $j = 1, 2, 3$.
- Y_w^0 becomes infected if:
 - Y_w^3 is infected; or
 - Y_w^2 is infected and there exists some $w' \in N_{Q_r}(w)$ such that $Y_{w'}^0$ is infected; or
 - Y_w^1 is infected and there exist distinct $w_1, w_2 \in N_{Q_r}(w)$ such that $Y_{w_1}^0$ and $Y_{w_2}^0$ are infected.

Using these rules, one can check that A_0 percolates.

By construction,

$$|A_0| = m(Q_{d-6}, 3) + 4m(Q_{d-6}, 2) + 5m(Q_{d-6}, 1)$$

which equals

$$\left\lceil \frac{(d-6)(d-3)}{6} \right\rceil + 1 + 4 \left(\frac{d-6}{2} + 1 \right) + 5 = \left\lceil \frac{d(d+3)}{6} \right\rceil + 1$$

by the inductive hypothesis. The result follows. \square

4.6 An Explicit Linear Algebraic Construction

Given integers k and ℓ with $k \geq \ell \geq 0$, we construct an explicit subspace X of \mathbb{R}^k of dimension $k - \ell$ such that $|\text{supp}(x)| \geq \ell + 1$ for every $x \in X \setminus \{0\}$. This can be seen as an alternative proof of Lemma 3.3.

The construction is based on a *Vandermonde matrix*. For $1 \leq i \leq k - \ell$ we let $\alpha_i \in \mathbb{R}^\ell$ be the vector such that, for $1 \leq j \leq \ell$,

$$\alpha_{i,j} := i^j.$$

Now, for $1 \leq i \leq k - \ell$ let e_i be the i th standard unit basis vector of $\mathbb{R}^{k-\ell}$ and define

$$v_i := \alpha_i \oplus e_i.$$

The space X is defined to be $\text{span}\{v_1, \dots, v_{k-\ell}\}$. It is clear that $\dim(X) = k - \ell$ by construction. All that remains is to show that $|\text{supp}(x)| \geq \ell + 1$ for every $x \in X \setminus \{0\}$. We require a few definitions.

Definition 6.1. Given a set $T \subseteq [k]$, let $\pi_T : \mathbb{R}^k \rightarrow \mathbb{R}^{|T|}$ be the standard projection $\pi_T : (x_1, \dots, x_k) \mapsto (x_i : i \in T)$.

Recall, from Chapter 3 that for $n \geq 1$, a collection $Z \subseteq \mathbb{R}^n$ is in *general position* in \mathbb{R}^n if any set of at most n vectors from Z is linearly independent.

We will use the following proposition to show $|\text{supp}(x)| \geq \ell + 1$. Our proof of the proposition follows an argument of Moshonkin [107].

Proposition 6.2. *For any set $T \subseteq [\ell]$, the vectors $\{\pi_T(\alpha_i) : 1 \leq i \leq k - \ell\}$ are in general position in $\mathbb{R}^{|T|}$.*

Proof. We assume that $|T| \geq 1$; otherwise, the result is trivial. Let $t := |T|$. Suppose that the proposition is false and let $I \subseteq [k - \ell]$ be a set of cardinality t for which there exists $\{c_i : i \in I\}$, not all of which are zero, such that

$$\sum_{i \in I} c_i \pi_T(\alpha_i) = 0.$$

Equivalently, for each $j \in T$,

$$\sum_{i \in I} c_i i^j = 0.$$

Since the determinant of a square matrix is equal to the determinant of its transpose, there must also exist scalars $\{c'_j : j \in T\}$, not all zero, such that

$$\sum_{j \in T} c'_j i^j = 0$$

for every $i \in I$.

Let $p(x)$ denote the real polynomial $\sum_{j \in T} c'_j x^j$. Then $p(x)$ is a polynomial with between one and t non-zero terms and at least t positive real roots (namely, each $i \in I$). We show, by induction on t , that no such polynomial can exist. The base case $t = 1$ is trivial. Now, let $p(x)$ be a polynomial with $t \geq 2$ non-zero terms and at least t positive real roots. Pick s to be the largest integer such that $p(x) = x^s q(x)$. By definition of $p(x)$, $s \geq 1$. For any $0 < y \in \mathbb{R}$, $q(y) = 0$ if and only if $p(y) = 0$. However, the derivative of $q(x)$ has at most $t - 1$ positive terms and at least $t - 1$ positive real roots, contradicting the inductive hypothesis. This completes the proof. \square

Now, suppose that $x \in X \setminus \{0\}$ such that $|\text{supp}(x)| \leq \ell$. Define

$$U_1 := \text{supp}(x) \cap [\ell],$$

$$U_2 := \text{supp}(x) \setminus [\ell] \text{ and}$$

$$T := [\ell] \setminus U_1.$$

Since $x \in X \setminus \{0\}$, we can write

$$x = \sum_{i=1}^{k-\ell} c_i v_i$$

for scalars c_i which are not all zero. But by construction of X , it must be the case that $c_i = 0$ for each i such that $i + \ell \notin U_2$. Thus,

$$0 = \pi_T(x) = \sum_{i:i+\ell \in U_2} c_i \pi_T(v_i) = \sum_{i:i+\ell \in U_2} c_i \pi_T(\alpha_i). \quad (6.3)$$

However, since $|U_1| + |U_2| \leq \ell$ we have $|T| = \ell - |U_1| \geq |U_2|$. Thus, (6.3) contradicts Proposition 6.2. It follows that $|\text{supp}(x)| \geq \ell + 1$ for every $x \in X \setminus \{0\}$, as required.

4.7 Concluding Remarks

In this chapter, we have determined the main asymptotics of $m(Q_d, r)$ for fixed r and d tending to infinity and obtained a sharper result for $r = 3$. We wonder whether sharper asymptotics are possible for general r .

Question 7.1. For fixed $r \geq 4$ and $d \rightarrow \infty$, does

$$\frac{m(Q_d, r) - \frac{d^{r-1}}{r!}}{d^{r-2}}$$

converge? If so, what is the limit?

As Theorem 1.8 illustrates, it may be possible to obtain an exact expression for $m(Q_d, r)$ for some small fixed values of r . The first open case is the following.

Problem 7.2. Determine $m(Q_d, 4)$ for all $d \geq 4$.

Using a computer, we have determined that $m(Q_5, 4) = 14$, which is greater than the lower bound of 13 implied by Theorem 1.3. Thus, Theorem 1.3 is not tight for general d and r . However, we wonder whether it could be tight when r is fixed and d is sufficiently large.

Question 7.3. For fixed $r \geq 4$, is it true that

$$m(Q_d, r) = 2^{r-1} + \left[\sum_{j=1}^{r-1} \binom{d-j-1}{r-j} \frac{j2^{j-1}}{r} \right]$$

provided that d is sufficiently large?

Chapter 5

Bootstrap Percolation in a Random Hypergraph

Joint work with Robert Morris and Jonathan A. Noel

5.1 Introduction

Recall the following definition from Section 1.4 of Chapter 1. Given a hypergraph \mathcal{H} , the \mathcal{H} -bootstrap process begins with an initial set of *infected* vertices of \mathcal{H} (a vertex that is not infected is *healthy*) and, at each step, a healthy vertex becomes infected if there exists a hyperedge of \mathcal{H} in which it is the unique healthy vertex. The set of initially infected vertices is said to *percolate* if every vertex of \mathcal{H} is eventually infected.

As we said in Chapter 1, our goal in this chapter is to estimate the *critical probability* of the \mathcal{H} -bootstrap process, denoted $p_c(\mathcal{H})$. Recall that,

$$p_c(\mathcal{H}) := \inf \{p \in (0, 1) : \mathbb{P}(V(\mathcal{H})_p \text{ percolates}) \geq 1/2\},$$

where for a finite set X and $p \in (0, 1)$, we let X_p denote a random subset of X obtained by including each element of X with probability p independently of one another. Recall also from Chapter 1 that, for $q \in (0, 1)$, \mathcal{H}_q denotes the hypergraph obtained from \mathcal{H} by including each hyperedge of \mathcal{H} independently with probability q . This chapter is concerned with estimating the quantity $p_c(\mathcal{H}_q)$ when the hypergraph \mathcal{H} is ‘sufficiently well behaved’. In order to state our results, we require a few definitions.

Definition 1.1. Given a hypergraph \mathcal{H} and a set $S \subseteq V(\mathcal{H})$, the *codegree* of S , denoted $\deg(S)$, is the number of hyperedges e of \mathcal{H} with $S \subseteq e$. For $v \in V(\mathcal{H})$, the *degree* of v is defined to be $\deg(\{v\})$ and is denoted by $\deg(v)$.

Definition 1.2. For an r -uniform hypergraph \mathcal{H} and $1 \leq \ell \leq r$, define

$$\Delta_\ell(\mathcal{H}) := \max\{\deg(S) : S \subseteq V(\mathcal{H}), |S| = \ell\}, \text{ and}$$

$$\delta_\ell(\mathcal{H}) := \min\{\deg(S) : S \subseteq V(\mathcal{H}), |S| = \ell\}.$$

We often write $\Delta_1(\mathcal{H})$ as $\Delta(\mathcal{H})$ and $\delta_1(\mathcal{H})$ as $\delta(\mathcal{H})$. We say that \mathcal{H} is d -regular if $\delta(\mathcal{H}) = \Delta(\mathcal{H}) = d$.

Definition 1.3. Given a vertex v of an r -uniform hypergraph \mathcal{H} , define the *neighbourhood* (or *link*) of v to be the set

$$N_{\mathcal{H}}(v) := \{e \setminus \{v\} : e \in E(\mathcal{H}) \text{ and } v \in e\}.$$

The following theorem is a corollary to our main result (Theorem 1.6 below), but contains the main essence of the chapter.

Theorem 1.4. Let $r, s, \alpha, \beta > 0$ be fixed. Let \mathcal{H} be a d -regular, r -uniform hypergraph on N vertices and let $q := \alpha \cdot d^{1/r-1}$. Suppose:

- (a) $N \leq d^\beta$,
- (b) $|N_{\mathcal{H}}(u) \cap N_{\mathcal{H}}(v)| \leq d^{1-s}$, and
- (c) $\Delta_\ell(\mathcal{H}) \leq d^{1-\frac{\ell-1}{r-1}-s}$ for $2 \leq \ell \leq r-1$.

Then

$$p_c(\mathcal{H}_q) = \left(\frac{r-2}{\alpha^{1/(r-2)}(r-1)^{(r-1)/(r-2)} + o(1)} \right) \cdot d^{-1/(r-1)},$$

with high probability as $d \rightarrow \infty$.

Our main theorem (Theorem 1.6) is a more general version of Theorem 1.4. We allow weaker conditions to be imposed on the underlying hypergraph \mathcal{H} . In order to easily state the conditions on our hypergraph, it is helpful to introduce the following definition.

Definition 1.5. Given an integer $r \geq 2$ and real numbers $d, \nu > 0$ and $\rho \in [0, 1]$, we say that an r -uniform hypergraph \mathcal{H} is (d, ρ, ν) -well behaved if the following conditions hold:

- (a) $\Delta(\mathcal{H}) \leq d$,
- (b) $\delta(\mathcal{H}) \geq d(1-\rho)$,
- (c) $\Delta_\ell(\mathcal{H}) \leq \rho \cdot d^{1-\frac{\ell-1}{r-1}}$ for $2 \leq \ell \leq r-1$,
- (d) $|N_{\mathcal{H}}(u) \cap N_{\mathcal{H}}(v)| \leq \rho \cdot d$ for distinct $u, v \in V(\mathcal{H})$, and
- (e) $|V(\mathcal{H})| \leq \nu$.

Observe that the conditions in Theorem 1.4 imposed on a d -regular r -uniform hypergraph \mathcal{H} simply amount to \mathcal{H} being (d, d^{-s}, d^β) -well behaved. We are now ready to state the main result of the chapter.

Theorem 1.6. *For fixed $r \geq 3$ and positive real numbers c, α, β and ε there exists a positive constant $K = K(r, c, \alpha)$ such that, for d sufficiently large, if \mathcal{H} is an r -uniform $(d, \log^{-K}(d), d^\beta)$ -well behaved hypergraph, then*

- if $c^{r-2}\alpha < \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$, then

$$\mathbb{P}\left(V(\mathcal{H})_{c/d^{1/(r-1)}} \text{ percolates in } \mathcal{H}_{\alpha/d^{1/(r-1)}}\right) \leq \varepsilon.$$

- if $c^{r-2}\alpha > \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$, then

$$\mathbb{P}\left(V(\mathcal{H})_{c/d^{1/(r-1)}} \text{ percolates in } \mathcal{H}_{\alpha/d^{1/(r-1)}}\right) \geq 1 - \varepsilon.$$

In particular, K will be chosen so that $\log^K(d)$ grows faster than any relevant constant power of $\log d$ that appears throughout the proof. We remark that we have not attempted to optimise the dependence of K on r, c and α .

Theorem 1.6 is stronger than Theorem 1.4 in two ways. Firstly, it applies to a much wider class of hypergraphs (allowing larger codegrees, neighbourhood intersections, etc.) and, secondly, it implies that the probability of percolation transitions from close to zero to close to one within a small window of the critical probability; i.e. the process exhibits a ‘sharp threshold’.

As discussed in Chapter 1, the main application of Theorem 1.6 is to obtain a sharp threshold for the $\mathcal{H}_{K_n, F}$ -bootstrap process where $\mathcal{H}_{K_n, F}$ is a hypergraph in which each vertex of $\mathcal{H}_{K_n, F}$ corresponds to an edge of K_n and the hyperedges of $\mathcal{H}_{K_n, F}$ are precisely the edge sets of copies of F in K_n . This hypergraph is $|E(F)|$ -uniform and $d(n, F)$ -regular, where $d(n, F) = \Theta(n^{|V(F)|-2})$ (the constant factor is related to the number of automorphisms of F which fix a hyperedge).

Recall the following definitions from Chapter 1.

Definition 1.7. For a graph F with at least two edges, the *2-density* of F is defined to be

$$d_2(F) := \frac{|E(F)| - 1}{|V(F)| - 2}.$$

Definition 1.8. A graph F with at least two edges is said to be *2-balanced* if $d_2(F) \geq d_2(F')$ for every proper subgraph F' of F with at least two edges. If the inequality is strict for every such F' , then we say that F is *strictly 2-balanced*.

For F a strictly 2-balanced graph, we will apply Theorem 1.6 to $\mathcal{H}_{K_n, F}$ to prove the following.

Theorem 1.9. *Let F be a strictly 2-balanced graph with at least three edges and define $r := |E(F)|$, $\mathcal{H} := \mathcal{H}_{K_n, F}$ and $d := d(n, F)$. Let $q := \alpha d^{-1/(r-1)}$ for some fixed $\alpha > 0$. Then*

$$p_c(\mathcal{H}_q) = \left(\frac{r-2}{\alpha^{1/(r-2)}(r-1)^{(r-1)/(r-2)} + o(1)} \right) \cdot d^{-1/(r-1)},$$

with high probability as $n \rightarrow \infty$.

Our proof of Theorem 1.6 involves an application of the ‘differential equations method’. This technique was introduced by Wormald [135, 136]. Some notable recent applications of this method include the analysis of the so called ‘triangle-free process’ by Bohman [25], Bohman and Keevash [27] and Fiz Pontiveros, Griffiths and Morris [55] and the more general ‘ H -free process’ for strictly 2-balanced graphs H by Bohman and Keevash [26]. For a fixed graph H , the H -free process is equivalent to the following ‘random greedy algorithm’ for producing a maximal independent set in $\mathcal{H}_{K_n, H}$: start with $I := \emptyset$ and, at each step, if there exists a vertex $v \in V(\mathcal{H}_{K_n, H}) \setminus I$ such that $I \cup \{v\}$ is an independent set of $\mathcal{H}_{K_n, H}$, then choose such a vertex uniformly at random and add it to I ; otherwise, terminate the algorithm.

Bennett and Bohman [24] generalised the result of [26] by proving that, with high probability, the random greedy algorithm gives an independent set of size $\Omega\left(N(\log N/d)^{1/(r-1)}\right)$ when applied to any r -uniform d -regular hypergraph \mathcal{H} with $N = d^{O(1)}$ vertices satisfying certain upper bounds on both $\Delta_\ell(\mathcal{H})$, for $2 \leq \ell \leq r-1$, and on $|N_{\mathcal{H}}(u) \cap N_{\mathcal{H}}(v)|$ for distinct $u, v \in V(\mathcal{H})$. In [23], Bennett and Bohman used the same method again (but with somewhat different conditions on \mathcal{H}) to prove a result on the random greedy algorithm for producing a maximal matching in a hypergraph; this result has applications in the problem of packing edge-disjoint copies of a graph H into a graph G (see, e.g., [85]).

We remark that the differential equations method was also used in the paper of Korándi, Peled and Sudakov [92] and that our proof uses some ideas from theirs. However, because we are working in much greater generality (and with much larger codegrees and neighbourhood intersections), we will require a number of different arguments. In particular, the collection of variables that we track differs from theirs, even in the case $r = 3$.

Our work generalises the result of [92] in a way which is analogous to the generalisation of [26] in [24] except that the result of [92] applies only to triangles and, because of the nature of the \mathcal{H} -bootstrap process, we are able to (simultaneously) handle larger codegrees and neighbourhood intersections than [24] and obtain an exact threshold for percolation.

The rest of the chapter is organised as follows. In the next section, we give a detailed outline of how the proof of Theorem 1.6 will proceed. In Section 5.3 we summarise the main probabilistic tools that will be used throughout the proof. A few preliminary lemmas will be proved in Section 5.4 before we can move on to the main meat of the proof. The proof of Theorem 1.6 is divided into four parts which are contained in Sections 5.5, 5.6, 5.7 and 5.8. Finally, in Section 5.9 we use Theorem 1.4 to prove Theorem 1.9 and a generalisation of it to strictly k -balanced hypergraphs (defined in the section itself).

5.2 Outline of the Proof

To study the \mathcal{H}_q -bootstrap process we divide it into two phases; in each phase, we analyse a different random process on \mathcal{H} . These processes will be equivalent to the \mathcal{H}_q -bootstrap process in the sense that the final set of infected vertices is the same. We will track¹ a family of variables through these processes and use this to determine whether or not the initial infection will percolate. In the first phase, one by one we sample certain hyperedges of \mathcal{H} to see if they are in \mathcal{H}_q . In the second phase we sample certain large sets of hyperedges of \mathcal{H} in rounds. Our analysis of the first phase involves an application of the differential equations method and is done in essentially the same way regardless of whether $c^{r-2}\alpha$ is smaller or larger than $\frac{(r-2)^{r-2}}{(r-1)^{r-1}}$. In the second phase, the analysis differs depending on which of these cases we are in.

The purpose of this section is to provide a fairly detailed outline of the proof of Theorem 1.6 and to prove some preliminary results. In particular, we will describe the random processes that we will analyse. We will also motivate and formally define the variables that we wish to track during the application of the differential equations method, and during the second phase.

Before diving into deeper details, we fix some parameters and notation that will be used throughout the chapter. Let $r, c, \alpha, \beta > 0$ be fixed, let K be large with respect to r, c and α and, for large d , let \mathcal{H} be an r -uniform $(d, \log^{-K}(d), d^\beta)$ -well behaved hypergraph (defined in Definition 1.5). Set $N := |V(\mathcal{H})|$. Note that, by property (e) of Definition 1.5, we have $N \leq d^\beta$. Since each set of size r contains exactly $\binom{r}{\ell}$ sets of size ℓ , by the pigeonhole principle, for $1 \leq \ell \leq r-1$ we have,

$$\Delta_\ell(\mathcal{H}) \geq \frac{|E(\mathcal{H})| \binom{r}{\ell}}{\binom{N}{\ell}} \geq \frac{\delta(\mathcal{H})N \binom{r}{\ell}}{r \binom{N}{\ell}}.$$

¹Throughout the chapter, when we say that we *track* a random variable, it will always mean one of two things: either (a) we show that it is concentrated or (b) we show that it satisfies a certain upper bound with high probability (i.e. with probability tending to 1 as $V(\mathcal{H}) \rightarrow \infty$).

Rearranging this expression gives

$$N^{\ell-1} \geq \frac{a\delta(\mathcal{H})}{\Delta_\ell(\mathcal{H})},$$

for some constant a . Now applying conditions (b) and (c) of Definition 1.5, and setting $\rho := \log^{-K}(d)$ we obtain

$$N^{\ell-1} \geq \frac{ad(1-\rho)}{d^{1-\frac{\ell-1}{r-1}}\rho} = ad^{\frac{\ell-1}{r-1}} \frac{1-\rho}{\rho}.$$

Setting $\ell = 2$ gives

$$N = \Omega\left(d^{1/(r-1)} \log^K(d)\right). \quad (2.1)$$

In particular,

$$\log(N) = \Theta(\log(d)). \quad (2.2)$$

In what follows, we will write $p := c \cdot d^{-1/(r-1)}$ and $q := \alpha \cdot d^{-1/(r-1)}$.

5.2.1 The First Phase and the Differential Equations Method

Here we define the random hypergraph process that we will analyse during the first phase. At time m , we will have a hypergraph $\mathcal{H}(m)$ formed by the unsampled hyperedges of \mathcal{H} and a set of infected vertices $I(m)$ (where $\mathcal{H}(m)$ and $I(m)$ will be formally defined below). Both $\mathcal{H}(m)$ and $I(m)$ depend on the outcomes of the process up to this point.

Due to the nature of the \mathcal{H} -bootstrap process, it should not come as a surprise to be told that the most important variable for us to ‘track’ is the cardinality of the set of *open* hyperedges

$$Q(m) := \{e \in E(\mathcal{H}(m)) : |e \cap I(m)| = r - 1\}. \quad (2.3)$$

Recall that \mathcal{H} is an r -uniform hypergraph, so $Q(m)$ is precisely the set of hyperedges that contain a single healthy vertex at time m . If, for some m , we have that $Q(m) = \emptyset$, then the \mathcal{H} -bootstrap process has terminated.

We will now formally introduce the random process through which we track $Q(m)$ during the first phase. In what follows, to *sample* a hyperedge e of \mathcal{H} means to determine whether or not it is contained in \mathcal{H}_q . The sampling of the hyperedge e is said to be *successful* if e is contained in \mathcal{H}_q .

The First Phase Process. At time zero, we let $I(0) := V(\mathcal{H})_p$ be the set of initially infected vertices and let $\mathcal{H}(0) := \mathcal{H}$. Now, for $m \geq 0$, given $I(m)$ and $\mathcal{H}(m)$, we obtain $I(m+1)$ and $\mathcal{H}(m+1)$ in the following way: if $Q(m) = \emptyset$, then set $\mathcal{H}(m+1) := \mathcal{H}(m)$ and $I(m+1) := I(m)$; otherwise, choose an open hyperedge e from $Q(m)$ uniformly at random and sample it. If the sampling is successful, then set $I(m+1) := I(m) \cup \{v\}$ where v is

the unique vertex of $e \setminus I(m)$ and, otherwise, set $I(m+1) := I(m)$. In either case, set $\mathcal{H}(m+1) := \mathcal{H}(m) \setminus \{e\}$.

In what follows, for a collection $X(m)$ of subhypergraphs of $\mathcal{H}(m)$ we will often write $|X(m)|$ simply as $X(m)$ (so for example, we will write $Q(m)$ for $|Q(m)|$). We may also write $I(m)$ for $|I(m)|$. In all cases, it should be clear from context whether we are referring to the collection $X(m)$ or its cardinality; we hope that this abuse of notation does not cause the reader any confusion.

We will run the first phase process to some time m where $Q(m)$ is either ‘large enough’ or ‘small enough’ to be able to complete the proof by sampling many open hyperedges at once in rounds. This is what we will do in the second phase (which will be summarised in Subsection 5.2.2 and the details will be given in Sections 5.7 and 5.8).

We would like to show that, at each step of the process, $Q(m)$ is close to its expected value with high probability. Roughly speaking, if we could bound (conditioned on what has occurred in the process so far) the expected change in the number of open hyperedges at each step and the maximum possible change, then we could apply a martingale concentration inequality to show that $Q(m)$ is concentrated around its expected value. However, in practice, the expected and maximum change depend on a number of other variables and it is only possible to get good enough bounds on these parameters if we also have control over the other variables. The differential equations method is a technique that can be used to simultaneously control such a collection of interdependent variables and this is what we will use to do this.

It is sometimes more convenient to think of our random variables as depending on a continuous variable t rather than the discrete variable m . The scaling that we will use when moving between discrete and continuous settings is

$$t := t_m = m/N, \tag{2.4}$$

for $m \geq 0$. Throughout the chapter we will alternate between the discrete and continuous settings without further comment. In the first phase, we will only consider values of m up to

$$\begin{cases} O(N) & \text{if } c^{r-2}\alpha < \frac{(r-2)^{r-2}}{(r-1)^{r-1}}, \\ O(N \log d) & \text{otherwise.} \end{cases} \tag{2.5}$$

The fact that m does not get ‘too large’ during the first phase will be used in some of the heuristic discussions which follow.

At time zero, each vertex is infected with probability p and, provided that $Q(m-1) \neq \emptyset$, at the m th step of the process a new vertex is infected with probability q . So if $Q(m-1) \neq \emptyset$ then we would expect

$$I(m) \approx pN + mq.$$

Letting M be the number of steps we run the first phase for, using the Chernoff bound (Theorem 3.2) we will prove the following (see Lemma 6.5 of Section 5.6).

Proposition 2.6. *For $0 \leq m \leq M$, with probability at least $1 - N^{-\Omega(\sqrt{\log N})}$,*

$$I(m) = O(\log N \cdot Nd^{-1/(r-1)}).$$

Our goal will be to use the fact that \mathcal{H} is well behaved to show that $I(m)$ behaves similarly to a random infection in which each vertex is infected independently with probability $p + qt_m$ (which we shall call a *uniformly random infection of density $p + qt_m$*), in the sense that $Q(m)$ is close to the value that one would expect in this case. First, let us determine the value of $Q(m)$ which we would expect if $I(m)$ were a uniformly random infection of density $p + qt_m$. We would expect a particular hyperedge of \mathcal{H} to contain exactly $r - 1$ infected vertices with probability $r \cdot (p + qt_m)^{r-1}(1 - (p + qt_m))$. Recall that at each step we sample precisely one open hyperedge and that hyperedge is deleted from the hypergraph. Putting this together, as \mathcal{H} is roughly d -regular and $p + qt_m = o(1)$, we have $|E(\mathcal{H})| \approx d \cdot N/r$ and thus we should expect

$$Q(m) \approx r \cdot (p + qt_m)^{r-1} |E(\mathcal{H})| - m \approx [(c + \alpha t_m)^{r-1} - t_m] \cdot N. \quad (2.7)$$

The main point of the first phase is to show that, up to a small error, $Q(m)$ follows this trajectory (see Theorem 2.13). Before stating this more precisely, let us discuss the motivation behind the number of steps for which we wish to run the process.

Define

$$\gamma(t) := (c + \alpha t)^{r-1} - t. \quad (2.8)$$

Observe that

$$\gamma'(t) = \alpha(r-1)(c + \alpha t)^{r-2} - 1.$$

Therefore, since $c, \alpha > 0$, we have that $\gamma'(t)$ has exactly one real root if r is odd and two real roots (one positive, one negative) if r is even. The rightmost root of $\gamma'(t)$ is a local minimum for $\gamma(t)$ located at

$$t_{\min} := \frac{1}{\alpha} \left(\left(\frac{1}{\alpha(r-1)} \right)^{1/(r-2)} - c \right).$$

Now,

$$\begin{aligned} \gamma(t_{\min}) &= \left(\frac{1}{\alpha(r-1)} \right)^{(r-1)/(r-2)} - \left(\frac{1}{\alpha} \left(\frac{1}{\alpha(r-1)} \right)^{1/(r-2)} - \frac{c}{\alpha} \right) \\ &= \left(\frac{2-r}{(\alpha(r-1))^{(r-1)/(r-2)}} \right) + \frac{c}{\alpha} \end{aligned}$$

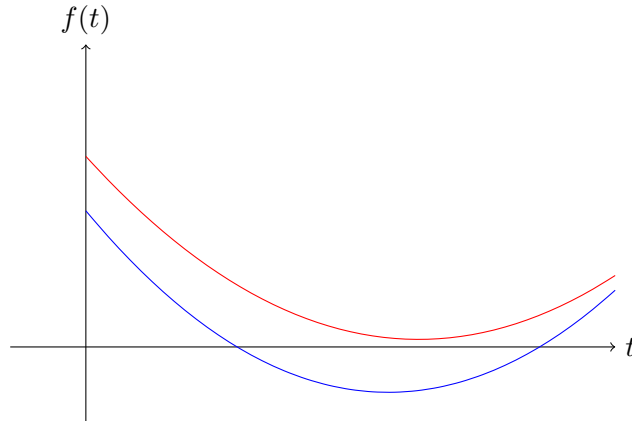


Figure 5.1: The red curve gives an example of the trajectory of $Q(m)$ when $r = 3$ and $c^{r-2}\alpha > \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$. The blue curve is an example of the trajectory of $Q(m)$ when $r = 3$ and $c^{r-2}\alpha < \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$.

which is negative if and only if $c^{r-2}\alpha < \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$.

From this and the fact that $\gamma(0) > 0$, we see that if $c^{r-2}\alpha < \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$ then $\gamma(t)$ has precisely two distinct positive real roots, say T_0 and T_1 where $0 < T_0 < t_{\min} < T_1$. Coming back to the random process, this tells us that if $c^{r-2}\alpha < \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$, then we expect the number of open hyperedges to become very small as t_m approaches T_0 from the left. What we will do in this case is track our variables until $\gamma(t) < \zeta$ where ζ is a constant chosen small with respect to r, c and α ; the value of ζ is given in Definition 7.3. At this point we can use other methods to verify that, with high probability, percolation does not occur.

On the other hand, if $c^{r-2}\alpha > \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$, then $\gamma(t_{\min}) > 0$ and $\gamma(t)$ has no positive real roots. Since $p + qt_m = o(1)$ for all values of m that we consider in the first phase (see (2.10)), we expect our supply of open hyperedges not to run out (in fact, when $t_m > t_{\min}$, we expect the number of open hyperedges to be typically increasing). What we will do in this case is track the above variables until step $N \left\lfloor \frac{\log(N)}{\alpha} \right\rfloor = O(N \log(d))$, at which point the number of open hyperedges will be large enough that we can deduce that percolation occurs with high probability. See Figure 5.1 for examples of how the trajectory of $Q(m)$ depends on the relationship between p and q .

We briefly discuss why we have chosen to focus on values of p and q of order $d^{-1/(r-1)}$. As we have argued above, if the infection at time m resembles a uniform infection of density $p + qt_m$, then we expect the variable $Q(m)$ to be roughly

$$N[d(p + qt_m)^{r-1} - t_m].$$

The nice thing about considering p and q of order $d^{-1/(r-1)}$ is that the expression inside the square brackets becomes a function of t_m only. Thus, as long as t_m is bounded by a constant,

open hyperedges are both being created and discarded at a ‘constant rate’ independent of d .

As we said above, the main aim of the first phase is to show that for $0 \leq m \leq M$, $Q(m)$ is within a small error term of $\gamma(t_m) \cdot N$. The following function describes the ‘relative error’ that we will allow ourselves in these bounds,

$$\epsilon(t) := \frac{(t+1)^{K/10}}{\log^{K/5}(d)}. \tag{2.9}$$

Note that $\epsilon(t) = o(1)$ when $t = O(\log(d))$. In what follows, we write an interval of the form $[(1-\epsilon)g(t), (1+\epsilon)g(t)]$ as $(1 \pm \epsilon)g(t)$ for brevity.

To summarise, we track the process for M steps, where

$$M := \begin{cases} N \left\lfloor \frac{\log(N)}{\alpha} \right\rfloor & \text{if } c^{r-2}\alpha > \frac{(r-2)^{r-2}}{(r-1)^{r-1}}, \\ \min\{m \geq 0 : (1 + \epsilon(t_m))\gamma(t) < \zeta\} & \text{otherwise,} \end{cases} \tag{2.10}$$

Define

$$T := M/N. \tag{2.11}$$

Remark 2.12. Observe that, by definition of M , on the interval $[0, T]$ the function $\gamma(t)$ is always bounded away from zero by a function of r, c and α .

We are now ready to formally state the bounds we will prove on $Q(m)$ in the first phase.

Theorem 2.13. *With high probability the following statement holds. For all $0 \leq m \leq M$,*

$$Q(m) \in (1 \pm 4\epsilon(t_m))\gamma(t_m) \cdot N.$$

One way to prove that $Q(m)$ is controlled in this way, or indeed to prove bounds for any of our variables, involves determining their expected and maximum change (conditioned on what has previously occurred during the process) at each time step and applying a martingale concentration inequality. Before thinking in more detail about the expected change of $Q(m)$ we introduce some notation that will be helpful.

For each $v \in V(\mathcal{H}) \setminus I(m)$ we write

$$Q_v(m) := \{e \in Q(m) : v \in e \setminus I(m)\}. \tag{2.14}$$

For $u \neq v \in V(\mathcal{H})$, the sets $Q_u(m)$ and $Q_v(m)$ are disjoint. Note that,

$$Q(m) = \bigcup_{v \in V(\mathcal{H}) \setminus I(m)} Q_v(m).$$

Let us now think about the expected change of $Q(m)$. Firstly, which open hyperedges from $Q(m)$ are not present in $Q(m+1)$? At each step, the hyperedge e we sample from

$\mathcal{H}(m)$ is deleted and is not present in $\mathcal{H}(m+1)$. Also, with probability q , the unique healthy vertex v of e becomes infected and so all the hyperedges in $Q(m)$ whose unique healthy vertex is v are no longer open. This results in a loss of $(Q_v(m) - 1)$ hyperedges (in addition to e). Now let us consider how we gain a new open hyperedge. This occurs when some hyperedge e is successfully sampled, and the vertex v of e that becomes infected is contained in a hyperedge e' with exactly $r - 2$ infected vertices. The hyperedge e' will now be open.

Observe that for each vertex $v \in V(\mathcal{H}) \setminus I(m)$, the probability an open hyperedge containing v is sampled is $Q_v(m)/Q(m)$. Given the above discussion, we can express the expectation of $Q(m+1) - Q(m)$ conditioned on $\mathcal{H}(m)$ and $I(m)$ as

$$-1 - \sum_{v \in V(\mathcal{H}) \setminus I(m)} q \cdot \frac{Q_v(m)}{Q(m)} (Q_v(m) - 1) + \sum_{v \in V(\mathcal{H}) \setminus I(m)} q \cdot \frac{Q_v(m)}{Q(m)} Y_v^{r-2}(m), \quad (2.15)$$

where

$$Y_v^{r-2}(m) := \{e \in \mathcal{H}(m) : |e \cap I(m)| = r - 2, v \in e\}.$$

So to be able to determine (2.15) we can see that we would need to have control over $Y_v^{r-2}(m)$. So let us consider how a new copy of Y^{r-2} is created at a time step. One way a member of $Y_v^{r-2}(m+1) \setminus Y_v^{r-2}(m)$ can be created is from a pair of hyperedges $\{e_1, e_2\} \subseteq E(\mathcal{H})$ where: $v \in e_1 \setminus e_2$; e_2 is open; e_1 has exactly $r - 3$ infected vertices and intersects e_2 on its unique healthy vertex; and e_2 is successfully sampled at time m . Thus to determine the expected change of $Y_v^{r-2}(m)$, we need to also have control over this family Z of pairs. And similarly, to do this there are a number of other variables that we must keep track of.

To summarise this train of thought, to prove Theorem 2.13 we must have control over a number of families of variables. In particular, variables of the two types described above. We briefly remark that, in our proof, we do not explicitly calculate the expected change of $Q(m)$ in the manner we have alluded to above. In fact, we show that having control over a more general family of variables will imply the required bounds on $Q(m)$ in a different way (see Lemma 4.1). However to prove bounds on our other variables, we do calculate their expected and maximum changes. The point of performing this thought exercise on $Q(m)$ was to illustrate its interdependence on a number of other variables and to motivate the following discussion.

In order to describe formally the families of variables we wish to track, it is helpful to introduce a few definitions. Each variable we wish to control is of the following form. It is the number of ‘copies’ of some particular subhypergraph $\mathcal{F} \subseteq \mathcal{H}(m)$ such that these copies of \mathcal{F} are ‘rooted’ at a particular subset $S \subseteq V(\mathcal{H})$ (in the sense that these vertices are contained within the copy) and some particular vertices of these copies are contained in

$I(m)$. We begin by introducing some notation to describe the particular sets (that we call *configurations*) we are interested in counting ‘copies’ of.

Definition 2.16. A *configuration* is a triple $X = (\mathcal{F}, R, D)$, where \mathcal{F} is an r -uniform hypergraph in which every vertex is contained in at least one hyperedge and R and D are disjoint subsets of $V(\mathcal{F})$. The vertices of R are called the *roots* of X , the vertices of D are called the *marked* vertices of X , and the vertices of $V(\mathcal{F}) \setminus (D \cup R)$ are called the *neutral* vertices of X .

Now we have a good way to describe the things we are interested in counting ‘copies’ of, we will formally define what we mean by a *copy* of a configuration.

Definition 2.17. For $m \geq 0$, given a configuration $X = (\mathcal{F}, R, D)$ and a set $S \subseteq V(\mathcal{H})$, a *copy* of X in $\mathcal{H}(m)$ *rooted* at S is a subhypergraph \mathcal{F}' of $\mathcal{H}(m)$ such that there exists an isomorphism $\phi : \mathcal{F} \rightarrow \mathcal{F}'$ with $\phi(R) = S$ and $\phi(D) \subseteq I(m)$. Also define $X_S(m)$ to be the collection of copies of X in $\mathcal{H}(m)$ rooted at S . We denote $X_{\{v\}}(m)$ by $X_v(m)$ for $v \in V(\mathcal{H})$.

Take note that in $\mathcal{H}(m)$, a copy of a configuration (\mathcal{F}, R, D) can contain elements of $I(m)$ apart from those in $\phi(D)$. In particular, it is even possible for the set $\phi(R)$ to contain elements of $I(m)$ (despite the fact that R and D are disjoint).

Before discussing specific families of configurations, let us discuss heuristically how many copies we expect there to be of some fixed configuration $X = (\mathcal{F}, R, D)$ in $\mathcal{H}(m)$ rooted at $S \subseteq V(\mathcal{H})$. If \hat{X}_S is the number of copies of $(\mathcal{F}, R, \emptyset)$ rooted at S in \mathcal{H} , then if $I(m)$ is a uniformly random infection of density $p + qt_m$, we would expect

$$X_S(m) \approx \hat{X}_S \cdot (p + qt_m)^{|D|},$$

as each vertex is independently infected with probability $p + qt_m$. That is, each infected vertex contributes a factor of $\log^{O(1)}(d) \cdot d^{-\frac{1}{r-1}}$ (as $t_m = O(\log d)$).

Now heuristically, how do we bound \hat{X}_S ? In our proof, for some families of configurations we will only require an upper bound, but for some we need to be more careful and also need a lower bound. All the configurations (\mathcal{F}, R, D) we are interested in tracking during the first phase will satisfy the following properties: \mathcal{F} is connected and contains at most $r + 1$ hyperedges, no vertex of \mathcal{F} is contained in the intersection of more than two hyperedges, no vertex in the intersection of two hyperedges is a root, and $|R| \geq 1$.

So suppose X satisfies these conditions. To find a bound on \hat{X}_S , we can break \mathcal{F} up into its hyperedges e_1, \dots, e_k , where $|e_1 \cap R| \geq 1$ and each e_i intersects $\bigcup_{\ell < i} e_\ell$, and bound the number of choices for each hyperedge. Let S_1, \dots, S_k be a fixed partition of S such that $|S_i| = |R \cap e_i \setminus \bigcup_{\ell < i} e_\ell|$. We will bound the number of members of \hat{X}_S in \mathcal{H} such that

$S_i \subseteq e_i \setminus \bigcup_{\ell < j} e_\ell$. As there are $O(1)$ such partitions of S , the total number of members of \hat{X}_S will be a constant factor away from this.

First consider the number of ways to chose e_1 . By conditions (a), (b) and (c) of Definition 1.5, if $|R| = 1$ this is within $(1 \pm \log^{-K}(d)) \cdot d$, and if $|R| \geq 1$ this is at most $d^{1 - \frac{|R|-1}{r-1}} \cdot \log^{-K}(d)$. Similarly, we can then bound the number of ways to choose e_2 . By our choice of hyperedge order, e_2 intersects e_1 . Given a choice of e_1 , there are $O(1)$ ways e_2 can intersect it. Defining $b := |e_2 \cap (R \cup e_1)|$ (by assumption on hyperedge order $b \geq 1$), by conditions (a) and (c) of Definition 1.5 there are at most

$$O(\Delta_b(\mathcal{H})) = O\left(d^{1 - \frac{b-1}{r-1}}\right) = O\left(d^{\frac{r-b}{r-1}}\right)$$

choices for e_2 . When $2 \leq b \leq r-1$, using condition (c) of Definition 1.5 gives a stronger bound of at most

$$O(\Delta_b(\mathcal{H})) = O\left(d^{\frac{r-b}{r-1}} \log^{-K}(d)\right)$$

choices for e_2 .

Given these bounds, the number of choices for e_2 can be thought of as being $O\left(d^{\frac{a}{r-1}}\right)$, where a is the number of vertices of e_2 that are not in e_1 or R (i.e. the number of ‘new’ vertices). We can bound the number of choices for e_3, \dots, e_k analogously. A more careful version of this argument will be applied later to give the bound in Lemma 2.26.

So heuristically, for most configurations X , up to a $\log^{O(1)}(d)$ factor we generally expect there to be about $d^{\frac{|V(\mathcal{F})| - |R| - |D|}{r-1}}$ copies of X rooted at S in $\mathcal{H}(m)$. One way of thinking about this is to imagine each hyperedge contributes a factor of d , but for each vertex that is either in the intersection of two hyperedges or not neutral we lose a factor of $d^{\frac{1}{r-1}}$ (up to some powers of $\log d$). Alternately, (again up to some powers of $\log d$) we get a factor of $d^{\frac{1}{r-1}}$ for each neutral vertex in the configuration. It will be helpful to bear this rough heuristic in mind throughout the calculations which come later.

We do care about having some control over the powers of $\log d$, as we will see when we discuss the precise configurations we wish to control. We remark that in practice, bounding the number of copies of a configuration in this manner (by adding in the hyperedges one by one and bounding them individually) will not always give a good enough bound for our purposes. This is why we need to track configurations consisting of multiple hyperedges.

We now introduce our most important family of configurations, the ‘ Y ’ configurations. These are a generalisation of the two variables Y_v^{r-2} and Z discussed above. The control we have over these more general variables in $\mathcal{H}(m)$ dictates the bounds we can prove on $Q(m)$ (see Lemma 4.1) and on the Y configurations in $\mathcal{H}(m+1)$. The two further sets of variables we will discuss below (see Definitions 2.24 and 2.28) do affect how the Y configurations

behave, but due to the codegree conditions on \mathcal{H} (see Definition 1.5), we can ensure that they only contribute lower order terms.

In general, the Y configurations consist of a hyperedge e containing a root and a fixed number of marked vertices, with ‘open’ hyperedges that are disjoint from one another and only intersect e on their unique unmarked vertex. See Figure 5.2 for a visualisation of some of these configurations in the case $r = 4$. See also Figure 5.3 for some examples of **copies** of Y configurations in the case $r = 6$.

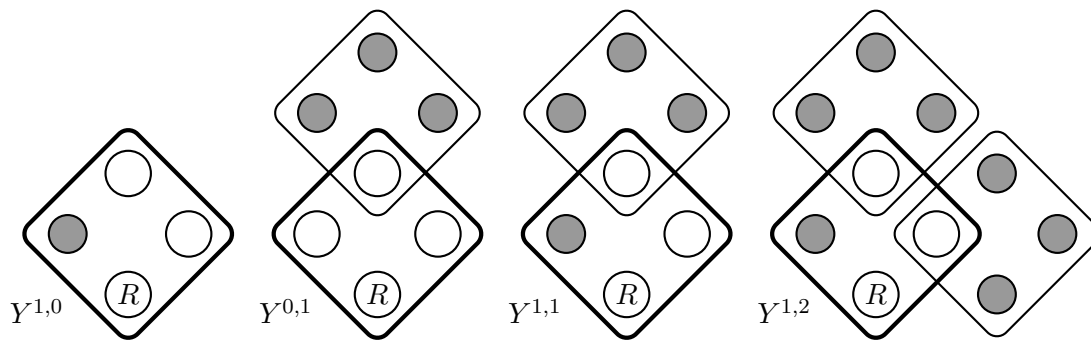


Figure 5.2: Some examples of Y configurations in the case $r = 4$. The marked vertices are shaded, the neutral vertices are white, the root is labelled with an R and the central hyperedge is drawn with a thick outline.

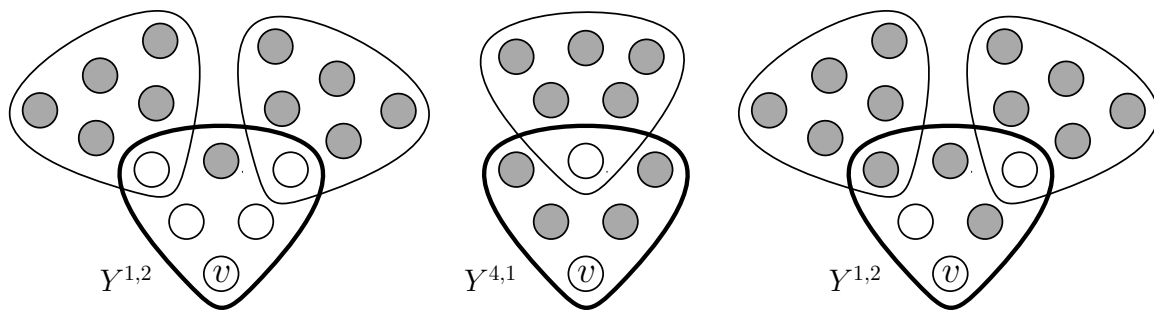


Figure 5.3: Some examples of **copies** of Y configurations rooted at v in the case $r = 6$. The infected vertices are shaded, the healthy vertices are white and the central hyperedge is drawn with a thick outline. Each copy is labelled by which configuration it is a copy of. Notice that a copy of a configuration could have more infections than marked vertices in the configuration itself, as is demonstrated by the third example.

We now formally define the family of Y configurations.

Definition 2.18. For non-negative integers i and j such that $i + j \leq r - 1$, let $Y^{i,j}$ denote the configuration (\mathcal{F}, R, D) such that \mathcal{F} is a hypergraph containing a hyperedge e , called the *central* hyperedge, such that

- (a) e contains exactly i marked vertices,
- (b) there is a unique root and e is the only hyperedge of \mathcal{F} containing the root,
- (c) \mathcal{F} has exactly j non-central hyperedges,
- (d) for each non-central hyperedge e' we have $|e \cap e'| = 1$ and the unique element of $e \cap e'$ is neutral,
- (e) every vertex of $V(\mathcal{F}) \setminus e$ is marked, and
- (f) no two non-central hyperedges intersect one another.

Observe that for $v \in V(\mathcal{H}) \setminus I(m)$, we have that $Y_v^{r-1,0}(m)$ is precisely the set of open hyperedges in which v is the unique healthy vertex, the set $Q_v(m)$. Given that $Q(m) \approx \gamma(t_m) \cdot N$, if the open hyperedges were distributed uniformly among the healthy vertices, then for each $v \in V(\mathcal{H}) \setminus I(m)$ we would expect

$$Q_v(m) \approx \gamma(t_m), \quad (2.19)$$

since by Proposition 2.6, at time $t_m < T$, we have $I(m) = o(N)$ with high probability (so at time t_m there are $(1 - o(1))N$ healthy vertices). However, when $M = O(N)$, the quantity $\gamma(t_m)$ is constant, and so we cannot hope to prove that $Q_v(m)$ is concentrated around $\gamma(t_m)$. However, as we will see in a moment, we should be able to prove concentration for $Y_v^{i,j}(m)$ when $i + j \neq r - 1$. This is why we need to track the configurations $Y_v^{i,j}(m)$, and cannot just bound such things by $Y_v^{i,0}(m) \cdot \binom{r-1-i}{j} \cdot Q^j$, where Q is an upper bound on $Q_v(m)$ for all $v \in V(\mathcal{H}) \setminus I(m)$. We would not be able to get a good enough bound on Q to prove bounds as tight as we would like on $Y_v^{i,j}(m)$.

Now we discuss how we expect the variables $Y_v^{i,j}(m)$ to behave. Consider first the variable $Y_v^{i,0}(m)$ for $0 \leq i \leq r-2$ and $v \in V(\mathcal{H})$. By properties (a) and (b) of Definition 1.5, every vertex has degree between $(1 - o(1))d$ and d . Therefore if $I(m)$ is a uniformly random infection of density $p + qt_m$, then we would expect

$$Y_v^{i,0}(m) \approx \binom{r-1}{i} d(p + qt_m)^i = \binom{r-1}{i} (c + \alpha t_m)^i d^{1 - \frac{i}{r-1}}. \quad (2.20)$$

Now, let us consider $Y_v^{i,j}(m)$ for $j \geq 1$. One can express $Y_v^{i,j}(m)$ as the sum over all $\mathcal{F}' \in Y_v^{i,0}(m)$ and all subsets U of $V(\mathcal{F}') \setminus \{v\}$ with $|U| = j$ and $|(V(\mathcal{F}') \cap I(m)) \setminus U| \geq i$ of the number of ways to choose one open hyperedge rooted at each element of U in such a way that (a) no two such hyperedges intersect and (b) each of them intersects $V(\mathcal{F}')$ on exactly one vertex. Given (2.20), most copies of $Y_v^{i,0}$ rooted at v have precisely i infected vertices. We will show by bounding other configurations (see Theorem 2.27) that the vast

majority of choices of U and open hyperedges rooted at vertices of U satisfy (a) and (b). So, if $I(m)$ is a uniformly random infection of density $p + qt_m$, then using (2.19) we would expect that

$$\begin{aligned} Y_v^{i,j}(m) &\approx Y_v^{i,0}(m) \binom{r-i-1}{j} \gamma(t_m)^j \\ &\approx \binom{r-1}{i} \binom{r-1-i}{j} (c + \alpha t_m)^i \gamma(t_m)^j d^{1-\frac{i}{r-1}}. \end{aligned}$$

Before stating the bounds we wish to prove on the Y configurations, it is helpful to introduce the following notation which will be used throughout the chapter. Define

$$y_{i,0}(t) := \binom{r-1}{i} (c + \alpha t)^i, \quad (2.21)$$

and

$$y_{i,j}(t) := \binom{r-1-i}{j} y_{i,0}(t) \gamma(t)^j. \quad (2.22)$$

We will prove the following.

Theorem 2.23. *With high probability the following statement holds. For all $0 \leq m \leq M$, for all $v \in V(\mathcal{H})$, for all $0 \leq i \leq r-2$ and all $0 \leq j \leq r-1-i$,*

$$Y_v^{i,j}(m) \in (1 \pm \epsilon(t_m)) y_{i,j}(t_m) d^{1-\frac{i}{r-1}}.$$

To track the Y configurations and thus prove Theorem 2.23 we will determine the expectation of $Y_v^{i,j}(m+1) - Y_v^{i,j}(m)$ given $\mathcal{H}(m)$ and $I(m)$. Let us consider how a new copy of $Y^{i,j}$ rooted at v is created. Let $Y^{i,j} = (\mathcal{F}, R, D)$. For each new copy Y of $Y^{i,j}$ rooted at v , there exists some $u \in D$ such that Y is created from a copy of a configuration $X = (\mathcal{F}, R \cup \{u\}, D \setminus \{u\})$, such that the image of u in $\mathcal{H}(m)$ is the unique healthy vertex of a hyperedge that gets successfully sampled. So in fact it is created from a copy of configuration $X' = (\mathcal{F}', R, D')$, where \mathcal{F}' is the union of \mathcal{F} and one new hyperedge e whose intersection with \mathcal{F} contains some $u \in D$, and D' is the union of $D \setminus \{u\}$ and $e \setminus \{u\}$. See Figure 5.4 for some examples of this.

Either e intersects \mathcal{F} on precisely one vertex, or e intersects \mathcal{F} on several vertices, including vertices of D . In the first case, as we will see in Section 5.6, X' can be expressed as a combination of Y configurations. However, the family of Y configurations does not contain the type of intersections we get in the second case, so we need to control another family of variables which includes those with such intersections. In Section 5.6.3 the calculation of the expectation of $Y_v^{i,j}(m+1) - Y_v^{i,j}(m)$ given $\mathcal{H}(m)$ and $I(m)$ will be presented in full detail; for now we have motivated the introduction of our next family, the *secondary* configurations. These configurations are so called because they will not usually contribute

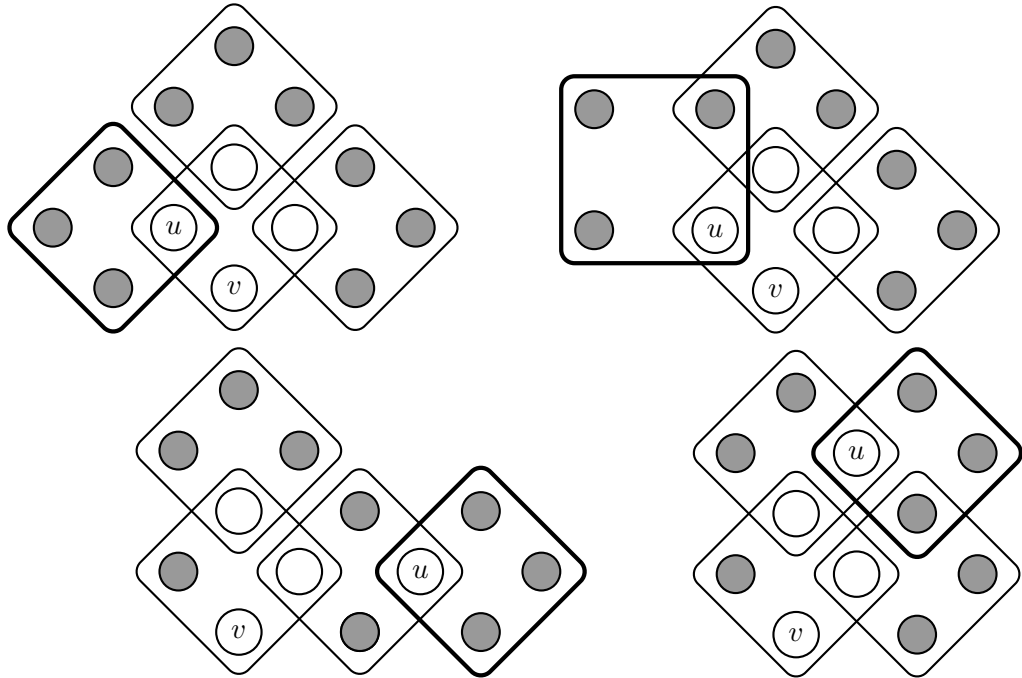


Figure 5.4: Here $r = 4$. We present four examples of ways that a copy of $Y^{1,2}$ rooted at v can be created. Such a copy is created if the open hyperedge containing u (drawn with a thick outline) is successfully sampled.

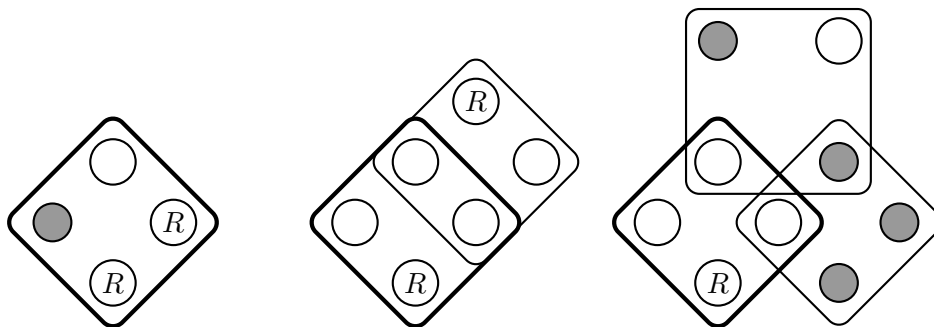


Figure 5.5: Three examples of secondary configurations X in the case $r = 4$. The marked vertices are shaded, the neutral vertices are white, the root is labelled with an R and the central hyperedge is drawn with a thick outline.

to the main order term in our calculations, but still need to be controlled. See Figure 5.5 for some examples of these configurations.

The family of secondary configurations contains configurations with multiple roots and ‘more complicated’ intersections of hyperedges than the Y configurations. Due to the codegree conditions on \mathcal{H} (see Definition 1.5), we are able to prove stronger upper bounds on secondary configurations than on Y configurations. Therefore we do not need to prove that the secondary configurations are concentrated. It will suffice to prove an upper bound that shows that they do not affect the main order term in our calculations for the Y configurations.

This will ensure that any way of creating Y configurations using secondary configurations (the second case above) is a lower order term than those terms given by the first case above. So the overall behaviour of the Y configurations is dictated purely by Y configurations. This is one reason why it is important for us to have such codegree conditions on \mathcal{H} .

Definition 2.24. Say that a configuration $X = (\mathcal{F}, R, D)$ is *secondary* if $1 \leq |E(\mathcal{F})| \leq 3$ and \mathcal{F} contains a hyperedge e called the *central* hyperedge such that

- (a) e contains at least one root and at least one neutral vertex,
- (b) for every non-central hyperedge e' we have that $e \cap e'$ contains at least one neutral vertex,
- (c) every vertex of \mathcal{F} is contained in at most two hyperedges of \mathcal{F} and
- (d) at least one of the following holds:
 - $|R| \geq 2$,
 - $|E(\mathcal{F})| = 2$ and the two hyperedges of \mathcal{F} intersect on more than one vertex, or
 - $|E(\mathcal{F})| = 3$, each non-central hyperedge intersects e on only one vertex and the two non-central hyperedges intersect one another.

Remark 2.25. If $X = (\mathcal{F}, R, D)$ is a secondary configuration and $u \in D$, then both $(\mathcal{F}, R, D \setminus \{u\})$ and $(\mathcal{F}, R \cup \{u\}, D \setminus \{u\})$ are secondary configurations.

In order to determine how we expect $X(m)$ to behave, we require the following lemma giving an upper bound on $X_S(m)$ when X is a secondary configuration with no marked vertices. This lemma will be proved in the next section and applied throughout the rest of the chapter.

Lemma 2.26. *Let $X = (\mathcal{F}, R, D)$ be a secondary configuration with $D = \emptyset$. Then, for any set $S \subseteq V(\mathcal{H})$ of cardinality $|R|$, the number of copies of X in $\mathcal{H}(m)$ rooted at S is*

$$O\left(d^{\frac{|V(\mathcal{F})|-|R|}{r-1}} \log^{-K}(d)\right).$$

We now consider what the expected number of copies of any secondary configuration $X = (\mathcal{F}, R, D)$ in $\mathcal{H}(m)$ rooted at S would be, if $I(m)$ were a uniformly random infection with density $p + qt_m$. The case $D = \emptyset$ is covered by Lemma 2.26, so now we consider the case that $D \neq \emptyset$. By Remark 2.25, the configuration $\tilde{X} := (\mathcal{F}, R, \emptyset)$ is also a secondary configuration and so, by Lemma 2.26, the expected number of copies of \tilde{X} is $O\left(d^{\frac{|V(\mathcal{F})|-|R|}{r-1}} \log^{-K}(d)\right)$. For any such copy, if $I(m)$ were a uniformly random infection with density $p + qt_m$, then the probability that every vertex of the image of D in this copy is infected would be precisely $(p + qt_m)^{|D|}$. Putting this together we get that, if $I(m)$ were uniform, then we would expect

$$X_S(m) = O\left((p + qt_m)^{|D|} d^{\frac{|V(\mathcal{F})|-|R|}{r-1}} \log^{-K}(d)\right).$$

Now we formally describe the upper bound that we will prove on $X_S(m)$ for a general secondary configuration $X = (\mathcal{F}, R, D)$ with central hyperedge e .

Theorem 2.27. *With high probability the following statement holds. For all $0 \leq m \leq M$, for all secondary configurations $X = (\mathcal{F}, R, D)$ and all $S \subseteq V(\mathcal{H})$ of cardinality $|R|$:*

$$X_S(m) \leq \log^{2|D|} r^4 (d) d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d).$$

By Definition 2.24, every secondary configuration contains at least one neutral vertex. To bound the maximum change of our variables we will also need to have control over configurations that consist of a single hyperedge with no neutral vertices. So now we will define the final set of variables we wish to control.

Definition 2.28. For $1 \leq i \leq r$, let W^i denote the configuration (\mathcal{F}, R, D) where \mathcal{F} is a hypergraph consisting of r vertices contained in a single hyperedge, $|R| = i$ and $D = V(\mathcal{F}) \setminus R$.

See Figure 5.6 for an illustration of some of these configurations in the case $r = 6$.

Of course, the variable $W_v^1(m)$ is the same as $Y_v^{r-1,0}(m)$, which is the same as $Q_v(m)$ if $v \notin I(m)$. For $i = 1$ and $S = \{v\}$, the degree of v in \mathcal{H} is at most d and so, if $I(m)$ were a uniformly random infection with density $p + qt_m$, then we would expect $W_S^1(m)$ to be at most $d \cdot (p + qt_m)^{r-1}$. For $i \geq 2$, by condition (c) of Definition 1.5, the number of hyperedges of $\mathcal{H}(m)$ containing S is at most $d^{1-\frac{i-1}{r-1}} \log^{-K}(d)$. So, if $I(m)$ were a uniformly random infection with density $p + qt_m$, then we would expect $W_S^i(m)$ to be at most

$$(p + qt_m)^{r-i} \cdot d^{1-\frac{i-1}{r-1}} \log^{-K}(d) = (p + qt_m)^{r-i} d^{\frac{r-i}{r-1}} \log^{-K}(d) = (\alpha + ct_m) \log^{-K}(d) = o(1).$$

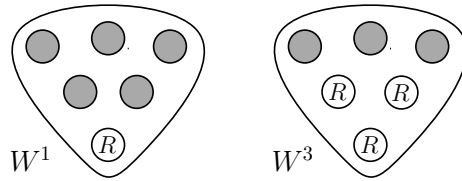


Figure 5.6: Some examples of W configurations in the case $r = 6$. The marked vertices are shaded and the roots are labelled with an R .

Here, we use the fact K is large and that we are only considering values of t up to $O(\log(d))$.

Thus we cannot hope to prove tight concentration bounds for these variables. However, we will prove the following.

Theorem 2.29. *With high probability the following statement holds. For all $1 \leq i \leq r$, for all $S \subseteq V(\mathcal{H})$ of cardinality i and all $0 \leq m \leq M$,*

$$W_S^i(m) \leq \log^{r^3(r-i)}(d).$$

We will show in Lemma 4.1 that to prove Theorem 2.13, it is sufficient to prove Theorems 2.23 and 2.29 and Proposition 2.6.

The parts of the chapter concerning the first phase are structured as follows. In Section 5.4 we show in Lemma 4.1 how Proposition 2.6 and Theorems 2.23 and 2.29 imply Theorem 2.13. Then we prove the case $m = 0$ of Theorems 2.13, 2.23, 2.27 and 2.29 in Section 5.5 using a version (Corollary 3.6) of the Kim–Vu Inequality (Theorem 3.5). The case $1 \leq m \leq M$ of Theorems 2.13, 2.23, 2.27 and 2.29 are proved in Section 5.6 using Freedman’s Inequality (Theorem 3.10) and the differential equations method. This concludes our discussion of the first phase of the proof of Theorem 1.6.

5.2.2 The Second Phase

To complete the proof of Theorem 1.6 we define a different process which involves sampling a large set of open hyperedges in each round, rather than sampling one hyperedge at a time like we did in the first phase. As a slight abuse of notation, in the second phase we let $I(0)$ denote $I(M)$ and $\mathcal{H}(0)$ denote $\mathcal{H}(M)$; that is, after the first phase, we ‘restart the clock’ from zero before running the second phase process. The second phase process will be defined slightly differently depending on whether we are in the subcritical or supercritical case.

For $m \geq 1$, in round m we will sample a set of open hyperedges and use the outcomes to define $I(m + 1)$ and $\mathcal{H}(m + 1)$. Again we will let $Q(m)$ be the set of hyperedges e of $\mathcal{H}(m)$

such that $|e \cap I(m)| = r - 1$ (the set of open hyperedges). Analogous to the first phase, for each configuration $X = (\mathcal{F}, R, D)$, $S \subseteq V(\mathcal{H})$ with $|S| = |R|$ and integer $m \geq 0$, we let $X_S(m)$ denote the set of copies of X in $\mathcal{H}(m)$ rooted at S . As before, we still write $X_S(m)$ when referring to $|X_S(m)|$ and $Q(m)$ when referring to $|Q(m)|$.

We now define the number of steps for which we will run the processes in the second phase. For $\lambda < 1/8$ depending on only r, c and α (λ is defined in (7.1)):

$$M_2 := \begin{cases} \min \left\{ m : (\log N)^{\left(\frac{3}{2}\right)^m} > d^{\frac{1}{r-1} + \frac{1}{10}} \right\} & \text{if } c^{r-2}\alpha > \frac{(r-2)^{r-2}}{(r-1)^{r-1}}, \\ 2 \lceil \log_{1/(1-\lambda)}(N) \rceil & \text{otherwise.} \end{cases} \quad (2.30)$$

5.2.3 The Subcritical Case

First, let us consider the ‘subcritical case’; i.e. when $c^{r-2}\alpha < \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$. Recall that, in this case, we track the first phase process until the number of open hyperedges is bounded above by ζN for some constant ζ chosen small with respect to r, c and α .

The Second Phase Process in the Subcritical Case. We obtain $I(m+1)$ and $\mathcal{H}(m+1)$ in the following way. For $m \geq 0$ we sample every hyperedge in $Q(m)$. We let $I(m+1)$ be the union of $I(m)$ and all of the vertices in hyperedges which were successfully sampled and let $\mathcal{H}(m+1) := \mathcal{H}(m) \setminus Q(m)$.

Our main result in the subcritical case is the following theorem, which immediately implies Theorem 1.6 in the case $c^{r-2}\alpha < \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$.

Theorem 2.31. *If $c^{r-2}\alpha < \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$, then with high probability,*

(i) $Q(M_2) = 0$, and

(ii) $|I(M_2)| = \frac{N \cdot \log^{O(1)}(N)}{d^{1/(r-1)}} = o(N)$.

The key part of the proof Theorem 2.31 is the following bound on $\mathbb{E}(Q(m))$:

$$\mathbb{E}(Q(m)) \leq (1 - \lambda)^m 2\zeta \cdot N, \quad (2.32)$$

for some $\lambda < 1/8$ depending only on r, c and α (λ is defined in (7.1)). For $m = M_2$, this bound implies that $\mathbb{E}(Q(M_2)) = o(1)$, and so (i) follows from (2.32) and Markov’s Inequality (Theorem 3.1).

To deduce (2.32), consider how a member of $Q(m+1)$ is created. (As every open hyperedge is sampled at each time step, $Q(m+1) \cap Q(m) = \emptyset$.) A member of $Q(m+1)$ is created from a hyperedge $e \in \mathcal{H}(m)$ such that $0 \leq |e \cap I(m)| \leq r - 2$, but $|e \cap I(m+1)| = r - 1$. So in particular, in round m all healthy vertices of e but one are contained in an

open hyperedge that is successfully sampled. Since $Q(m+1) \cap Q(m) = \emptyset$, e contains at least one healthy vertex that is contained in a successfully sampled hyperedge of $Q(m)$.

So a member of $Q(m+1)$ is created when we have a hypergraph consisting of:

- (1) a hyperedge $e \notin Q(m)$ such that $e \not\subset I(m)$,
- (2) a hyperedge $e_u \in Q_u(m)$, for each healthy vertex u in e except one,

and each open hyperedge chosen in (2) is successfully sampled at time m .

Given such a hypergraph, picking a hyperedge $e' \neq e$ and deleting it gives a hypergraph that may be a copy of a Y configuration (rooted at the healthy vertex of $e \cap e'$), or may not be (because of additional overlaps between the non-central hyperedges). In order to track $Q(m)$ in the second phase, we wish to track these sorts of configurations. The reason that we consider configurations of this type (where the edge e' is not present rather than with it also included), is that we wish to express $Q(m+1)$ in terms of $Q(m)$ and breaking up the configuration this way allows us to do this (see for example (2.35) below). This motivates the introduction of another family of configurations: the Z configurations.

Definition 2.33. Given $v \in V(\mathcal{H})$, $0 \leq i \leq r-2$, $0 \leq j \leq r-1-i$ and $0 \leq m \leq M_2$ let $Z_v^{i,j}(m)$ be the union of $Z_v(m)$ over all configurations $Z = (\mathcal{F}, R, D)$ such that \mathcal{F} is a hypergraph containing a hyperedge e , called the *central* hyperedge, such that

- (a) e contains exactly i marked vertices,
- (b) there is a unique root and e is the only hyperedge of \mathcal{F} containing the root,
- (c) \mathcal{F} has exactly j non-central hyperedges,
- (d) every non-central hyperedge contains exactly $r-1$ marked vertices,
- (e) for each non-central hyperedge e' , the unique neutral vertex of e' is contained in e ,
and
- (f) no neutral vertex is contained in two non-central hyperedges.

See Figure 5.7 for some examples in the case $r = 6$. These configurations can be thought of as a more general version of the Y configurations, in fact the main order term of $Z_v^{i,j}(m)$ comes from $Y_v^{i,j}(m)$. The following simple observation provides a helpful way of thinking about the Z configurations.

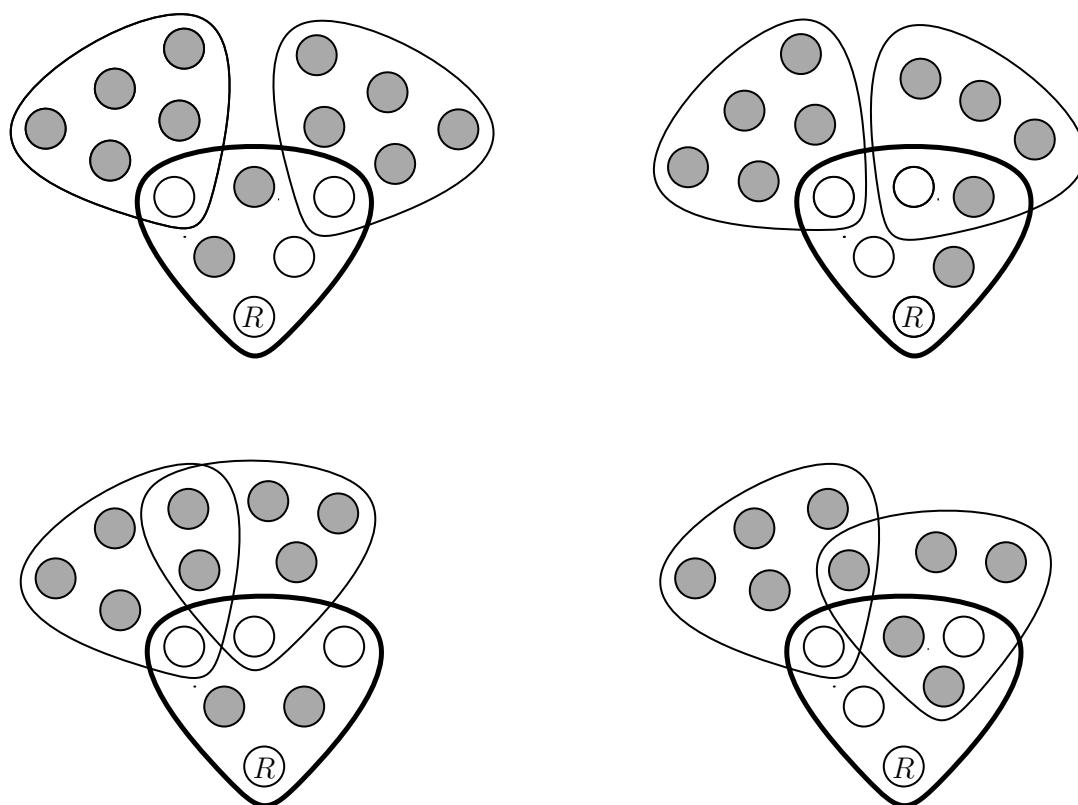


Figure 5.7: Some examples of possible members of $Z_v^{2,2}(m)$ in the case $r = 6$. The root is labelled by R , the infected vertices are shaded, the healthy vertices are white and the central hyperedge is drawn with a thick outline.

Observation 2.34. Let Z be a member of $Z_v^{i,j}(m)$. If each of the non-central hyperedges intersects the central hyperedge e on only one vertex and no pair of them intersect one another then Z is a member of $Y_v^{i,j}(m)$. Also observe that $Y_v^{i,j}(m) \subseteq Z_v^{i,j}(m)$, as every member of $Y_v^{i,j}(m)$ is a member of $Z_v^{i,j}(m)$ of this type. If one of the non-central hyperedges intersects e on more than one vertex or two of them intersect one another, then Z consists of a copy \mathcal{F}' of a secondary configuration with one root, $r - 1 - i$ neutral vertices and a set of at most $j - 1$ copies of W^1 rooted at vertices of \mathcal{F}' .

We will see in Lemma 7.9 that the members of $Z_v^{i,j}(m)$ that come from secondary configurations just contribute a lower order term. Given the above discussion, conditioned on $\mathcal{H}(m)$ and $I(m)$ the expected value of $Q(m + 1)$ is at most

$$\sum_{Q \in Q(m)} \sum_{w \in Q \setminus I(m)} \sum_{j=0}^{r-2} Z_w^{r-2-j,j}(m) q^{j+1}. \tag{2.35}$$

The proof of Theorem 2.31 comes down to proving that $Z_w^{r-2-j,j}(m)$ satisfies strong enough upper bounds (with high probability) so that evaluating this expression gives the bound in (2.32).

So to prove Theorem 2.31, it suffices to control $Z_v^{i,j}(m)$. To do this, we must also keep control over $Y_v^{i,j}(m)$, $X_S(m)$ and $W_S^i(m)$, as before. This is achieved via multiple applications of a version of the Kim–Vu Inequality (Corollary 3.6). Full details will be given in Section 5.7.

5.2.4 The Supercritical Case

The strategy in the ‘supercritical case’, i.e. when $c^{r-2}\alpha > \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$, is somewhat similar, but the details are a little different.

The Second Phase Process in the Supercritical Case. We obtain $I(m + 1)$ and $\mathcal{H}(m + 1)$ in the following way. Each round contains two steps. In the first step we choose some $Q'(m) \subseteq Q(m)$ and sample every open hyperedge of $Q'(m)$. We define $Q'(0) = Q(0)$. For $m > 0$ we choose $Q'_v(m)$ to be a subset of $Q_v(m)$ with cardinality $\min \left\{ Q_v(m), (\log N)^{\left(\frac{3}{2}\right)^m} \right\}$ and define

$$Q'(m) := \bigcup_{v \in V(\mathcal{H}) \setminus I(m)} Q'_v(m).$$

We let $I_0(m + 1)$ be the union of $I(m)$ and all of the hyperedges of $Q'(m)$ that were successfully sampled. We let $\mathcal{H}_0(m + 1) := \mathcal{H}(m) \setminus Q'(m)$.

In the second step for $i \geq 0$, if there exists a healthy vertex v contained in at least $d^{\frac{1}{r-1} + \frac{1}{10}}$ open hyperedges of $\mathcal{H}_i(m + 1)$, we sample every open hyperedge in $\mathcal{H}_i(m + 1)$ containing v .

Call this set $Q_v^i(m+1)$ and define $\mathcal{H}_{i+1}(m+1) := \mathcal{H}_i(m+1) \setminus Q_v^i(m+1)$. Let $I_{i+1}(m+1)$ be the union of $I_i(m)$ and the vertices of all of the hyperedges of $Q_v^i(m+1)$ that were successfully sampled. The second step ends when we reach j such that no healthy vertex of $\mathcal{H}_j(m+1)$ is contained in $d^{\frac{1}{r-1} + \frac{1}{10}}$ open hyperedges. We define $\mathcal{H}(m+1) := \mathcal{H}_j(m+1)$, and $I(m+1) := I_j(m+1)$.

In the second step, each $Q_v^i(m)$ is large enough such that, with high probability, v will become infected when we sample every hyperedge in $Q_v^i(m)$.

The proof of Theorem 1.6 relies on the following bound:

$$Q_v(m) \geq (\log N)^{\left(\frac{3}{2}\right)^m}, \quad (2.36)$$

for each $0 \leq m \leq M_2$ and all $v \in V(\mathcal{H}) \setminus I(m)$. To prove this bound, we use Janson's inequality for the lower tail (Theorem 3.9). Given (2.36), after at most $O(\log \log N)$ rounds, every healthy vertex has at least $d^{1/(r-1)+1/10}$ open hyperedges containing it. Now, by the Chernoff Bound (Theorem 3.2) and the fact that $N = d^{O(1)}$ (by property (b) of Definition 1.5), with high probability every vertex is infected after only one additional round. Therefore, percolation occurs with high probability. See Section 5.8 for full details. This concludes our outline of the proof.

5.3 Probabilistic Tools

Here, for convenience, we collect together the probabilistic tools we apply throughout the proof. We will also formally define the probability space we use throughout the proof.

5.3.1 Standard Concentration Inequalities

The following two theorems will be repeatedly applied throughout the proof. The first is Markov's Inequality, which is one of the simplest concentration inequalities in probability theory.

Theorem 3.1 (Markov's Inequality). *If X is a non-negative random variable and $a > 0$, then*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

The second is a version of the Chernoff Bound taken from [46].

Theorem 3.2 (The Chernoff Bound). *Let X_1, \dots, X_n be a sequence of independent $[0, 1]$ -valued random variables and let $X = \sum_{i=1}^n X_i$. Then, for $0 < \varepsilon < 1$,*

$$\begin{aligned} \mathbb{P}(X < (1 - \varepsilon)\mathbb{E}(X)) &\leq e^{-\frac{\varepsilon^2 \mathbb{E}(X)}{2}}, \\ \mathbb{P}(X > (1 + \varepsilon)\mathbb{E}(X)) &\leq e^{-\frac{\varepsilon^2 \mathbb{E}(X)}{3}}. \end{aligned}$$

5.3.2 Kim–Vu and Janson Inequalities

A central theme in the study of large deviation inequalities is that if a random variable X depends on a sequence of independent trials in which, for any outcome of the trials, changing the result of a small set of the trials does not influence the value of X too much, then X is often concentrated (see, e.g., [46, 98, 133] for further discussion). In our case, it is clear that the value of any of the variables that we track at time zero depends on the N independent random trials which determine whether or not each vertex of $V(\mathcal{H})$ is contained in $I(0)$. However, as it turns out, our codegree and neighbourhood similarity conditions (conditions (c) and (d) of Definition 1.5) are not strong enough to obtain good control over the ‘worst case’ influence of changing a small set of trials. Fortunately for us, there exist a number of different tools for obtaining strong concentration when the worst case influence is rather large, but the ‘typical’ influence is small.

The tool that we use in Section 5.5 to prove bounds on our variables when $m = 0$ is a version of the Kim–Vu Inequality [88] due to Vu [132] which is particularly well suited to our situation. Other such tools include large deviation versions of Janson’s Inequality (Theorem 3.9), which we apply in Section 5.8, and the ‘method of typical bounded differences’ developed by Warnke [133]. Before stating the Kim–Vu inequality, we require some definitions.

Definition 3.3. Let V be a finite index set and let f be a multivariate polynomial in variables $\{x_v : v \in V\}$. Given a multiset A of indices from V , let $\partial_A f$ denote the partial derivative of f with respect to the variables $\{x_v : v \in A\}$ (with multiplicity).

Definition 3.4. Let X be a random variable of the form $f(\xi_v : v \in V)$ where V is an index set with N elements, f is a multivariate polynomial of degree k in variables $\{x_v : v \in V\}$ with coefficients in $[0, 1]$ and $\{\xi_v : v \in V\}$ is a collection of independent random variables taking values in $\{0, 1\}$. For $0 \leq j \leq k$, define

$$\mathbb{E}_j(X) := \max_{|A| \geq j} \mathbb{E}(\partial_A f(\xi_v : v \in V))$$

where the maximum is taken over all multisets of indices from V with cardinality at least j .

Theorem 3.5 (Vu [132]). *There exist positive constants c_k and C_k such that if X is a random variable as in Definition 3.4 and $\mathcal{E}_0 > \mathcal{E}_1 > \dots > \mathcal{E}_k = 1$ and ℓ are positive numbers such that $\mathbb{E}_j(X) \leq \mathcal{E}_j$ for $0 \leq j \leq k$, and $\mathcal{E}_j/\mathcal{E}_{j+1} \geq \ell + j \log N$ for $0 \leq j \leq k-1$, then*

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geq \sqrt{\ell \mathcal{E}_0 \mathcal{E}_1}\right) \leq C_k e^{-c_k \ell}.$$

In order to apply Theorem 3.5 to some random variable of the form $f(\xi_v : v \in V)$, we require that the coefficients of f are in $[0, 1]$. Therefore, in practice, to apply this theorem to most of our variables we first need to rescale them, then apply the theorem, then scale them back to get the bounds we want on the original variable. For this reason, we prove a corollary to Theorem 3.5 which applies this theorem in precisely the form we will use it throughout the chapter. This should make the later calculations easier to follow.

Corollary 3.6. *Let $X = f(\xi_v : v \in V)$ be a random variable where V is an index set with N elements, f is a multivariate polynomial of degree k in variables $\{x_v : v \in V\}$ with non-negative coefficients and no variable in f has an exponent greater than 1. Let τ and $\mathcal{E}_0 \geq \log^{2k+1} N$ be positive numbers such that*

$$(i) \mathbb{E}(X) \leq \tau \cdot \mathcal{E}_0, \text{ and}$$

$$(ii) \mathbb{E}_j(X) \leq \tau, \text{ for } 1 \leq j \leq k.$$

Then for N sufficiently large, with probability at least $1 - N^{-20\sqrt{\log N}}$,

$$X \in \mathbb{E}(X) \pm O\left(\tau \log^k(N) \sqrt{\mathcal{E}_0}\right).$$

Proof. Define

$$Z = g(\xi_v : v \in V) := \tau^{-1} \cdot f(\xi_v : v \in V).$$

We show that Z is close to its expectation with high probability and use this to obtain bounds on X which hold with high probability. Using the definition of Z , from (i) we obtain that

$$\mathbb{E}(Z) = \tau^{-1} \cdot \mathbb{E}(X) \leq \mathcal{E}_0 \tag{3.7}$$

and from (ii) we obtain that for $1 \leq j \leq k$,

$$\mathbb{E}_j(Z) = \tau^{-1} \cdot \mathbb{E}_j(X) \leq 1. \tag{3.8}$$

In particular this implies that, when N is sufficiently large, every term of g has a coefficient which is at most one. Indeed, for a monomial $\prod_{u \in A} x_u$ appearing with a non-zero coefficient in g , its coefficient is precisely $\mathbb{E}(\partial_A g) \leq \mathbb{E}_{|A|}(Z)$.

Set $\ell := \log^2(N)$, $\mathcal{E}_k := 1$ and $\mathcal{E}_j := (\ell + j \log N) \mathcal{E}_{j+1}$ for $1 \leq j < k$. By hypothesis, $\mathcal{E}_0 \geq \log^{2k+1}(N)$, which is at least $(\ell + j \log N) \mathcal{E}_1$ for N sufficiently large, and so $\mathcal{E}_j / \mathcal{E}_{j+1} \geq \ell + j \log N$ for $0 \leq j \leq k - 1$.

By (3.7) and (3.8), we have $\mathbb{E}_j(Z) \leq \mathcal{E}_j$ for all $0 \leq j \leq k$. Therefore, we may apply Theorem 3.5 to obtain that

$$\mathbb{P}\left(|Z - \mathbb{E}(Z)| \geq \sqrt{\ell \mathcal{E}_0 \mathcal{E}_1}\right) \leq e^{-\Omega(\log^2 N)} \leq N^{-20\sqrt{\log N}}.$$

Since by definition $\ell \cdot \mathcal{E}_1 = O(\log^{2k}(N))$, we have $\sqrt{\ell \mathcal{E}_0 \mathcal{E}_1} = O(\log^k(N)\sqrt{\mathcal{E}_0})$. Therefore with probability at least $1 - N^{-20\sqrt{\log N}}$,

$$Z \in \mathbb{E}(Z) \pm O\left(\log^k(N)\sqrt{\mathcal{E}_0}\right).$$

Now rescaling by τ gives that with probability at least $1 - N^{-20\sqrt{\log N}}$,

$$X \in \mathbb{E}(X) \pm O\left(\tau \log^k(N)\sqrt{\mathcal{E}_0}\right),$$

as required. \square

We will also use Corollary 3.6 to prove bounds on our variables for the ‘subcritical’ case during the second phase. For the ‘supercritical’ case we will apply the following lower tail version of Janson’s Inequality.

Theorem 3.9 (Janson’s Inequality for the Lower Tail [77]). *Let \mathcal{G} be a hypergraph and, for $p \in (0, 1)$ and $e \in E(\mathcal{G})$, let I_e be the indicator variable for the event $e \subseteq V(\mathcal{G})_p$. Set*

$$X := \sum_{e \in E(\mathcal{G})} I_e,$$

$$\mu := \mathbb{E}(X), \text{ and}$$

$$\delta := \sum_{\substack{(e, e') \in E(\mathcal{G})^2 \\ e \neq e' \\ e \cap e' \neq \emptyset}} \mathbb{E}(I_e I_{e'}),$$

where the final sum is over all ordered pairs, (so each pair is counted twice). Then, for any $\varepsilon \in [0, 1]$,

$$\mathbb{P}(X \leq (1 - \varepsilon)\mathbb{E}(X)) \leq \exp(-\varphi(-\varepsilon)\mu^2/(\mu + \delta)),$$

where $\varphi(x) = (1 + x) \log(1 + x) - x$.

5.3.3 Martingales and Concentration

In Section 5.6 we will use standard martingale concentration inequalities to prove bounds on our variables throughout the first phase for $m > 0$.

Recall that a sequence $0 = B(0), B(1), \dots$ of random variables is said to be a *supermartingale* with respect to a filtration $\mathcal{F}(0), \mathcal{F}(1), \dots$ if, for all $m \geq 0$, we have that $B(m)$ is $\mathcal{F}(m)$ -measurable and

$$\mathbb{E}(B(m+1) \mid \mathcal{F}(m)) \leq B(m).$$

In what follows, when we say that $B(0), B(1), \dots$ is a supermartingale, it is always with respect to the natural filtration corresponding to our process (which will be formally defined

in the next subsection). A sequence $B(0), B(1), \dots$ is a *submartingale* if the sequence $-B(0), -B(1), \dots$ is a supermartingale. Also, a sequence $B(0), B(1), \dots$ is said to be η -*bounded* if, for all $m \geq 0$,

$$-\eta \leq B(m+1) - B(m) \leq \eta.$$

Our main tool in Section 5.6 is the following concentration inequality of Freedman [59].

Theorem 3.10 (Freedman [59]). *Let $0 := B(0), B(1), \dots$ be a η -bounded supermartingale and let*

$$V(m) := \sum_{\ell=0}^{m-1} \text{Var}(B(\ell+1) - B(\ell) \mid \mathcal{F}_\ell).$$

Then, for all $a, \nu > 0$,

$$\mathbb{P}(B(m) \geq a \text{ and } V(m) \leq \nu \text{ for some } m) \leq \exp\left(-\frac{a^2}{2(\nu + a\eta)}\right).$$

5.3.4 The Probability Space

A natural candidate for the probability space on which to view our process is

$$\Omega := \{0, 1\}^N \times E(\mathcal{H})^{|E(\mathcal{H})|} \times \mathcal{P}(E(\mathcal{H}))^{|E(\mathcal{H})|} \times \{0, 1\}^{|E(\mathcal{H})|}$$

where $\mathcal{P}(E(\mathcal{H}))$ is the collection of all subsets of $E(\mathcal{H})$. For any point in Ω , the first N coordinates determine the infection at time zero, the next $|E(\mathcal{H})|$ coordinates list the hyperedges sampled during the first phase process (although, note that the first phase stops before $|E(\mathcal{H})|$ hyperedges have been sampled), the next $|E(\mathcal{H})|$ coordinates list the sets of hyperedges sampled during the second phase process and the last $|E(\mathcal{H})|$ coordinates determine which hyperedges of \mathcal{H} are contained in \mathcal{H}_q .

One should notice that Ω contains a large number of ‘infeasible’ points (i.e. points of measure zero); for example, it contains points corresponding to evolutions of the processes in which some hyperedges are sampled more than once, or the m th hyperedge sampled in the first phase is not even chosen from $Q(m-1)$, etc. We let Ω' be the subspace of Ω consisting of only those points which have positive measure.

For $m \geq 0$, let \mathcal{F}_m be the σ -algebra generated by the partitioning of Ω' in which two points are in the same class if they correspond to evolutions of the processes which have the same initial infection and which are indistinguishable after ℓ steps of the first phase process for every ℓ in the range $1 \leq \ell \leq m$. For example, any two points of Ω' corresponding to evolutions in which the first phase process runs for fewer than m steps are in the same class if and only if they are indistinguishable at every step of the first phase. Similarly, for $m \geq 0$, let \mathcal{F}'_m be the σ -algebra generated by the partitioning of Ω' in which two points are in the same class if they are indistinguishable at every step of the first phase and, for every

$1 \leq \ell \leq m$, they are indistinguishable after the ℓ th step of the second phase process. We will work in this probability space throughout the proof without further comment.

5.4 Preliminaries

In this section, we will prove four preliminary results. First we prove Proposition 2.6, which gives a bound on the number of infected vertices at each step of the first phase. Then we deduce that to track $Q(m)$, it is enough to have control over the Y and W configurations and the number of infected vertices at time m . After that, we will prove Lemma 2.26, which bounds the number of copies of any secondary configuration $(\mathcal{F}, R, \emptyset)$ in $\mathcal{H}(m)$ rooted at S (where $|S| = |R|$), for any $0 \leq m \leq M$. At the end of the section, we will prove an analytic lemma that will be used in the application of the differential equations method in Section 5.6.

We restate here Proposition 2.6 from Section 2, to aid the reader.

Proposition 2.6 (Restated). *For $0 \leq m \leq M$, with probability at least $1 - N^{-\Omega(\sqrt{\log N})}$,*

$$I(m) = O(\log N \cdot Nd^{-1/(r-1)}).$$

Proof. The expected number of vertices which are infected at time zero is $cNd^{-1/(r-1)}$. By the Chernoff bound (Theorem 3.2) with $\epsilon = 1/2$, we have that, with probability at least $1 - e^{-\Omega(Nd^{-1/(r-1)})}$, there are at most $\frac{3c}{2}Nd^{-1/(r-1)}$ vertices infected at time zero.

At each step, one hyperedge is sampled and becomes infected with probability q . As the total number of hyperedges sampled is $M = O(\log N \cdot N)$, the expected number of vertices infected by successfully sampling an open hyperedge is $O\left(\frac{N \cdot \log(N)}{d^{1/(r-1)}}\right)$.

Applying the Chernoff bound with $\epsilon = 1/2$, we get that with probability at least $1 - N^{-\Omega(\sqrt{\log N})}$ there are at most $O\left(\frac{N \cdot \log(N)}{d^{1/(r-1)}}\right)$ vertices infected by successfully sampling an open hyperedge. By choosing K large enough and (2.1), we have $1 - e^{-\Omega(Nd^{-1/(r-1)})} \geq 1 - N^{-10\sqrt{\log N}}$. So with probability at least $1 - N^{-10\sqrt{\log N}}$,

$$I(M) = \frac{3c}{2}Nd^{-1/(r-1)} + O\left(\frac{N \cdot \log(N)}{d^{1/(r-1)}}\right) = O\left(\frac{N \cdot \log(N)}{d^{1/(r-1)}}\right).$$

As $I(\ell) \leq I(M)$ for any $\ell \leq M$, this completes the proof. \square

As mentioned in Section 5.2, to prove Theorem 2.13, it is sufficient to prove Theorems 2.23 and 2.29 and Proposition 2.6. More formally:

Lemma 4.1. *If for every choice of $0 \leq m \leq M$, $v \in V(\mathcal{H})$, $0 \leq i \leq r-2$ and $0 \leq j \leq r-1$ we have:*

$$(i) Y_v^{i,j}(m) \in (1 \pm \epsilon(t_m)) y_{i,j}(t_m) d^{1 - \frac{i}{r-1}},$$

$$(ii) W_v^1(m) \leq \log^{r^4}(d), \text{ and}$$

$$(iii) I(m) = O(\log N \cdot N d^{-1/(r-1)}),$$

then for every $0 \leq m \leq M$,

$$Q(m) \in (1 \pm 4\epsilon(t_m)) \gamma(t_m) \cdot N.$$

Proof. The sum

$$S := \sum_{w \in V(\mathcal{H}) \setminus I(m)} (r-1) Q_w(m) (Y_w^{0,0}(m) - Y_w^{1,0}(m)) \quad (4.2)$$

counts the number of ways of choosing

- (1) a vertex $w \in V(\mathcal{H}) \setminus I(m)$,
- (2) a hyperedge $e' \in Q_w(m)$,
- (3) a hyperedge $e \in \mathcal{H}(m)$ disjoint from $I(m)$ containing w , and
- (4) a vertex $v \in e \setminus \{w\}$.

Viewing v as the root, such a configuration is a copy of $Y^{0,1}$ rooted at v but not a copy of $Y^{1,1}$. Define $S_2 := \sum_{v \in V(\mathcal{H}) \setminus I(m)} (Y_v^{0,1}(m) - Y_v^{1,1}(m))$. S_2 counts the number of ways of choosing

- (1) a vertex $v \in V(\mathcal{H}) \setminus I(m)$,
- (2) a hyperedge $e \in \mathcal{H}(m)$ containing v and at most one infected vertex,
- (3) a vertex $w \in e \setminus \{v\}$ such that if $e \cap I(m) \neq \emptyset$ then $w \in I(m)$, and
- (4) a hyperedge $e' \in W_w^1(m)$,

which contains everything that S counts. So $S \leq S_2$. The configurations counted by S_2 but not S are those given by choosing

- (1) a vertex $v \in V(\mathcal{H}) \setminus I(m)$,
- (2) a hyperedge $e \in \mathcal{H}(m)$ containing v such that e has a unique infected vertex w ,
- (3) a hyperedge $e' \in W_w^1(m)$.

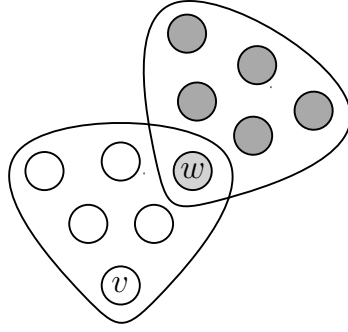


Figure 5.8: Here $r = 6$. Copies of this hypergraph are counted by S only when w is healthy. Copies are counted by S_2 whether w is infected or not. The (unlabelled) vertices shaded dark grey are infected, the unshaded vertices are healthy.

See Figure 5.8 for an illustration of the difference between what is counted in S and S_2 when $r = 6$.

Applying hypotheses (i) and (ii) to bound $Y_v^{1,0}(m)$ and $W_w^1(m)$ for all $v \in V(\mathcal{H}) \setminus I(m)$ and $w \in I(m)$, we see that $S_2 - S$ is bounded above by

$$\sum_{v \in V(\mathcal{H}) \setminus I(m)} (Y_v^{1,0}(m) - Y_v^{2,0}(m)) \log^{r^4}(d) = N \cdot d^{\frac{r-2}{r-1}} \log^{O(1)}(d).$$

Thus,

$$S \in S_2 \pm N \cdot d^{1 - \frac{1}{r-1}} \log^{O(1)}(d).$$

Using the definition of S_2 and applying the bounds on $Y_v^{1,1}(m)$ given by the hypotheses of the lemma, we have

$$S \in \sum_{v \in V(\mathcal{H}) \setminus I(m)} Y_v^{0,1}(m) \pm N \cdot d^{1 - \frac{1}{r-1}} \log^{O(1)}(d). \quad (4.3)$$

Using the bounds on $Y_v^{1,0}(m)$ and (for $w \in V(\mathcal{H}) \setminus I(m)$) the bounds on $Q_w(m) = W_w^1(m)$ given by the hypotheses of the lemma, from (4.2) we get

$$S \in \sum_{w \in V(\mathcal{H}) \setminus I(m)} (r-1)Q_w(m)Y_w^{0,0}(m) \pm N \cdot d^{1 - \frac{1}{r-1}} \log^{O(1)}(d). \quad (4.4)$$

Combining (4.4) and (4.3) gives

$$\sum_{w \in V(\mathcal{H}) \setminus I(m)} (r-1)Q_w(m)Y_w^{0,0}(m) \in \sum_{v \in V(\mathcal{H}) \setminus I(m)} Y_v^{0,1}(m) \pm N \cdot d^{1 - \frac{1}{r-1}} \log^{O(1)}(d).$$

We also have by hypothesis that

$$\sum_{w \in V(\mathcal{H}) \setminus I(m)} (r-1)Q_w(m)Y_w^{0,0}(m) \in \sum_{w \in V(\mathcal{H}) \setminus I(m)} (r-1)Q_w(m) (1 \pm \epsilon(t_m)) d$$

$$= Q(m)(r-1)(1 \pm \epsilon(t_m))d$$

and

$$\sum_{v \in V(\mathcal{H}) \setminus I(m)} Y_v^{0,1}(m) \in |V(\mathcal{H}) \setminus I(m)|(1 \pm \epsilon(t_m))(r-1)\gamma(t_m) \cdot d.$$

Putting all this together gives

$$\begin{aligned} Q(m) &\in \frac{|V(\mathcal{H}) \setminus I(m)|(1 \pm \epsilon(t_m))(r-1)\gamma(t_m) \pm d^{-\frac{1}{r-1}}N \log^{O(1)}d}{(r-1)(1 \mp \epsilon(t_m))} \\ &\subseteq N(1 \pm 4\epsilon(t_m))\gamma(t_m) \end{aligned}$$

since $\frac{1 \pm x}{1 \mp x} \in 1 \pm 3x$ for x sufficiently small and $I(m) = O(\log N \cdot Nd^{-1/(r-1)})$ by hypothesis (iii). The result follows. \square

We will now present the proof of Lemma 2.26. It may be helpful to first recall the definition of a secondary configuration from Definition 2.24. We restate here the result from Section 2 to aid the reader.

Lemma 2.26 (Restated). *Let $X = (\mathcal{F}, R, D)$ be a secondary configuration with $D = \emptyset$. Then, for any set $S \subseteq V(\mathcal{H})$ of cardinality $|R|$, the number of copies of X in $\mathcal{H}(m)$ rooted at S is*

$$O\left(d^{\frac{|V(\mathcal{F})| - |R|}{r-1}} \log^{-K}(d)\right).$$

Proof. Let $X = (\mathcal{F}, R, D)$ be a secondary configuration. First we see that, for any ordering $e_1, \dots, e_{|E(\mathcal{F})|}$ of the hyperedges of \mathcal{F} , we have that

$$|V(\mathcal{F})| = |R| + \sum_{i=1}^{|E(\mathcal{F})|} \left(r - \left| e_i \cap \left(R \cup \bigcup_{j=1}^{i-1} e_j \right) \right| \right). \quad (4.5)$$

To see this, count the number of vertices by first counting the vertices of R and then, for each i in turn, count the vertices of e_i which have not yet been counted. This will be used several times in the calculations below.

Our goal is to bound the number of copies of X in $\mathcal{H}(m)$ rooted at a set $S \subseteq V(\mathcal{H})$ of cardinality $|R|$. By construction, $\mathcal{H}(m)$ is a subhypergraph of $\mathcal{H}(0) = \mathcal{H}$. Therefore, it suffices to upper bound the number of copies of X in $\mathcal{H}(0) = \mathcal{H}$ rooted at S . We do this in the way we described in the previous section, by breaking \mathcal{F} up into individual hyperedges and bounding the number of ways to choose each one individually, given the previous choices. We will consider a number of different cases.

First, suppose there exists a hyperedge $e_1 \in E(\mathcal{F})$ such that $|e_1 \cap R| \geq 2$. By definition of a secondary configuration, every hyperedge of a secondary configuration contains a neutral

vertex, so $e_1 \not\subseteq R$ and $|e_1 \cap R|$ is between 2 and $r-1$. Thus, by condition (c) of Definition 1.5 the number of hyperedges f_1 in \mathcal{F} intersecting S in exactly $|e_1 \cap R|$ vertices is

$$\binom{|S|}{|e_1 \cap R|} \cdot \Delta_{|e_1 \cap R|}(\mathcal{H}) = O\left(d^{\frac{r-|e_1 \cap R|}{r-1}} \log^{-K}(d)\right).$$

Note that this bound is already enough to complete the proof in the case $|E(\mathcal{F})| = 1$ since, by definition of a secondary configuration, the unique hyperedge of \mathcal{F} contains at least two roots. So, in what follows, we may assume that $|E(\mathcal{F})| \geq 2$.

Let e_2 be a hyperedge which intersects e_1 (which exists by definition of a secondary configuration) and, if $|E(\mathcal{F})| \geq 3$, then let e_3 be the remaining hyperedge. The number of copies of X rooted at S is at most the number of ways to choose a hyperedge f_1 of \mathcal{H} intersecting S on exactly $|e_1 \cap R|$ vertices, a hyperedge f_2 intersecting $S \cup f_1$ on exactly $|e_2 \cap (R \cup e_1)|$ vertices and, if $|E(\mathcal{F})| \geq 3$, a hyperedge f_3 intersecting $S \cup f_1 \cup f_2$ on exactly $|e_3 \cap (R \cup e_1 \cup e_2)|$ vertices. Using the bound that we have already proven for the number of ways of choosing f_1 , we get that this is

$$O\left(\prod_{i=1}^{|E(\mathcal{F})|} \Delta_{|e_i \cap (R \cup \bigcup_{j=1}^{i-1} e_j)|}(\mathcal{H})\right) = O\left(\log^{-K}(d) \cdot \prod_{i=1}^{|E(\mathcal{F})|} d^{\frac{r-|e_i \cap (R \cup \bigcup_{j=1}^{i-1} e_j)|}{r-1}}\right)$$

By (4.5), the exponent of d in the above expression is precisely $\frac{|V(\mathcal{F})| - |R|}{r-1}$, and so we are done when there exists some hyperedge e_1 such that $|e_1 \cap R| \geq 2$.

So from now on we assume that every hyperedge of \mathcal{F} contains at most one root. In particular, by definition of a secondary configuration, the central hyperedge has exactly one root.

Now, let e_1 denote the central hyperedge and suppose that there is a non-central hyperedge e_2 such that $e_2 \subseteq R \cup e_1$. Then, since e_2 contains at most one root, we must have that $|e_1 \cap e_2| = r-1$ and that the unique vertex of $e_2 \setminus e_1$ is a root. The vertex of $e_1 \setminus e_2$ is also a root because, by definition of a secondary configuration, e_1 contains a root and this root cannot be contained in e_2 (as every hyperedge contains at most one root). By condition (d) of Definition 1.5, the number of ways to choose two vertices x, y of S and two hyperedges f_1 and f_2 of \mathcal{H} such that $f_1 \Delta f_2 = \{x, y\}$ is

$$O(d \cdot \log^{-K}(d)).$$

If e_1 and e_2 are the only two hyperedges of \mathcal{F} , then $|V(\mathcal{F})| - |R| = r-1$ and so this bound is what we wanted to prove. If $|E(\mathcal{F})| \geq 3$, then there are

$$O\left(\Delta_{|e_3 \cap (R \cup e_1 \cup e_2)|}(\mathcal{H})\right) = O\left(d^{\frac{r-|e_3 \cap (R \cup e_1 \cup e_2)|}{r-1}}\right)$$

ways to choose a third hyperedge to form a copy of \mathcal{F} . Combining this with the bound on the number of ways to choose the first two hyperedges and applying (4.5) gives the desired bound.

So every non-central hyperedge contains at least one non-root vertex which is not contained in the central hyperedge. We can now conclude the proof in the case $|E(\mathcal{F})| = 2$. Indeed, let e_1 be the central hyperedge and e_2 be the non-central hyperedge. We can bound the number of copies of X by

$$O(\Delta(\mathcal{H}) \cdot \Delta_{|e_2 \cap (e_1 \cup R)|}(\mathcal{H})).$$

By definition of a secondary configuration and the fact that e_1 has at most one root, we know that $|e_2 \cap (e_1 \cup R)|$ must be at least two. Also, it is at most $r - 1$ by the result of the previous paragraph. So, by condition (c) of Definition 1.5, we get an upper bound of

$$O\left(d \cdot d^{\frac{r - |e_2 \cap (e_1 \cup R)|}{r-1}} \cdot \log^{-K}(d)\right)$$

which, by (4.5) and the fact that $|e_1 \cap R| = 1$, is the desired bound.

It remains to consider the case $|E(\mathcal{F})| = 3$. Let e_1 be the central hyperedge and let e_2 and e_3 be the other two hyperedges. Suppose that $|R| \geq 2$. Then, since e_1 contains exactly one root, there must be a non-central hyperedge, say e_2 , such that $e_2 \setminus e_1$ contains a root. We can now bound the number of copies of X by

$$O(\Delta(\mathcal{H}) \cdot \Delta_{|e_2 \cap (R \cup e_1)|}(\mathcal{H}) \cdot \Delta_{|e_3 \cap (R \cup e_1 \cup e_2)|}(\mathcal{H})).$$

Since $|e_2 \cap (R \cup e_1)|$ is at least two and at most $r - 1$, this is bounded above by

$$O\left(d \cdot d^{\frac{r - |e_2 \cap (R \cup e_1)|}{r-1}} \cdot d^{\frac{r - |e_3 \cap (R \cup e_1 \cup e_2)|}{r-1}} \cdot \log^{-K}(d)\right)$$

and so, in this case, we are again done by (4.5) and the fact that $|e_1 \cap R| = 1$.

Thus, there is exactly one root and it is contained in e_1 . By definition of a secondary configuration, this implies that $e_1 \cap e_2 \cap e_3 = \emptyset$ and that e_2 intersects e_3 . In particular, it implies that $|e_3 \cap (e_1 \cup e_2)| \geq 2$. We assume that e_2 was chosen to be the non-central hyperedge such that $|e_1 \cap e_2|$ is maximised. As above, the number of copies of X is bounded above by

$$O(\Delta(\mathcal{H}) \cdot \Delta_{|e_2 \cap e_1|}(\mathcal{H}) \cdot \Delta_{|e_3 \cap (e_1 \cup e_2)|}(\mathcal{H}))$$

which gives the desired bound by condition (c) of Definition 1.5 unless $|e_2 \cap e_1| = 1$ and $|e_3 \cap (e_1 \cup e_2)| = r$ (since we already know that $|e_3 \cap (e_1 \cup e_2)| \geq 2$). So, we assume that \mathcal{F} satisfies these conditions. By our choice of e_2 , we also get that $|e_3 \cap e_1| = 1$ as well. The last case to consider is therefore when e_2 and e_3 each intersect e_1 on a single vertex (where

these two vertices are distinct) and $|e_2 \cap e_3| = r - 1$. In this case, the number of copies of X is bounded above by the number of ways to choose a hyperedge f_1 containing S , choose two vertices x, y of $f_1 \setminus S$ and then choose two hyperedges f_2, f_3 such that $f_2 \triangle f_3 = \{x, y\}$. By condition (d) of Definition 1.5, this is bounded above by

$$O(d^2 \log^{-K}(d)),$$

which is what we needed since $|V(\mathcal{F})| - |R| = 2(r - 1)$ in this case. This completes the proof. \square

In our application of the differential equations method in Section 5.6, it is often useful for us to approximate certain sums by a related integral. For this, we use the following simple lemma. We remark that a very similar statement is derived in [92] using the same proof.

Lemma 4.6. *For $T > 0$, let $s(t)$ be a function which is differentiable and has bounded derivative on $[0, T]$. Then, for non-negative integers N and m such that $m \leq TN$, we have*

$$\left| \int_0^{m/N} s(t) dt - \frac{1}{N} \sum_{i=0}^{m-1} s(i/N) \right| \leq \frac{m \cdot \sup_{t \in [0, T]} |s'(t)|}{2N^2}.$$

Proof. Let $a, b \in [0, T]$ with $a \leq b$. As for all $t \in [a, b]$, $s(t) \leq s(a) + (t - a) \sup_{t \in [a, b]} |s'(t)|$, we have that

$$\begin{aligned} \int_a^b s(t) dt &\leq \int_a^b \left(s(a) + (t - a) \sup_{x \in [a, b]} |s'(x)| \right) dt \\ &= (b - a)s(a) + \left(\frac{(b - a)^2}{2} \right) \sup_{t \in [a, b]} |s'(t)| \end{aligned}$$

and, similarly,

$$\int_a^b s(t) dt \geq (b - a)s(a) - \left(\frac{(b - a)^2}{2} \right) \sup_{t \in [a, b]} |s'(t)|.$$

So, setting $a = i/N$ and $b = (i + 1)/N$ for $0 \leq i \leq m - 1$, we obtain

$$\left| \int_{i/N}^{(i+1)/N} s(t) dt - \frac{s(i/N)}{N} \right| \leq \left(\frac{1}{2N^2} \right) \sup_{t \in [0, T]} |s'(t)|.$$

Summing up these expressions and applying the triangle inequality, we have

$$\left| \int_0^{m/N} s(t) dt - \sum_{i=0}^{m-1} \frac{s(i/N)}{N} \right| \leq \left(\frac{m}{2N^2} \right) \sup_{t \in [0, T]} |s'(t)|$$

as desired. \square

5.5 Concentration at Time Zero

Our goal in this section is to prove Theorems 2.23, 2.27 and 2.29 in the case $m = 0$. Theorem 2.13 will follow from Theorems 2.23 and 2.29 and Proposition 2.6 via Lemma 4.1. In fact, we will actually prove the following stronger bounds in order to give ourselves some ‘extra room’ in the next section.

Lemma 5.1. *With probability at least $1 - N^{-10\sqrt{\log(N)}}$ the following statement holds. For each $1 \leq i \leq r$ and set $S \subseteq V(\mathcal{H})$, we have*

$$W_S^i(0) \leq \log^{2r}(d).$$

Lemma 5.2. *With probability at least $1 - N^{-10\sqrt{\log(N)}}$ the following statement holds. For every secondary configuration $X = (\mathcal{F}, R, D)$ and set $S \subseteq V(\mathcal{H})$, we have*

$$X_S(0) \leq d^{\frac{|V(\mathcal{F})| - |R| - |D|}{r-1}} \log^{-4K/5}(d)$$

Recall the definitions of $y_{i,0}(t)$ and $y_{i,j}(t)$ from (2.21) and (2.22). We will prove the following.

Lemma 5.3. *With probability at least $1 - N^{-10\sqrt{\log(N)}}$ the following statement holds. For every pair of non-negative integers i and j such that $i \leq r - 2$ and $i + j \leq r - 1$ and any vertex $v \in V(\mathcal{H})$, we have*

$$Y_v^{i,j}(0) \in \left(1 \pm \log^{-3K/10}(d)\right) y_{i,j}(0) d^{1 - \frac{i}{r-1}}.$$

Note that we get the following concentration result for $Q(0)$ from Lemmas 5.1, 5.3 and Proposition 2.6 via Lemma 4.1.

Lemma 5.4. *With probability at least $1 - N^{-9\sqrt{\log(N)}}$, we have*

$$Q(0) \in (1 \pm 4\epsilon(0))\gamma(0) \cdot N.$$

We will prove Lemmas 5.1, 5.2 and 5.3 by applying Corollary 3.6. Although the Y configurations are arguably the most important, we save proving Lemma 5.3 until last; the proofs of the first two lemmas involve more simple applications of Corollary 3.6 and hence provide a more gentle introduction for the reader to the style of arguments we will be using throughout the section. We remark that in the proof of Lemma 5.1, we technically do not need to rescale our random variable, and so could apply Theorem 3.5 directly. However it is marginally simpler to apply Corollary 3.6, so this is what we shall do.

We will use the following random variables throughout the rest of the section. Given $w \in V(\mathcal{H})$, let ξ_w be the Bernoulli random variable which is equal to one if and only if $w \in I(0)$. Without further ado we present the proofs of Lemmas 5.1, 5.2 and 5.3.

Proof of Lemma 5.1. Let $S \subseteq V(\mathcal{H})$ be a set of cardinality i . If $i = r$, then clearly $W_S^i(0) \leq 1$ and so we may assume that $1 \leq i \leq r - 1$. Observe that $W_S^i(0)$ can be written as $f(\xi_v : v \in V(\mathcal{H}))$ where

$$f(x_v : v \in V(\mathcal{H})) := \sum_{\substack{e \in E(\mathcal{H}) \\ S \subseteq e}} \left(\prod_{v \in e \setminus S} x_v \right).$$

Observe that no variable in f has an exponent greater than 1 and the degree of f is $r - i$. We wish to apply Corollary 3.6 to obtain an upper bound on $W_S^i(0)$ which holds with high probability. In order to do this, we must bound $\mathbb{E}_j(W_S^i(0))$ for $0 \leq j \leq r - i$.

Let A be a set of at most $r - i$ vertices of \mathcal{H} disjoint from S . We have

$$\partial_A f(x_v : v \in V(\mathcal{H})) = \sum_{\substack{e \in E(\mathcal{H}) \\ A \cup S \subseteq e}} \left(\prod_{v \in e \setminus (S \cup A)} x_v \right).$$

Therefore, by linearity of expectation and independence we have

$$\mathbb{E}(\partial_A f(x_v : v \in V(\mathcal{H}))) = \sum_{\substack{e \in E(\mathcal{H}) \\ A \cup S \subseteq e}} p^{r-i-|A|}.$$

In the case that $|A| = r - i$, the above expression is simply equal to 0 or 1 (depending on whether $A \cup S$ is a hyperedge of \mathcal{H} or not). Otherwise,

$$\begin{aligned} \mathbb{E}(\partial_A f(\xi_v : v \in V(\mathcal{H}))) &= \sum_{\substack{e \in E(\mathcal{H}) \\ A \cup S \subseteq e}} p^{r-i-|A|} \leq \Delta_{|A|+i}(\mathcal{H}) \cdot p^{r-i-|A|} \\ &= \Delta_{|A|+i}(\mathcal{H}) c^{r-i-|A|} d^{\frac{|A|+i-1}{r-1}-1}. \end{aligned}$$

By conditions (a) and (c) of Definition 1.5, this expression is $o(1)$ if $|A| + i \geq 2$ and is at most c^{r-1} otherwise (i.e. if $A = \emptyset$ and $i = 1$).

This analysis gives

$$\mathbb{E}(W_S^i(0)) \leq c^{r-1}$$

and for $1 \leq j \leq r - i$,

$$\mathbb{E}_j(W_S^i(0)) \leq 1.$$

Set $\tau := 1$ and $\mathcal{E}_0 := \log^{2r} N$. As $\mathbb{E}(W_s(0)) = O(1)$ and

$$\tau \log^{r-i} N \sqrt{\mathcal{E}_0} = o(\log^{2r}(d)),$$

applying Corollary 3.6 gives that with probability at least $1 - N^{-20\sqrt{\log N}}$,

$$W_S^i(0) \leq c^{r-1} + o(\log^{2r}(d)) < \log^{2r}(d).$$

The result now follows by taking a union bound over all values of i and all subsets of $V(\mathcal{H})$ of cardinality i . \square

Before proving Lemmas 5.2 and 5.3 it is helpful to introduce the following definition and simple claim.

Definition 5.5. Let $X = (\mathcal{F}, R, D)$ be a configuration. Let \mathcal{T}_X be the collection of all pairs (\mathcal{F}', D') , with \mathcal{F}' a subhypergraph of \mathcal{H} and $D' \subseteq V(\mathcal{F}')$, such that there exists an isomorphism ϕ from \mathcal{F} to \mathcal{F}' such that $\phi(R) = S$ and $\phi(D) = D'$.

Claim 5.6. Let $X = (\mathcal{F}, R, D)$ be a configuration. Let $S \subseteq V(\mathcal{H})$ be a set of cardinality $|R|$. Then $X_S(0)$ is bounded above by $f(\xi_v : v \in V(\mathcal{H}))$, where

$$f(x_v : v \in V(\mathcal{H})) := \sum_{(\mathcal{F}', D') \in \mathcal{T}_X} \left(\prod_{v \in D'} x_v \right).$$

Proof. Let $\mathcal{F}' \subseteq \mathcal{H}(0)$. By Definition 2.17, \mathcal{F}' is a copy of X rooted at S in $\mathcal{H}(0)$ only if there exists an isomorphism $\phi : \mathcal{F} \rightarrow \mathcal{F}'$ with $\phi(R) = S$ and $D' := \phi(D) \subseteq I(0)$. For such a ϕ and D' , say that (\mathcal{F}', ϕ, D') is a *witness triple* for \mathcal{F}' .

The number of copies of X rooted at S in $\mathcal{H}(0)$ is at most the number of $\mathcal{F}' \subseteq \mathcal{H}(0)$ such that there exists some ϕ and D' where (\mathcal{F}', ϕ, D') is a witness triple for \mathcal{F}' . This is at most the number of pairs (\mathcal{F}', D') in \mathcal{T}_X such that $D' \subseteq I(0)$.

Therefore $X_S(0)$ is bounded above by $f(\xi_v : v \in V(\mathcal{H}))$, where

$$f(x_v : v \in V(\mathcal{H})) := \sum_{(\mathcal{F}', D') \in \mathcal{T}_X} \left(\prod_{v \in D'} x_v \right),$$

as required. □

Observation 5.7. Let \mathcal{F}' be a copy of $X = (\mathcal{F}, R, D)$ rooted at S in $\mathcal{H}(0)$. Let us consider how \mathcal{F}' may be counted multiple times by $f(\xi_v : v \in V(\mathcal{H}))$. This will happen precisely when there exist two witness triples (defined in the proof above) for \mathcal{F}' of the form $(\mathcal{F}', \phi_1, D_1)$ and $(\mathcal{F}', \phi_2, D_2)$ such that $D_1 \neq D_2$ (and so $\phi_1 \neq \phi_2$).

When $|I(0) \cap \mathcal{F}'| = |D|$, there is only one choice for the set D' in a witness triple and so no such pair $(\mathcal{F}', \phi_1, D_1)$ and $(\mathcal{F}', \phi_2, D_2)$ exists. However, if $|I(0) \cap \mathcal{F}'| > |D|$, then there may exist subsets $D_1 \neq D_2$ of \mathcal{F}' (and isomorphisms ϕ_1 and ϕ_2) such that $(\mathcal{F}', \phi_1, D_1)$ and $(\mathcal{F}', \phi_2, D_2)$ are both witness triples for X . In this case both (\mathcal{F}', D_1) and (\mathcal{F}', D_2) are in \mathcal{T}_X and \mathcal{F}' is counted multiple times by $f(\xi_v : v \in V(\mathcal{H}))$.

So the difference between $X_S(0)$ and $f(\xi_v : v \in V(\mathcal{H}))$ is at most $O(1)$ times the number of copies of configurations $X' = (\mathcal{F}, R, D')$, where $D' := D \cup \{u\}$, for some $u \in V(\mathcal{F}) \setminus (D \cup R)$.

We will now return to proving Lemmas 5.2 and 5.3.

Proof of Lemma 5.2. Let $X = (\mathcal{F}, R, D)$ be a secondary configuration and let $S \subseteq V(\mathcal{H})$ be a set of cardinality $|R|$. If $D = \emptyset$, $X_S(0)$ is simply bounded above by the number of copies of X in \mathcal{H} rooted at S . By Lemma 2.26, this is $O\left(d^{\frac{|V(\mathcal{F})|-|R|}{r-1}} \log^{-K}(d)\right)$, and this bound is actually stronger than we need. So, from now on, we assume that $D \neq \emptyset$.

By Claim 5.6, setting $\mathcal{T} := \mathcal{T}_X$ gives that the variable $X_S(0)$ is bounded above by $f(\xi_v : v \in V(\mathcal{H}))$ where

$$f(x_v : v \in V(\mathcal{H})) := \sum_{(\mathcal{F}', D') \in \mathcal{T}} \left(\prod_{v \in D'} x_v \right).$$

Observe that the degree of f is $|D|$ and no variable in f has an exponent greater than 1.

We wish to apply Corollary 3.6 with

$$\tau := d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-9K/10}(d)$$

and $\mathcal{E}_0 := \log^{2|D|+1}(N)$ to obtain an upper bound for $X_S(0)$ which holds with high probability. As above, in order to apply Corollary 3.6 we must bound $\mathbb{E}_j(X_S(0))$ for $0 \leq j \leq |D|$.

If A contains an element of S , then $\partial_A f = 0$. On the other hand, if A is a subset of $V(\mathcal{H}) \setminus S$ of cardinality at most $|D|$, then

$$\partial_A f(x_v : v \in V(\mathcal{H})) = \sum_{\substack{(\mathcal{F}', D') \in \mathcal{T} \\ A \subseteq D'}} \left(\prod_{v \in D' \setminus A} x_v \right)$$

and so, by linearity of expectation and independence,

$$\mathbb{E}(\partial_A f(\xi_v : v \in V(\mathcal{H}))) = \sum_{\substack{(\mathcal{F}', D') \in \mathcal{T} \\ A \subseteq D'}} p^{|D' \setminus A|}. \quad (5.8)$$

Recalling Remark 2.25, we see that the number of $(\mathcal{F}', D') \in \mathcal{T}$ with $A \subseteq D'$ is at most the sum of $X'_{S \cup A}(0)$ over all secondary configurations X' with $|V(\mathcal{F})|$ vertices, $|R| + |A|$ roots and zero marked vertices multiplied by a constant factor (as there is a choice for which $|A|$ roots are in D'). So, by Lemma 2.26, we get that the right side of (5.8) is bounded above by

$$\begin{aligned} & O\left(d^{\frac{|V(\mathcal{F})|-|R|-|A|}{r-1}} \log^{-K}(d) p^{|D|-|A|}\right) \\ &= O\left(d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-K}(d)\right) = O\left(\tau \log^{-K/10}(d)\right) = o(\tau). \end{aligned}$$

So for $0 \leq j \leq |D|$,

$$\mathbb{E}_j(X_S(0)) = o(\tau).$$

As X is secondary, $|D| \leq 3r - 1$. So for K large with respect to r ,

$$\begin{aligned} \tau \log^{|D|}(N) \sqrt{\mathcal{E}_0} &\leq d^{\frac{|V(\mathcal{F})| - |D| - |R|}{r-1}} \log^{6r}(d) \log^{-9K/10}(d) \\ &= o\left(d^{\frac{|V(\mathcal{F})| - |D| - |R|}{r-1}} \log^{-4K/5}(d)\right). \end{aligned}$$

Using this and the fact that $\mathbb{E}(X_S(0)) = o(\tau)$, applying Corollary 3.6 gives that

$$X_S(0) = o\left(d^{\frac{|V(\mathcal{F})| - |D| - |R|}{r-1}} \log^{-4K/5}(d)\right).$$

with probability at least $1 - N^{-20\sqrt{\log(N)}}$. The result follows by taking a union bound over all secondary configurations and choices of S . \square

5.5.1 Proof of Lemma 5.3

First, note that it suffices to consider the case that i and j are not both zero, since $Y_v^{0,0}(0) = \deg(v)$ and so the bounds hold for $Y_v^{0,0}(0)$ by conditions (a) and (b) of Definition 1.5. Thus, from now on, we assume $i + j \geq 1$.

Write the configuration $Y^{i,j}$ as (\mathcal{F}, R, D) . By Claim 5.6, setting $\mathcal{T} := \mathcal{T}_{Y^{i,j}}$ gives that the variable $Y_v^{i,j}(0)$ is bounded above by $\tilde{Y}_v^{i,j} := f(\xi_u : u \in V(\mathcal{H}))$ where

$$f(x_u : u \in V(\mathcal{H})) := \sum_{(\mathcal{F}', D') \in \mathcal{T}} \left(\prod_{u \in D'} x_u \right).$$

Note that by definition of $Y^{i,j}$, f has degree $i + j(r - 1)$. Observe that no variable in f has an exponent greater than 1.

We will prove the following.

Proposition 5.9. *For each $v \in V(\mathcal{H})$, $0 \leq i \leq r - 2$ and $0 \leq j \leq r - 1 - i$ such that $i + j \geq 1$, with probability $1 - N^{-20\sqrt{\log(N)}}$ we have*

$$\tilde{Y}_v^{i,j} \in \left(1 \pm o\left(\log^{-3K/10}(d)\right)\right) y_{i,j}(0) d^{1 - \frac{i}{r-1}}.$$

We now show that $\tilde{Y}_v^{i,j}$ is a good approximation for $Y_v(0)$, and hence it suffices to prove Proposition 5.9.

Proposition 5.10. *If for all $v \in V(\mathcal{H})$, $0 \leq i \leq r - 2$ and $0 \leq j \leq r - 1 - i$ such that $i + j \geq 1$,*

$$\tilde{Y}_v^{i,j} \in \left(1 \pm o\left(\log^{-3K/10}(d)\right)\right) y_{i,j}(0) d^{1 - \frac{i}{r-1}},$$

and for all $w \in V(\mathcal{H})$,

$$W_w^1(0) \leq \log^{2r}(d),$$

then

$$Y_v^{i,j}(0) \in \left(1 \pm \log^{-3K/10}(d)\right) y_{i,j}(0) d^{1 - \frac{i}{r-1}}.$$

The proof of Lemma 5.3 follows from Propositions 5.9 and 5.10 and Lemma 5.1 by applying the union bound.

Proof of Proposition 5.10. Fix $v \in V(\mathcal{H})$. By Observation 5.7, it may be the case that $\tilde{Y}_v^{i,j}$ counts an element \mathcal{F}' of $Y_v^{i+1,0}(0)$ more than once if it contains more than $i + (r - 1)j$ elements of $I(0)$. However, we have

$$\bar{Y}_v^{i,j} := \tilde{Y}_v^{i,j} - Y_v^{i,j}(0) = O \left(\sum_{\mathcal{F}'' \in Y_v^{i+1,0}(0)} \sum_{\substack{U \subseteq V(\mathcal{F}'') \setminus \{v\} \\ |U|=j}} \prod_{u \in U} W_u^1(0) \right). \quad (5.11)$$

We now prove that

$$\bar{Y}_v^{i,j} = o \left(d^{1 - \frac{i}{r-1}} \log^{-3K/10}(d) \right),$$

from which the claim follows. If $i = r - 2$, then $Y_v^{i+1,0}(0) = W_v^1(0)$, so by hypothesis (using the fact that $j < r$) we have $\bar{Y}_v^{i,j} = O \left(\log^{2r^2}(d) \right)$. Otherwise, for $i < r - 2$ we have by hypothesis that

$$Y_v^{i+1,0}(0) \leq \tilde{Y}_v^{i+1,0} = O \left(d^{1 - \frac{i+1}{r-1}} \right).$$

Using this, as $j < r$, from (5.11) we get

$$\begin{aligned} \bar{Y}_v^{i,j} &= O \left(\sum_{\mathcal{F}'' \in Y_v^{i+1,0}(0)} \sum_{\substack{U \subseteq V(\mathcal{F}'') \setminus \{v\} \\ |U|=j}} \prod_{u \in U} W_u^1(0) \right) \\ &= O \left(Y_v^{i+1,0}(0) \log^{2rj}(d) \right) \\ &= O \left(d^{1 - \frac{i+1}{r-1}} \log^{2r^2}(d) \right) \\ &= o \left(d^{1 - \frac{i}{r-1}} \log^{-3K/10}(d) \right), \end{aligned}$$

as required. The claim follows. □

It remains to prove Proposition 5.9.

Proof of Proposition 5.9. Fix $v \in V(\mathcal{H})$. We wish to apply Corollary 3.6 with

$$\tau := d^{1 - \frac{i}{r-1}} \log^{-9K/10}(d)$$

and $\mathcal{E}_0 := \log^K(d)$ to obtain bounds on $\tilde{Y}_v^{i,j}$ that hold with high probability. As before, in order to apply Corollary 3.6, we must bound $\mathbb{E}(\tilde{Y}_v^{i,j})$ for $0 \leq \ell \leq i + (r - 1)j$.

Claim 5.12. *We have*

$$\mathbb{E}(\tilde{Y}_v^{i,j}) = (1 \pm 3r \log^{-K}(d)) y_{i,j}(0) d^{1-\frac{i}{r-1}},$$

and for $1 \leq \ell \leq i + (r-1)j$ we have

$$\mathbb{E}_\ell(\tilde{Y}_v^{i,j}) = o(\tau).$$

Proof. If A contains v , then $\partial_A f = 0$. On the other hand, if A is a subset of $V(\mathcal{H}) \setminus \{v\}$ of cardinality at most $i + (r-1)j$, then

$$\partial_A f(x_u : u \in V(\mathcal{H})) = \sum_{\substack{(\mathcal{F}', D') \in \mathcal{T} \\ A \subseteq D'}} \left(\prod_{u \in D' \setminus A} x_u \right)$$

and so, by linearity of expectation and independence,

$$\mathbb{E}(\partial_A f(\xi_u : u \in V(\mathcal{H}))) = \sum_{\substack{(\mathcal{F}', D') \in \mathcal{T} \\ A \subseteq D'}} p^{|D' \setminus A|}. \quad (5.13)$$

Now let us bound the number of $(\mathcal{F}', D') \in \mathcal{T}$ with $A = \emptyset$. This is at most the number of ways to:

- (1) choose a hyperedge $e \in \mathcal{H}$ such that $v \in e$,
- (2) choose vertices $\{v_1, \dots, v_i\}$ in $e \setminus \{v\}$,
- (3) choose vertices $\{w_1, \dots, w_j\}$ in $e \setminus \{v, v_1, \dots, v_i\}$,
- (4) choose hyperedges e_1, \dots, e_j , sequentially, such that $e_k \cap (e \cup \bigcup_{a < k} e_a) = w_k$.

By condition (a) of Definition 1.5, there are $(1 \pm \log^{-K}(d))d$ choices of e for (1). Given the choice of e , by conditions (b) and (c) of Definition 1.5, there are $(1 \pm \log^{-K}(d))d$ choices for each of e_1, \dots, e_j . Therefore

$$\begin{aligned} |\mathcal{T}| &\in (1 \pm \log^{-K}(d))^{j+1} \binom{r-1}{i} \binom{r-i-1}{j} d^{j+1} \\ &\subseteq (1 \pm 3r \log^{-K}(d)) \binom{r-1}{i} \binom{r-i-1}{j} d^{j+1} \end{aligned}$$

Therefore, applying (5.13) to the case $A = \emptyset$, we have

$$\begin{aligned} \mathbb{E}(\tilde{Y}_v^{i,j}) &\in (1 \pm 3r \log^{-K}(d)) \binom{r-1}{i} \binom{r-i-1}{j} d^{j+1} p^{i+j(r-1)} \\ &= (1 \pm 3r \log^{-K}(d)) y_{i,j}(0) d^{1-\frac{i}{r-1}}. \end{aligned}$$

Now, if $A \neq \emptyset$, then the number of $(\mathcal{F}', D') \in \mathcal{T}$ with $A \subseteq D'$ is at most the number of ways to partition A into $j+1$ sets A_0, \dots, A_j (some of which may be empty) and do the following:

- (1) choose a hyperedge e such that $e \cap (A \cup \{v\}) = A_0 \cup \{v\}$,
- (2) choose a subset $W = \{w_1, \dots, w_j\}$ of $e \setminus (A_0 \cup \{v\})$,
- (3) for $k = 1, \dots, j$, choose a hyperedge e_k of \mathcal{H} containing $A_k \cup \{w_k\}$.

The number of such partitions of A is $O(1)$. For any such partition A_0, \dots, A_j , by condition (c) of Definition 1.5 and the fact that $|A| \geq 1$, the number of elements of \mathcal{T} generated by the above procedure is at most

$$O\left(\prod_{k=0}^j \Delta_{|A_k|+1}(\mathcal{H})\right) = O\left(d^{j+1-\frac{|A|}{r-1}} \log^{-K}(d)\right)$$

Combining this with (5.13), we get that

$$\begin{aligned} \mathbb{E}(\partial_A f(\xi_u : u \in V(\mathcal{H}))) &= O\left(d^{j+1-\frac{|A|}{r-1}} \log^{-K}(d) p^{i+j(r-1)-|A|}\right) \\ &= O\left(d^{1-\frac{i}{r-1}} \log^{-K}(d)\right) = O\left(\tau \log^{-K/10}(d)\right) = o(\tau). \end{aligned}$$

This completes the proof of Claim 5.12. □

The degree of f can be crudely bounded by r^2 . So as for K sufficiently large with respect to r ,

$$\tau \log^{r^2} N \sqrt{\mathcal{E}_0} = o\left(d^{1-\frac{i}{r-1}} \log^{-3K/10}\right),$$

using the value of $\mathbb{E}(\tilde{Y}_v^{i,j})$ given by Claim 5.12 and applying Corollary 3.6 gives that with probability at least $1 - N^{-20\sqrt{\log N}}$ we have

$$\tilde{Y}_v^{i,j}(0) \in \left(1 \pm o\left(\log^{-3K/10}(d)\right)\right) y_{i,j}(0) d^{1-\frac{i}{r-1}}.$$

This completes the proof of Proposition 5.9. □

Proposition 5.9 was the final piece for the proof of Lemma 5.3. Thus this completes the proof of Lemma 5.3 and hence our analysis of $m = 0$ is now complete.

5.6 The First Phase After Time Zero

In this section, we will use the differential equations method to prove Theorems 2.23, 2.27 and 2.29 for $1 \leq m \leq M$, where M is defined in (2.10).

Definition 6.1. For $0 \leq m \leq M$, let \mathcal{B}_m be the event (in Ω' , which was defined in Subsection 5.3.4) that there exists $0 \leq \ell \leq m$ such that, for some $v \in V(\mathcal{H})$ or $S \subseteq V(\mathcal{H})$, one of the following four statements fails to hold.

(B.1) For all $0 \leq i \leq r-2$ and $0 \leq j \leq r-1-i$:

$$Y_v^{i,j}(\ell) \in (1 \pm \epsilon(t_\ell)) y_{i,j}(t_m) \cdot d^{1-\frac{i}{r-1}}.$$

(B.2) For any secondary configuration $X = (\mathcal{F}, R, D)$,

$$X_S(\ell) \leq \log^{2|D|r^4}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d).$$

(B.3) For all $1 \leq i \leq r$,

$$W_S^i(\ell) \leq \log^{r^3(r-i)}(d).$$

(B.4) $I(m) = O(\log N \cdot Nd^{-1/(r-1)})$.

Note that these are precisely the bounds we wish to prove for Theorems 2.29, 2.27 and 2.23 and Proposition 2.6. One should think of \mathcal{B}_m as the event that one of the variables that we are tracking has ‘strayed’ far from its expected trajectory at or before the m th step.

In this section, we prove the following theorem.

Theorem 6.2.

$$\mathbb{P}(\mathcal{B}_M) \leq N^{-2\sqrt{\log(N)}}.$$

Observe that Theorems 2.23, 2.27 and 2.29 all follow immediately from Theorem 6.2. By Lemma 4.1, when $\omega \notin \mathcal{B}_M$, we obtain

$$Q(m) \in (1 \pm 4\epsilon(t_m))\gamma(t_m) \cdot N. \quad (6.3)$$

so Theorem 2.13 is implied as well.

Now, given a point $\omega \in \mathcal{B}_M$, let

$$\mathcal{J} = \mathcal{J}(\omega) := \min\{i : \omega \in \mathcal{B}_i\}. \quad (6.4)$$

That is, ω corresponds to a trajectory of the process in which at least one of the variables strays far from its expectation at step \mathcal{J} but not before. Define \mathcal{Z} to be the set of all $\omega \in \mathcal{B}_M$ such that (B.4) is violated at time $\mathcal{J}(\omega)$ for some set $S \subseteq V(\mathcal{H})$. Similarly, define \mathcal{Y} , \mathcal{X} and \mathcal{W} to be the events that (B.1), (B.2) and (B.3) are respectively violated at time $\mathcal{J}(\omega)$. By definition,

$$\mathcal{B}_M = \mathcal{W} \cup \mathcal{X} \cup \mathcal{Y} \cup \mathcal{Z}.$$

Our goal is to show that the probability of each of the events \mathcal{W} , \mathcal{X} , \mathcal{Y} and \mathcal{Z} is small, from which Theorem 6.2 will follow.

Getting a sufficient bound on $\mathbb{P}(\mathcal{Z})$ follows directly from Proposition 2.6. We obtain the following.

Lemma 6.5.

$$\mathbb{P}(\mathcal{Z}) \leq N^{-5\sqrt{\log N}}.$$

Proof. The event \mathcal{Z} is contained within the event that there exists some $0 \leq \ell \leq M$, such that the bound

$$I(\ell) = O(\log N \cdot Nd^{-1/(r-1)})$$

fails to hold. By Proposition 2.6, the result follows. \square

We devote the rest of the section to proving that each of \mathcal{W} , \mathcal{X} and \mathcal{Y} occurs with probability at most $N^{-\Omega(\sqrt{\log N})}$. To prove this, we will apply the differential equations method and Theorem 3.10. The sequences of variables that we track are not themselves supermartingales or submartingales and so we cannot apply Theorem 3.10 to them directly. What we do is show that the difference between each variable in the sequence and its expected trajectory, plus or minus some appropriate (growing) error function, is bounded above by an η -bounded supermartingale and below by an η -bounded submartingale (actually we only need to bound $W_S^i(m)$ and $X_S(m)$ from above). As in many applications of the differential equations method, the trick to verifying that these ‘bounding’ sequences are indeed η -bounded sub- or supermartingales is to define them in such a way that, if none of our sequences have strayed far from their expected trajectory, then we can use the fact that they have not strayed to prove that the properties hold and, otherwise, the properties hold for trivial reasons.

As in Section 5.5, despite the Y configurations being the most important, we first consider the W configurations, then the X configurations, then the Y . This is because the proofs of their respective lemmas increase in complexity and it is helpful for the reader to first see a more simple application of Theorem 3.10, before diving into the proof of Lemma 6.29. In the proof of Lemma 6.29, we need to be more careful than for Lemmas 6.19 and 6.6. This is because we are proving that the Y configurations are tightly concentrated, whereas we just prove weak upper bounds on the X and W configurations.

5.6.1 Tracking the W Configurations

Roughly speaking, our goal is to determine the probability that (B.3) is the first bound to be violated. More precisely, to determine the probability that (B.3) is violated at the first time any of (B.1), (B.2), (B.3) and (B.4) are violated. We will prove the following.

Lemma 6.6.

$$\mathbb{P}(\mathcal{W}) \leq N^{-4\sqrt{\log N}}.$$

When $i = r$, (B.3) cannot be violated, so we assume $i < r$. For $0 \leq i \leq r - 1$ and $S \subseteq V(\mathcal{H})$ with $|S| = i$, define $\mathcal{W}(S)$ to be the set of all $\omega \in \mathcal{B}_M$ such that the bound

$$W_S^i(m) \leq \log^{r^3(r-i)}(d) \quad (6.7)$$

is violated at time $m = \mathcal{J}(\omega)$, where $\mathcal{J}(\omega)$ is defined in (6.4). Observe that the bound (6.7) is precisely the bound (B.3) for our fixed choices of i and S . It follows that

$$\mathcal{W} = \bigcup_{\substack{S \subseteq V(\mathcal{H}) \\ 1 \leq |S| \leq r-1}} \mathcal{W}(S).$$

Therefore the following proposition will imply the lemma, via an application of the union bound over all choices of i and S .

Proposition 6.8. *For all $0 \leq i \leq r - 1$ and $S \subseteq V(\mathcal{H})$ with $|S| = i$,*

$$\mathbb{P}(\mathcal{W}(S)) \leq N^{-6\sqrt{\log N}}.$$

Proof. For $\ell \geq 0$ define \mathcal{E}_ℓ to be the event that $W_S^i(\ell + 1) \geq W_S^i(\ell)$. We remark that it is possible for $W_S^i(\ell)$ to decrease in a step. This will happen if a copy of W^i rooted at S is an open hyperedge which is successfully sampled. However, our choice of martingale will reflect the fact that we are only concerned with proving an upper bound on $W_S^i(\ell)$.

Given an event E , we let $\mathbb{1}_E$ denote the indicator function of E . For $\ell \geq 0$, define

$$a_\ell := \frac{\alpha y_{r-2,1}(t_\ell) (1 + \epsilon(t_\ell)) + \alpha \log^{-K/2}(d)}{N \gamma(t_\ell) (1 - 4\epsilon(t_\ell))}$$

and

$$A_S(\ell) := \begin{cases} \mathbb{1}_{\mathcal{E}_\ell} (W_S^i(\ell + 1) - W_S^i(\ell)) - a_\ell & \text{if } \omega \notin \mathcal{B}_\ell, \\ 0 & \text{otherwise,} \end{cases}$$

where $y_{r-2,1}(t)$ is defined in (2.22) and $\gamma(t)$ is defined in (2.8). Also, set

$$B_S(m) := \sum_{\ell=0}^{m-1} A_S(\ell).$$

Note that, by definition, $B_S(0) = 0$. Also, if $\omega \notin \mathcal{B}_{m-1}$, then

$$W_S^i(m) \leq B_S(m) + W_S^i(0) + \sum_{\ell=0}^{m-1} a_\ell. \quad (6.9)$$

Therefore, to obtain an upper bound on $W_S^i(m)$, it suffices to bound the three quantities on the right side of this expression. Since $\epsilon(t_\ell) = o(1)$ for all $\ell \leq M$ and $\gamma(t)$ is bounded

away from zero (by Remark 2.12) the sum can be bounded above in the following way:

$$\begin{aligned} \sum_{\ell=0}^{m-1} a_\ell &\leq \sum_{\ell=0}^{m-1} \frac{2\alpha y_{r-2,1}(t_\ell)}{N\gamma(t_\ell)} \\ &= 2\alpha(r-1) \left(\frac{1}{N} \sum_{\ell=0}^{m-1} (c + \alpha t_\ell)^{r-2} \right) \\ &= o\left(\log^{r^3(r-i)}(d)\right) \end{aligned}$$

as $m \leq M \leq \frac{N \log(N)}{\alpha} = O(N \log(d))$ by (2.2). Still assuming $\omega \notin \mathcal{B}_{m-1}$, by the above analysis and (6.9), if $W_S^i(0) \leq \log^{2r}(d) = o\left(\log^{r^3(r-i)}(d)\right)$ we have

$$W_S^i(m) \leq B_S(m) + o\left(\log^{r^3(r-i)}(d)\right).$$

It follows that the event $\mathcal{W}(S)$ is contained within the event that either $W_S^i(0) > \log^{2r}(d)$ or that $B_S(m) > \frac{1}{2} \log^{r^3(r-i)}(d)$ for some $0 \leq m \leq M$.

By Lemma 5.1,

$$\mathbb{P}\left(W_S^i(0) > \log^{2r}(d)\right) \leq N^{-10\sqrt{\log N}},$$

so to prove Proposition 6.8 it suffices to show that $B_S(m)$ is unlikely to be large. We will show that

$$\mathbb{P}\left(B_S(m) > \frac{1}{2} \log^{r^3(r-i)}(d) \text{ for some } 0 \leq m \leq M\right) \leq N^{-6\sqrt{\log N}}. \quad (6.10)$$

We wish to apply Theorem 3.10 to the sequence $B_S(0), \dots, B_S(M)$. In order to do this, we must show that $B_S(0), \dots, B_S(M)$ is an η -bounded supermartingale and we must also bound the sum $\sum_{\ell=0}^{m-1} \text{Var}(A_S(\ell) \mid \mathcal{F}_\ell)$.

Claim 6.11. $B_S(0), \dots, B_S(M)$ is a supermartingale.

Proof. This is equivalent to showing that, for $0 \leq \ell \leq M-1$, the expectation of $A_S(\ell)$ given \mathcal{F}_ℓ is non-positive. For $\omega \in \mathcal{B}_\ell$ we have $A_S(\ell) = 0$, and so it suffices to consider $\omega \notin \mathcal{B}_\ell$. That is, we can assume that none of the variables that we track has ‘strayed’ at or before time ℓ . The only hyperedges e which can be counted by $W_S^i(\ell+1)$ but not by $W_S^i(\ell)$ are those which contain S and have the property that there is a unique vertex $x \in e \setminus S$ such that $x \notin I(\ell)$. Also, such a hyperedge e contributes to $W_S^i(\ell+1) - W_S^i(\ell)$ if and only if an open hyperedge $e^* \neq e$ containing x is successfully sampled at the $(\ell+1)$ th step. Let \mathcal{T} be the set of all such pairs (e, e^*) . As the probability that a particular open hyperedge is successfully sampled is $\frac{q}{Q(\ell)}$ and as $\omega \notin \mathcal{B}_\ell$, using (6.3) we have

$$\begin{aligned} \mathbb{E}\left[\mathbb{1}_{\mathcal{E}_\ell}(W_S^i(\ell+1) - W_S^i(\ell)) \mid \mathcal{F}_\ell\right] &\leq \sum_{(e, e^*) \in \mathcal{T}} \frac{q}{Q(\ell)} \\ &\leq \frac{\alpha |\mathcal{T}|}{N\gamma(t_\ell)(1 - 4\epsilon(t_\ell))d^{1/(r-1)}}, \end{aligned} \quad (6.12)$$

where $\gamma(t)$ is defined in (2.8). Therefore it suffices to bound $|\mathcal{T}|$.

If $i \geq 2$, then e is a copy of a secondary configuration (\mathcal{F}, R, D) where \mathcal{F} is a single hyperedge, $|R| = i$ and $|D| = r - i - 1$. Then since $\omega \notin \mathcal{B}_\ell$ we can use (B.2) to bound the number of such e and (B.3) to bound the number of choices for e^* to get

$$|\mathcal{T}| \leq \log^{2r^4(r-1)}(d) \cdot d^{\frac{1}{r-1}} \log^{-3K/5}(d) \cdot \log^{r^3(r-1)}(d) \leq d^{\frac{1}{r-1}} \log^{-K/2}(d), \quad (6.13)$$

for K chosen large with respect to r .

Now, suppose that $i = 1$ and let v be the unique element of S . The number of such pairs (e, e^*) with $|e \cap e^*| = 1$ is precisely $Y_v^{r-2,1}(\ell)$ which, since $\omega \notin \mathcal{B}_\ell$, is at most $(1 + \epsilon(t_\ell))y_{r-2,1}(t_\ell)d^{1/(r-1)}$. When $|e \cap e^*| \geq 2$, we have that $e \cup e^*$ is a copy of a secondary configuration with a single neutral vertex. So, as $\omega \notin \mathcal{B}_\ell$, by (B.2) the number of pairs (e, e^*) with $|e \cap e^*| \geq 2$ is at most $d^{\frac{1}{r-1}} \log^{-K/2}(d)$, as above. Therefore when $i = 1$, we have

$$|\mathcal{T}| \leq (1 + \epsilon(t_\ell))y_{r-2,1}(t_\ell)d^{1/(r-1)} + d^{\frac{1}{r-1}} \log^{-K/2}(d). \quad (6.14)$$

Putting together (6.12), (6.13) and (6.14) gives

$$\mathbb{E} \left[\mathbb{1}_{\mathcal{E}_\ell} (W_S^i(\ell + 1) - W_S^i(\ell)) \mid \mathcal{F}_\ell \right] \leq a_\ell, \quad (6.15)$$

which implies that the expectation of $A_S(\ell)$ given \mathcal{F}_ℓ is non-positive, as desired. This completes the proof of the claim. \square

Claim 6.16. $B_S(0), \dots, B_S(M)$ is η -bounded for

$$\eta := \log^{r^3(r-i-1)}(d).$$

Proof. First we bound the maximum value of $|A_S(\ell)|$. Again, we can assume that $\omega \notin \mathcal{B}_\ell$ as, otherwise, $|A_S(\ell)|$ is simply equal to zero. By definition of $A_S(\ell)$, the minimum possible value of $A_S(\ell)$ is

$$-a_\ell = -o(1).$$

Now we bound the maximum possible value of $W_S^i(\ell + 1) - W_S^i(\ell)$. The only way that this quantity can be positive is if some vertex, say x , becomes infected in the $(\ell + 1)$ th step. Given that x becomes infected, the maximum value that $W_S^i(\ell + 1) - W_S^i(\ell)$ can achieve is precisely $W_{S \cup \{x\}}^{i+1}(\ell)$. This is at most $\log^{r^3(r-i-1)}(d)$ by (B.3) since $\omega \notin \mathcal{B}_\ell$. So $|A_S(\ell)| \leq \eta$ for $0 \leq \ell \leq M$ and $B_S(0), \dots, B_S(M)$ is η -bounded, as required. \square

Claim 6.17.

$$\sum_{\ell=0}^{m-1} \text{Var} (A_S(\ell) \mid \mathcal{F}_\ell) = O \left(\log^{r^3(r-i-1)+(r-1)}(d) \right)$$

Proof. When $\omega \in \mathcal{B}_\ell$, we have that $\text{Var}(A_S(\ell) \mid \mathcal{F}_\ell) = 0$. So now consider when $\omega \notin \mathcal{B}_\ell$. Since for a constant c and any random variable X we have $\text{Var}(X - c) = \text{Var}(X) \leq \mathbb{E}(X^2)$, by definition of $A_S(\ell)$ we have

$$\text{Var}(A_S(\ell) \mid \mathcal{F}_\ell) \leq \mathbb{E} \left(\mathbb{1}_{\mathcal{E}_\ell} (W_S^i(\ell + 1) - W_S^i(\ell))^2 \mid \mathcal{F}_\ell \right).$$

Now,

$$\mathbb{E} \left(\mathbb{1}_{\mathcal{E}_\ell} (W_S^i(\ell + 1) - W_S^i(\ell))^2 \mid \mathcal{F}_\ell \right) = \sum_{k=1}^{\infty} k^2 \mathbb{P}(W_S^i(\ell + 1) - W_S^i(\ell) = k \mid \mathcal{F}_\ell). \quad (6.18)$$

By Claim 6.16, $W_S^i(\ell + 1) - W_S^i(\ell) \leq \eta$ for $\omega \notin \mathcal{B}_\ell$. So, the right hand side of (6.18) can be rewritten as

$$\sum_{k=1}^{\eta} k^2 \mathbb{P}(W_S^i(\ell + 1) - W_S^i(\ell) = k \mid \mathcal{F}_\ell).$$

The sum

$$\sum_{k=1}^{\eta} k \mathbb{P}(W_S^i(\ell + 1) - W_S^i(\ell) = k \mid \mathcal{F}_\ell)$$

is precisely the expected value of $\mathbb{1}_{\mathcal{E}_\ell} (W_S^i(\ell + 1) - W_S^i(\ell))$ given \mathcal{F}_ℓ . So by (6.15), definition of η , $\gamma(t)$ and $y_{r-2,1}(t)$, and the fact that $t = O(\log(d))$ we have,

$$\begin{aligned} & \sum_{k=1}^{\eta} k^2 \mathbb{P}(W_S^i(\ell + 1) - W_S^i(\ell) = k \mid \mathcal{F}_\ell) \leq \eta \cdot a_\ell \\ & = O \left(\frac{\log^{r^3(r-i-1)}(d) y_{r-2,1}(t_\ell)}{N \gamma(t_\ell)} \right) = O \left(\frac{\log^{r^3(r-i-1)}(d) (c + \alpha t_\ell)^{r-2}}{N} \right) \\ & = O \left(\frac{\log^{r^3(r-i-1)}(d) \log^{r-2}(d)}{N} \right). \end{aligned}$$

So by the above analysis,

$$\sum_{\ell=0}^{m-1} \text{Var}(A_S(\ell) \mid \mathcal{F}_\ell) = O \left(\frac{M \log^{r^3(r-i-1)+(r-2)}(d)}{N} \right) = O \left(\log^{r^3(r-i-1)+(r-1)}(d) \right),$$

since $M = O(N \log(d))$. This completes the proof of the claim. \square

Set $\nu := \log^{r^3(r-i-1)+r}(d)$ and $a := \frac{1}{2} \log^{r^3(r-i)}(d)$. Using Claims 6.11, 6.16 and 6.17, we can apply Theorem 3.10 to show that

$$\mathbb{P}(B_S(m) \geq a \text{ for some } 0 \leq m \leq M) \leq \exp \left(-\frac{a^2}{2(\nu + a\eta)} \right) \ll N^{-6\sqrt{\log(N)}},$$

as required for (6.10). This completes the proof of Proposition 6.8. \square

Hence Lemma 6.6 is proved.

5.6.2 Tracking the X Configurations

Now we determine the probability that (B.2) is violated at the first time any of (B.1), (B.2), (B.3) and (B.4) are violated. We will prove the following.

Lemma 6.19.

$$\mathbb{P}(\mathcal{X}) \leq N^{-4\sqrt{\log N}}.$$

For a secondary configuration $X = (\mathcal{F}, R, D)$ and $S \subseteq V(\mathcal{H})$ with $|S| = |R|$, define $\mathcal{X}(X, S)$ to be the set of all $\omega \in \mathcal{B}_M$ such that the bound

$$X_S(m) \leq \log^{2|D|r^4}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d) \quad (6.20)$$

is violated at time $m = \mathcal{J}(\omega)$, where $\mathcal{J}(\omega)$ is defined in (6.4). Observe that the bound (6.20) is precisely the bound (B.2) for our fixed choices of X and S . Therefore,

$$\mathcal{X} = \bigcup_{X, S} \mathcal{X}(X, S).$$

So the following proposition will imply Lemma 6.19, via an application of the union bound over all choices of X and S .

Proposition 6.21. *For all secondary configurations $X = (\mathcal{F}, R, D)$ and $S \subseteq V(\mathcal{H})$ such that $|S| = |R|$,*

$$\mathbb{P}(\mathcal{X}(X, S)) \leq N^{-6\sqrt{\log N}}.$$

Proof. Define \mathcal{E}_ℓ to be the event that $X_S(\ell + 1) \geq X_S(\ell)$. For $\ell \geq 0$, define

$$a_\ell := \frac{\log^{r^4(2|D|-1)}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}}}{N \log^{3K/5}(d)}$$

and

$$A_S(\ell) := \begin{cases} \mathbb{1}_{\mathcal{E}_\ell} (X_S(\ell + 1) - X_S(\ell)) - a_\ell, & \text{if } \omega \notin \mathcal{B}_\ell, \\ 0 & \text{otherwise} \end{cases}$$

and let

$$B_S(m) := \sum_{\ell=0}^{m-1} A_S(\ell).$$

If $\omega \notin \mathcal{B}_{m-1}$, then

$$X_S(m) \leq B_S(m) + X_S(0) + m \cdot a_\ell.$$

So if $X_S(0) \leq d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-4K/5}(d) = o\left(\log^{2r^4|D|}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d)\right)$, then

$$X_S(m) \leq B_S(m) + o\left(\log^{2r^4|D|}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d)\right) \quad (6.22)$$

by the fact that $m = O(N \log(d))$.

It follows that the event $\mathcal{X}(X, S)$ is contained within the event that either

$$X_S(0) > d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-4K/5}(d)$$

or

$$B_S(m) > \frac{1}{2} \log^{2|D|r^4}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d)$$

for some $0 \leq m \leq M$.

By Lemma 5.2,

$$\mathbb{P}\left(X_S(0) > d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-4K/5}(d)\right) \leq N^{-10\sqrt{\log N}},$$

so to prove Proposition 6.8 it suffices to show that $B_S(m)$ is unlikely to be large. We will show that

$$\mathbb{P}\left(B_S(m) > \frac{1}{2} \log^{2|D|r^4-3K/5}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \text{ for some } 0 \leq m \leq M\right) \leq N^{-6\sqrt{\log N}}. \quad (6.23)$$

We will apply Theorem 3.10. In order to apply Theorem 3.10 we must show that the sequence $B_S(0), \dots, B_S(M)$ is an η -bounded supermartingale and we also need to bound the sum $\sum_{\ell=0}^{m-1} \text{Var}(A_s(\ell) \mid \mathcal{F}_\ell)$.

Claim 6.24. $B_S(0), \dots, B_S(M)$ is a supermartingale.

Proof. This is equivalent to showing that, for $0 \leq \ell \leq M-1$, the expectation of $A_S(\ell)$ given \mathcal{F}_ℓ is non-positive. For $\omega \in \mathcal{B}_\ell$ we have $A_S(\ell) = 0$, and so it suffices to consider $\omega \notin \mathcal{B}_\ell$. For each $u \in D$, let X^u denote the configuration $(\mathcal{F}, R, D \setminus \{u\})$. By Remark 2.25, X^u is a secondary configuration. Every element of $X_S(\ell+1) \setminus X_S(\ell)$ comes from an element of $X_S^u(\ell)$, for some $u \in D$, and an open hyperedge e containing the image of u such that e is successfully sampled. As $\omega \notin \mathcal{B}_\ell$, using (6.3) and the fact that $\gamma(t_\ell)$ is bounded below by a function of r, c and α (see Remark 2.12) gives that the probability that any particular open hyperedge is successfully sampled is

$$\frac{q}{Q(\ell)} \leq \frac{\alpha}{\gamma(t_\ell)(1-4\epsilon(t_\ell))d^{1/(r-1)}N} = O\left(\frac{1}{d^{1/(r-1)}N}\right).$$

Also, since $\omega \notin \mathcal{B}_\ell$, we can apply (B.2) to $X_S^u(\ell)$ and (B.3) to the number of open hyperedges containing the image of u to get that the expected number of copies of X created in the $(\ell+1)$ th step is

$$O\left(\frac{1}{d^{1/(r-1)}N} \cdot \frac{\log^{2r^4(|D|-1)}(d) d^{\frac{|V(\mathcal{F})|-|R|-|D|+1}{r-1}}}{\log^{3K/5}(d)} \cdot \log^{r^3(r-1)}(d)\right)$$

$$= O\left(a_\ell \cdot \frac{\log^{r^3(r-1)}(d)}{\log^{r^4}(d)}\right) = o(a_\ell) \quad (6.25)$$

and therefore the expectation of $A_S(\ell)$ given \mathcal{F}_ℓ is negative. This proves that the sequence $B_S(0), \dots, B_S(M)$ is a supermartingale. \square

Claim 6.26. $B_S(0), \dots, B_S(M)$ is η -bounded for

$$\eta := \log^{r^4(2|D|-1)+1}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d).$$

Proof. We first bound the maximum value of $|A_S(\ell)|$. Again, assume $\omega \notin \mathcal{B}_\ell$. By (2.1) and the definition of $A_S(\ell)$, we have

$$A_S(\ell) \geq -a_\ell = o\left(d^{\frac{|V(\mathcal{F})|-|R|-|D|-1}{r-1}}\right).$$

For $u \in D$, define \tilde{X}^u to be the configuration $(\mathcal{F}, R \cup \{u\}, D \setminus \{u\})$. By Remark 2.25 this configuration is secondary. The value of $A_S(\ell)$ can only be positive if some vertex, say x , becomes infected in the $(\ell+1)$ th step. Given that x becomes infected, the number of copies of X rooted at S created is at most $\sum_{u \in D} \tilde{X}_{S \cup \{x\}}^u(\ell)$. Since $\omega \notin \mathcal{B}_\ell$, by (B.2) this is

$$O\left(\log^{2r^4(|D|-1)}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d)\right).$$

So we have $|A_S(\ell)| \leq \eta$ for $0 \leq \ell \leq M$, as required. \square

Claim 6.27.

$$\sum_{\ell=0}^{m-1} \text{Var}(A_S(\ell) \mid \mathcal{F}_\ell) \leq \frac{\log^{2r^4(2|D|-1)+1}(d) \cdot d^{\frac{2(|V(\mathcal{F})|-|R|-|D|)}{r-1}}}{\log^{6K/5}(d)}.$$

Proof. When $\omega \in \mathcal{B}_\ell$, we have that $\text{Var}(A_S(\ell) \mid \mathcal{F}_\ell) = 0$. So now consider when $\omega \notin \mathcal{B}_\ell$. We have

$$\text{Var}(A_S(\ell) \mid \mathcal{F}_\ell) \leq \mathbb{E}(A_S(\ell)^2 \mid \mathcal{F}_\ell) \leq \mathbb{E}\left(\mathbb{1}_{\mathcal{E}_\ell} (X_S(\ell+1) - X_S(\ell))^2 \mid \mathcal{F}_\ell\right).$$

As in the proof of Claim 6.17, we have

$$\mathbb{E}\left(\mathbb{1}_{\mathcal{E}_\ell} (X_S(\ell+1) - X_S(\ell))^2 \mid \mathcal{F}_\ell\right) = \sum_{k=1}^{\eta} k^2 \mathbb{P}(X_S(\ell+1) - X_S(\ell) = k \mid \mathcal{F}_\ell),$$

and

$$\mathbb{E}(\mathbb{1}_{\mathcal{E}_\ell} (X_S(\ell+1) - X_S(\ell)) \mid \mathcal{F}_\ell) = \sum_{k=1}^{\eta} k \mathbb{P}(X_S(\ell+1) - X_S(\ell) = k \mid \mathcal{F}_\ell).$$

So, by (6.25) and definition of η , we get that $\text{Var}(A_S(\ell) \mid \mathcal{F}_\ell)$ is at most

$$o(\eta \cdot a_\ell) = o\left(\frac{\log^{2r^4(2|D|-1)}(d) \cdot d^{\frac{2(|V(\mathcal{F})|-|R|-|D|)}{r-1}}}{N \log^{6K/5}(d)}\right).$$

By the above analysis, $\sum_{\ell=0}^{m-1} \text{Var}(A_S(\ell) \mid \mathcal{F}_\ell)$ is therefore at most

$$\nu := \frac{\log^{2r^4(2|D|-1)+1}(d) \cdot d^{\frac{2(|V(\mathcal{F})|-|R|-|D|)}{r-1}}}{\log^{6K/5}(d)}, \quad (6.28)$$

as $M = O(N \log d)$. \square

Set ν as in (6.28) and $a := \frac{1}{2} \log^{2|D|r^4}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d)$. Applying Theorem 3.10 shows that

$$\mathbb{P}(B_S(m) \geq a \text{ for some } 0 \leq m \leq M) \leq \exp\left(-\frac{a^2}{2(\nu + a\eta)}\right) \ll N^{-6\sqrt{\log N}},$$

as required for (6.23). This completes the proof of Proposition 6.21. \square

Therefore, the proof of Lemma 6.19 is concluded.

5.6.3 Tracking the \mathcal{Y} Configurations

Much of the analysis in this subsection is similar to the previous two; however, we must be more careful since we are aiming at tight concentration bounds (not just crude upper bounds). As we require both upper and lower bounds, we will be dealing with both a supermartingale and a submartingale.

We wish to bound the probability that (B.1) is violated at the first time any of (B.1), (B.2), (B.3) and (B.4) are violated. We will prove the following.

Lemma 6.29.

$$\mathbb{P}(\mathcal{Y}) \leq N^{-2\sqrt{\log N}}.$$

For $0 \leq i \leq r-2$, $0 \leq j \leq r-1-i$ and $v \in V(\mathcal{H})$, define $\mathcal{Y}(i, j, v)$ to be the set of all $\omega \in \mathcal{B}_M$ such that the bound

$$Y_v^{i,j}(m) \in (1 \pm \epsilon(t)) d^{1-\frac{i}{r-1}} y_{i,j}(t_m) \quad (6.30)$$

is violated at time $m = \mathcal{J}(\omega)$, where $\mathcal{J}(\omega)$ is defined in (6.4) and $y_{i,j}(t)$ is defined in (2.22). Observe that the bound (6.30) is precisely the bound (B.1) for our fixed choices of i, j and v . Note that

$$\mathcal{Y} = \bigcup_{i,j,v} \mathcal{Y}(i, j, v).$$

Thus the following proposition will imply the lemma, via an application of the union bound over all choices of i, j and v .

Proposition 6.31. *For $0 \leq i \leq r-2$, $0 \leq j \leq r-1-i$ and $v \in V(\mathcal{H})$,*

$$\mathbb{P}(\mathcal{Y}(i, j, v)) \leq N^{-3\sqrt{\log N}}.$$

The proof of the proposition relies on two claims (Claim 6.33 and Claim 6.34), which will be stated where they are needed once the relevant variables have been defined. They will be proved later after completing the proof of the proposition assuming the claims.

Proof of Proposition 6.31. For $\ell \geq 0$, define

$$a_\ell^\pm := d^{1-\frac{i}{r-1}} \left(\frac{y'_{i,j}(t_\ell)}{N} \mp \frac{20 \cdot |y'_{i,j}(t_\ell)| \cdot \epsilon(t_\ell)}{N} \right)$$

and

$$A_v^\pm(\ell) := \begin{cases} \left(Y_v^{i,j}(\ell+1) - Y_v^{i,j}(\ell) \right) - a_\ell^\pm & \text{if } \omega \notin \mathcal{B}_\ell, \\ 0 & \text{otherwise,} \end{cases}$$

where $y_{i,j}(t)$ is defined in (2.22). We clarify that we are defining two different random variables $A_v^+(\ell)$ and $A_v^-(\ell)$. The superscript denotes whether we expect the variable to be ‘typically positive’ or ‘typically negative’. Given this definition, we define the following pair of random variables:

$$B_v^\pm(m) := \sum_{\ell=0}^{m-1} A_v^\pm(\ell).$$

Thus if $\omega \notin \mathcal{B}_{m-1}$, by definition,

$$Y_v^{i,j}(m) = B_v^\pm(m) + Y_v^{i,j}(0) + d^{1-\frac{i}{r-1}} \left(\sum_{\ell=0}^{m-1} \frac{y'_{i,j}(t_\ell)}{N} \mp 20 \sum_{\ell=0}^{m-1} \frac{|y'_{i,j}(t_\ell)| \cdot \epsilon(t_\ell)}{N} \right). \quad (6.32)$$

Our strategy is to obtain a concentration result for $Y_v^{i,j}(m)$ by analysing each of the terms on the right side of this expression.

The proposition will follow from the next two claims.

Claim 6.33. *If $\omega \notin \mathcal{B}_{m-1}$ and*

$$Y_v^{i,j}(0) \in \left(1 \pm \log^{-3K/10}(d) \right) y_{i,j}(0) d^{1-\frac{i}{r-1}},$$

then the following bounds hold:

$$\begin{aligned} Y_v^{i,j}(m) &\leq B_v^-(m) + d^{1-\frac{i}{r-1}} \left(1 + \frac{\epsilon(t_m)}{2} \right) y_{i,j}(t_m), \\ Y_v^{i,j}(m) &\geq B_v^+(m) + d^{1-\frac{i}{r-1}} \left(1 - \frac{\epsilon(t_m)}{2} \right) y_{i,j}(t_m). \end{aligned}$$

Claim 6.34. *With probability at least $1 - N^{-4\sqrt{\log(N)}}$, both the following bounds hold for all $0 \leq m \leq M$:*

$$\begin{aligned} B_v^-(m) &\leq \frac{1}{2} \epsilon(t_m) d^{1-\frac{i}{r-1}} y_{i,j}(t_m), \\ B_v^+(m) &\geq -\frac{1}{2} \epsilon(t_m) d^{1-\frac{i}{r-1}} y_{i,j}(t_m). \end{aligned}$$

Indeed, by Claim 6.33, the event $\mathcal{Y}(i, j, v)$ is contained within the event that either

$$Y_v^{i,j}(0) \notin \left(1 \pm \log^{-3K/10}(d)\right) y_{i,j}(0) d^{1-\frac{i}{r-1}},$$

or one of the following bounds hold:

$$\begin{aligned} B_v^-(m) &> \frac{1}{2} \epsilon(t_m) d^{1-\frac{i}{r-1}} y_{i,j}(t_m) \\ B_v^+(m) &< -\frac{1}{2} \epsilon(t_m) d^{1-\frac{i}{r-1}} y_{i,j}(t_m) \end{aligned}$$

By Lemma 5.3,

$$\mathbb{P}\left(Y_v^{i,j}(0) \notin \left(1 \pm \log^{-3K/10}(d)\right) y_{i,j}(0) d^{1-\frac{i}{r-1}}\right) \leq N^{-10\sqrt{\log N}}.$$

Combining this with Claim 6.34 completes the proof of the proposition. \square

We now prove Claims 6.33 and 6.34.

Proof of Claim 6.33. We will analyse each term in the right hand side of (6.32). By Lemma 4.6 with $s(t) = y'_{i,j}(t)$, we have

$$\begin{aligned} \sum_{\ell=0}^{m-1} \frac{y'_{i,j}(t_\ell)}{N} &\in \int_0^{t_m} y'_{i,j}(t) dt \pm \frac{m \cdot \sup_{t \in [0, T]} |y''_{i,j}(t)|}{2N^2} \\ &\in y_{i,j}(t_m) - y_{i,j}(0) \pm \frac{\log^{r^2}(d)}{N}, \end{aligned} \quad (6.35)$$

as $m = O(N \log d)$.

Next, we analyse the second summation in the parentheses of (6.32). First, by definition of $\epsilon(t)$ (given in (2.9)),

$$20 \sum_{\ell=0}^{m-1} \frac{|y'_{i,j}(t_\ell)| \cdot \epsilon(t_\ell)}{N} = \left(\frac{20}{\log^{K/5}(d)}\right) \sum_{\ell=0}^{m-1} \frac{|y'_{i,j}(t_\ell)| \cdot (t_\ell + 1)^{K/10}}{N}. \quad (6.36)$$

The function $y_{i,j}(t)$ is a polynomial with positive leading coefficient and, by Remark 2.12, $y_{i,j}(t)$ is bounded away from zero by a function of r, c and α for all $t \in [0, T]$. Therefore, there exists a positive constant $C = C(r, c, \alpha)$ such that

$$|y'_{i,j}(t_\ell)| \cdot (t_\ell + 1)^{K/10} \leq C \cdot y_{i,j}(t_\ell) \cdot (t_\ell + 1)^{(K/10)-1}.$$

Combining this with (6.36) gives

$$20 \sum_{\ell=0}^{m-1} \frac{|y'_{i,j}(t_\ell)| \cdot \epsilon(t_\ell)}{N} \leq \left(\frac{20 \cdot C}{\log^{K/5}(d)}\right) \sum_{\ell=0}^{m-1} \frac{y_{i,j}(t_\ell) \cdot (t_\ell + 1)^{(K/10)-1}}{N}.$$

Now, applying Lemma 4.6 and choosing K sufficiently large with respect to r ,

$$\sum_{\ell=0}^{m-1} \frac{y_{i,j}(t_\ell) \cdot (t_\ell + 1)^{(K/10)-1}}{N} \leq \int_0^{t_m} y_{i,j}(t) \cdot (t + 1)^{(K/10)-1} dt + \frac{\log^{K/9}(d)}{N}.$$

Consider the integral on the right side of the above inequality. By definition of $y_{i,j}(t)$ (given in (2.21) and (2.22)) and definition of $\gamma(t)$ (2.8), we have

$$\int_0^{t_m} y_{i,j}(t) \cdot (t + 1)^{(K/10)-1} dt \leq \int_0^{t_m} \binom{r-1}{i} \binom{r-1-i}{j} (t + 1)^{K/10-1} (c + \alpha t)^{i+(r-1)j} dt.$$

Letting $C' := \binom{r-1}{i} \binom{r-1-i}{j} \cdot (\max\{1, c, \alpha\})^{r^3}$, gives that this is at most

$$\begin{aligned} C' \int_0^{t_m} (t + 1)^{i+(r-1)j+K/10-1} dt &\leq \frac{10C'}{K} (t_m + 1)^{i+(r-1)j+K/10} \\ &\leq \tilde{C} \cdot y_{i,j}(t_m) \cdot (t_m + 1)^{K/10}, \end{aligned}$$

for some positive constant \tilde{C} depending on r, c, α and K . Moreover, by choosing K large with respect to r, c and α , we may take \tilde{C} arbitrarily close to zero. So, provided that K is large enough, we have

$$\begin{aligned} 20 \sum_{\ell=0}^{m-1} \frac{|y'_{i,j}(t_\ell)| \cdot \epsilon(t_\ell)}{N} &\leq \left(\frac{20 \cdot C}{\log^{K/5}(d)} \right) \left(\tilde{C} \cdot y_{i,j}(t_m) \cdot (t_m + 1)^{K/10} + \frac{\log^{K/9}(d)}{N} \right) \\ &\leq \frac{y_{i,j}(t_m) \epsilon(t_m)}{16}, \end{aligned} \quad (6.37)$$

as we may choose \tilde{C} arbitrarily close to zero. Now let us combine (6.32), (6.35) and (6.37) with our hypothesis to give an upper bound for $Y_v^{i,j}(m)$:

$$\begin{aligned} Y_v^{i,j}(m) &\leq B_v^-(m) + \left(1 + \log^{-3K/10}(d) \right) y_{i,j}(0) d^{1-\frac{i}{r-1}} \\ &\quad + d^{1-\frac{i}{r-1}} \left(y_{i,j}(t_m) - y_{i,j}(0) + \frac{\log^{r^2}(d)}{N} + \frac{y_{i,j}(t_m) \epsilon(t_m)}{4} \right). \end{aligned}$$

By the definition of $\epsilon(t)$ in (2.9) we can absorb the error term $\log^{-3K/10}(d) y_{i,j}(0) d^{1-\frac{i}{r-1}}$ into the main error term. This gives

$$Y_v^{i,j}(m) \leq B_v^-(m) + \left(1 + \frac{\epsilon(t_m)}{2} \right) d^{1-\frac{i}{r-1}} y_{i,j}(t_m).$$

Similarly, we get

$$Y_v^{i,j}(m) \geq B_v^+(m) + \left(1 - \frac{\epsilon(t_m)}{2} \right) d^{1-\frac{i}{r-1}} y_{i,j}(t_m),$$

as required. \square

It remains to prove Claim 6.34.

Proof of Claim 6.34. To prove these bounds, we show that the sequence $B_v^-(0), \dots, B_v^-(m)$ is a supermartingale and that $B_v^+(0), \dots, B_v^+(m)$ is a submartingale which satisfy certain ‘maximum change’ and ‘variance increment’ bounds and apply Theorem 3.10. This amounts to bounding the expectation, maximum/minimum possible values and variance of $A_v^\pm(\ell)$. As in the previous two subsections, we will always assume $\omega \notin \mathcal{B}(\ell)$; otherwise, we have $A_v^\pm(\ell) = 0$ and so all of the required bounds hold trivially. Recall the definition of $y_{i,j}(t)$ from (2.21) and (2.22) and the definition of $\gamma(t)$ from (2.8). The following expression, obtained by differentiating $y_{i,j}(t)$, will be useful in what follows:

$$\begin{aligned} y'_{i,j}(t) &= \binom{r-1}{i} \binom{r-1-i}{j} (\alpha \cdot i(c + \alpha t)^{i-1} \gamma(t)^j + (c + \alpha t)^i j \gamma'(t) \gamma(t)^{j-1}) \\ &= \frac{(j+1)\alpha y_{i-1,j+1}(t)}{\gamma(t)} + \frac{(r-i-j)\alpha y_{i,j-1}(t) y_{r-2,1}(t)}{\gamma(t)} - \frac{j y_{i,j}(t)}{\gamma(t)}. \end{aligned} \quad (6.38)$$

Subclaim 6.39. $B_v^+(0), \dots, B_v^+(M)$ is a submartingale and $B_v^-(0), \dots, B_v^-(M)$ is a supermartingale.

Proof of Subclaim 6.39. First, we bound the expectation of $Y_v^{i,j}(\ell+1) - Y_v^{i,j}(\ell)$ given \mathcal{F}_ℓ . To do this we must consider copies of $Y^{i,j}$ rooted at v that are created by successfully sampling an open hyperedge at the $(\ell+1)$ th step, and also copies that are destroyed. Let $C_{\ell+1}$ be the number of copies of $Y^{i,j}$ rooted at v which are present in the $(\ell+1)$ th step but not in the ℓ th step. Any such copy must come from either:

- (1) an element \mathcal{F}' of $Y_v^{i-1,j}(\ell) \setminus Y_v^{i,j}(\ell)$ such that a healthy vertex of \mathcal{F}' , say x , contained only in the central hyperedge becomes infected in the $(\ell+1)$ th step, or
- (2) an element \mathcal{F}' of $Y_v^{i,j-1}(\ell)$ and a copy \mathcal{G}' of $Y^{r-2,0}$ in $\mathcal{H}(\ell)$ where:
 - \mathcal{G}' contains precisely $r-2$ infected vertices,
 - \mathcal{G}' is rooted at some non-root $u \in \mathcal{F}'$, where u is contained only in the central hyperedge of \mathcal{F}' ,
 - $\mathcal{G}' \cap \mathcal{F}' = \{u\}$,
 - the unique healthy non-root x of \mathcal{G}' becomes infected at the $(\ell+1)$ th step.

See Figure 5.9 for examples of (1) and Figure 5.10 for examples of (2) (both in the case $r=6$). See also Figure 5.4 for other examples of these (u there plays the role of x here). Both (1) and (2) will contribute to the main term in the expectation of $Y_v^{i,j}(\ell+1) - Y_v^{i,j}(\ell)$ given \mathcal{F}_ℓ .



Figure 5.9: Here $r = 6$. Some examples of ways that a copy of $Y^{3,1}$ rooted at v can be created via (1) when the hyperedge shaded light grey becomes successfully sampled and x becomes infected. The central hyperedge of the copy of $Y^{3,1}$ created is drawn with a thick outline. Infected vertices are shaded dark grey and healthy vertices are unshaded.

The only way that a vertex can become infected is if one of the open hyperedges containing that vertex is successfully sampled. The probability that any given open hyperedge is successfully sampled is

$$\frac{q}{Q(\ell)} \in \frac{\alpha}{(1 \mp 4\epsilon(t_\ell))\gamma(t_\ell)Nd^{1/(r-1)}}. \quad (6.40)$$

Let us count the number of ways that we can create a copy of $Y^{i,j}$ rooted at v that is not present in the ℓ th step using (1). We will break this into two cases. If e^* is the hyperedge that is successfully sampled at the $(\ell + 1)$ th step, then either $|e^* \cap \mathcal{F}'| = 1$, or $|e^* \cap \mathcal{F}'| > 1$. Consider the first case.

We wish to count the number of ways to choose:

- an element \mathcal{F}' of $Y_v^{i-1,j}(\ell) \setminus Y_v^{i,j}(\ell)$;
- a healthy vertex x contained only in the central hyperedge of \mathcal{F}' ; and
- an open hyperedge e^* containing x such that e^* intersects the copy of $Y^{i-1,j}$ only on x .

Observe that, in this case, $e^* \cup \mathcal{F}'$ is a member of $Y_v^{i-1,j+1}(\ell)$ that contains exactly $i - 1$ infected vertices in its central hyperedge. For each such member of $Y_v^{i-1,j+1}(\ell)$, there are $j + 1$ possible choices of the hyperedge e^* .

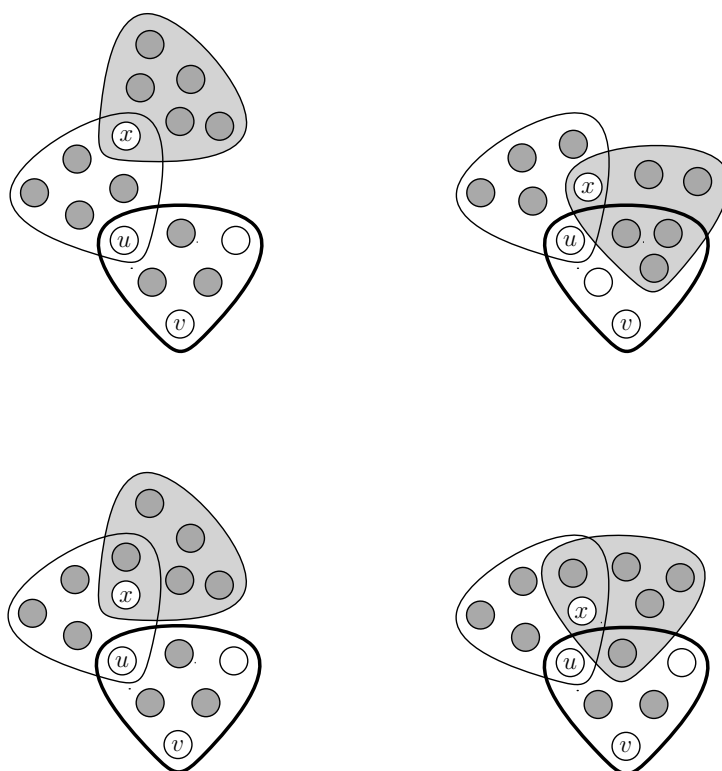


Figure 5.10: Here $r = 6$. Some examples of ways that a copy of $Y^{3,1}$ rooted at v can be created via (2) when the hyperedge shaded light grey becomes successfully sampled and x becomes infected. The central hyperedge of the copy of $Y^{3,1}$ created is drawn with a thick outline. Infected vertices are shaded dark grey and healthy vertices are unshaded.

Therefore the number of ways to create a member of $Y_v^{i,j}(\ell + 1)$ via (1) when $|e^* \cap \mathcal{F}'| = 1$ is at most $(j + 1)Y_v^{i-1,j+1}(\ell)$. However, this also unnecessarily counts $(j + 1)$ times the number of copies of $Y^{i-1,j+1}$ rooted at v that contain at least i infected vertices in their central hyperedge. Therefore we need to subtract off $(j + 1)$ times the number of such copies. Let us now crudely bound these. Such a copy consists of a member \mathcal{F}'' of $Y_v^{i,0}(\ell)$ and $j + 1$ copies of W^1 rooted at vertices of \mathcal{F}'' . Since $\omega \notin \mathcal{B}_\ell$, by (B.3) we have that $W_u^1(\ell) = \log^{O(1)}(d)$ for all $u \in V(\mathcal{H})$.

So the number of ways to create a member of $Y_v^{i,j}(\ell + 1)$ via (1) when $|e^* \cap \mathcal{F}'| = 1$, is within

$$(j + 1)Y_v^{i-1,j+1}(\ell) \pm Y_v^{i,0}(\ell) \log^{O(1)}(d).$$

Since $\omega \notin \mathcal{B}_\ell$, we have

$$(j + 1)Y_v^{i-1,j+1}(\ell) \pm Y_v^{i,0}(\ell) \log^{O(1)}(d) \in (j + 1)(1 \pm 2\epsilon(t_\ell))y_{i-1,j+1}(t_\ell)d^{1-\frac{i-1}{r-1}}. \quad (6.41)$$

Now consider the case when $|e^* \cap \mathcal{F}'| > 1$. As we will see, this will give a lower order term. In this case, $e^* \cup \mathcal{F}'$ is a member of $Z_v^{i-1,j+1}(m) \setminus Y_v^{i-1,j+1}(m)$. So by Observation 2.34, $e^* \cup \mathcal{F}'$ consists of:

- (a) a copy of a secondary configuration $X = (\mathcal{F}^*, R^*, D^*)$ in $\mathcal{H}(\ell)$ with a unique root and precisely $i - 1$ marked vertices in the central hyperedge such that every non-central hyperedge has a unique neutral vertex (also contained in the central hyperedge), and
- (b) at most j copies of W^1 rooted at vertices of the central hyperedge.

So we can bound the number of ways of creating a member of $Y_v^{i,j}(\ell + 1)$ via (1) in the second case by the sum of $X_v(\ell)$ multiplied by $\log^{O(1)}(d)$ (for the copies of W^1), over all secondary configurations X satisfying (a).

Since $\omega \notin \mathcal{B}_\ell$ and $|V(\mathcal{F}^*)| - |R^*| - |D^*| = r - i$ for each secondary configuration X satisfying (a), by (B.2) this is at most

$$\log^{O(1)}(d) \cdot d^{1-\frac{i-1}{r-1}} \log^{-3K/5}(d) = o\left(d^{1-\frac{i-1}{r-1}}\epsilon(t_\ell)\right). \quad (6.42)$$

Putting (6.41) and (6.42) together with (6.40) shows that the expected number of copies of $Y^{i,j}$ rooted at v created via (1) is

$$\frac{\alpha(j + 1)(1 \pm 5\epsilon(t_\ell))y_{i-1,j+1}(t_\ell)d^{1-\frac{i}{r-1}}}{(1 \mp 4\epsilon(t_\ell))\gamma(t_\ell)N} \subseteq \frac{\alpha(j + 1)(1 \pm 20\epsilon(t_\ell))y_{i-1,j+1}(t_\ell)d^{1-\frac{i}{r-1}}}{\gamma(t_\ell)N}, \quad (6.43)$$

since $\left|\frac{1+5x}{1-4x}\right| \leq 1 \pm 20x$ for x sufficiently small.

Counting the number of ways to create a copy of $Y^{i,j}$ rooted at v via (2) (recall that Figure 5.10 provides examples of this) is equivalent to counting

- the number of ways to choose an element \mathcal{F}' of $Y_v^{i,j-1}(\ell)$;
- a healthy vertex $u \in V(\mathcal{F}') \setminus \{v\}$ contained only in the central hyperedge;
- an element \mathcal{G}' of $Y_u^{r-2,0}(\ell)$ (a single edge) that contains exactly $r - 2$ infected vertices and intersects \mathcal{F}' precisely on u ,
- an element e^* of $Q_x(\ell)$, where x is the unique healthy vertex of $\mathcal{G}' \setminus \{u\}$. (e^* is the hyperedge shaded light grey in the examples in Figure 5.10.)

As $\omega \notin \mathcal{B}_\ell$, there are at most $O\left(d^{1-\frac{i+1}{r-1}} \log^{O(1)}(d)\right)$ members of $Y_v^{i,j-1}(\ell)$ that do not contain exactly i infected vertices in their central hyperedge (as we just did in (1), we can bound the number of possible central hyperedges by $Y_v^{i+1,0}(\ell)$ and other hyperedges (which are copies of W^1) using (B.3)). For each \mathcal{F}' , the number of choices of u is precisely the number of healthy vertices in the central hyperedge of \mathcal{F}' . So, the number of ways to choose such an \mathcal{F}' and u is contained in

$$(r - i - j)Y_v^{i,j-1}(\ell) \pm O\left(d^{1-\frac{i+1}{r-1}} \log^{O(1)}(d)\right).$$

As $\omega \notin \mathcal{B}_\ell$, by (B.1) this is contained in

$$(r - i - j)(1 \pm 2\epsilon(t))y_{i,j-1}(t)d^{1-\frac{i}{r-1}}. \tag{6.44}$$

Let $C(\mathcal{F}', u, \mathcal{G}'')$ be the number of ways to pick such an \mathcal{F}' and u multiplied by the number of ways to pick some $\mathcal{G}'' = \mathcal{G}' \cup \{e^*\} \in Y_v^{r-2,1}(\ell)$. So as $\omega \notin \mathcal{B}_\ell$, by (B.1) and (6.44) we have

$$C(\mathcal{F}', u, \mathcal{G}'') = (r - i - j)(1 \pm 4\epsilon(t))y_{i,j-1}(t)y_{r-2,1}(t)d^{1-\frac{i-1}{r-1}}. \tag{6.45}$$

The quantity $C(\mathcal{F}', u, \mathcal{G}'')$ both fails to count some triples we wish to count and also counts some triples we do not want to count. However, we will show that both these terms are of a lower order. So up to an error term, the number of ways to create a copy of $Y^{i,j}$ rooted at v via (2) is $C(\mathcal{F}', u, \mathcal{G}'')$. See Figure 5.11 for examples of the triples we fail to count and Figure 5.12 for examples of the triples we count unwantedly.

The triples which $C(\mathcal{F}', u, \mathcal{G}'')$ fails to count are those where e^* intersects \mathcal{G}' in more than one vertex. However, for such a pair \mathcal{G}' and e^* , we have that $\mathcal{G}' \cup e^*$ is a copy in $\mathcal{H}(\ell)$ of a secondary configuration X rooted at u with precisely one neutral vertex. As $\omega \notin \mathcal{B}_\ell$, by (B.2) we have $X_u(\ell) \leq d^{\frac{1}{r-1}} \log^{-2K/5}$. Summing over all such configurations X , and using (6.44) gives that there are

$$O\left(d^{1-\frac{i-1}{r-1}} \log^{-2K/5}\right) \tag{6.46}$$

triples that are not counted in $C(\mathcal{F}', u, \mathcal{G}'')$.

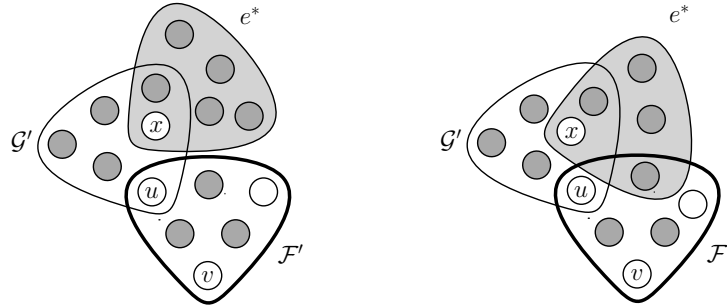


Figure 5.11: Here $r = 6$, $i = 3$ and $j = 1$. Two examples of hypergraphs that are not counted in $C(\mathcal{F}', u, \mathcal{G}'')$. If the shaded hyperedge were successfully sampled, this would create a copy of $Y^{3,1}$ whose central hyperedge is drawn with a thick outline. Infected vertices are shaded dark grey and healthy vertices are unshaded.

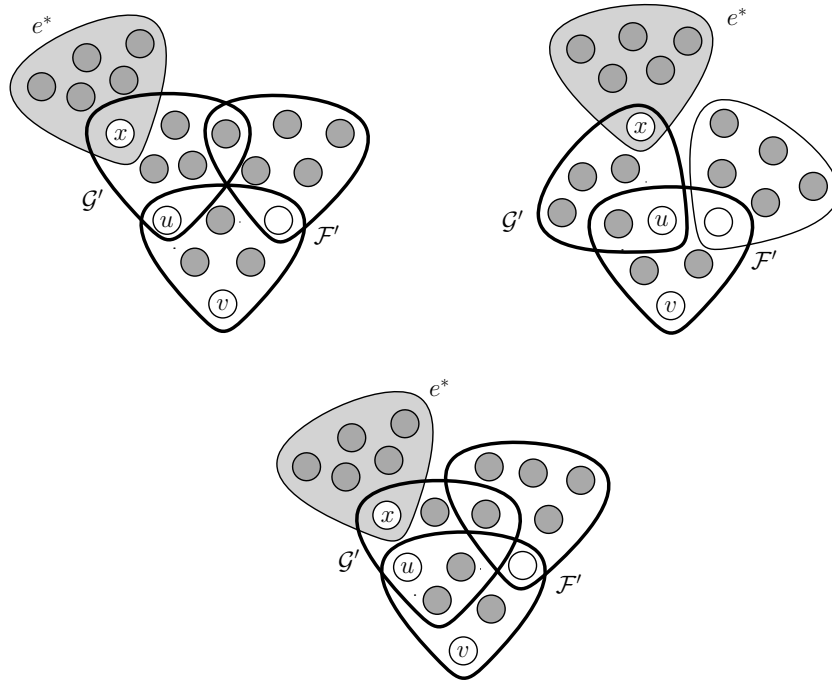


Figure 5.12: Here $r = 6$, $i = 3$ and $j = 2$. Three examples of hypergraphs that are unwantedly counted in $C(\mathcal{F}', u, \mathcal{G}'')$. Observe that $|\mathcal{F}' \cap \mathcal{G}'| \geq 2$ and $\mathcal{F}' \cup \mathcal{G}'$ consists of a copy X of a secondary configuration (whose edges are drawn with a thick outline) and some additional copies of W^1 rooted at healthy vertices of X . The central hyperedge of X is the hyperedge containing v . In any of these cases, if the shaded hyperedge were successfully sampled, it would not turn $\mathcal{F}' \cup \mathcal{G}'$ into a copy of $Y^{3,2}$.

Now consider the triples which $C(\mathcal{F}', u, \mathcal{G}'')$ counts that we do not want (Figure 5.12 provides examples of these). The triples we wish to exclude are precisely those where $|\mathcal{F}' \cap \mathcal{G}'| \geq 2$. \mathcal{G}' could intersect either the central hyperedge or a non-central hyperedge of \mathcal{F}' (or both). However for any such pair, $\mathcal{F}' \cup \mathcal{G}' \cup e^*$ consists of:

- (a) a copy in $\mathcal{H}(\ell)$ of a secondary configuration X with a unique root (v) and exactly i marked (so at least i infected) vertices in the central hyperedge, and
- (b) at most j copies of W^1 (one is e^* and the others are the non-central hyperedges of \mathcal{F}'). (The reason we have *at most* j is because the secondary configuration could contain 2 or 3 hyperedges.)

So similarly to the second case of (1) above, since $\omega \notin \mathcal{B}_\ell$ and $|V(\mathcal{F}^*)| - |R^*| - |D^*| = r - (i + 1)$ for each secondary configuration $X = (\mathcal{F}^*, R^*, D^*)$ satisfying (a), by (B.2) and (B.3) the number of triples we wish to exclude is

$$\log^{O(1)}(d)d^{1-\frac{i}{r-1}} \log^{-3K/5}(d) = O\left(d^{1-\frac{i}{r-1}} \log^{-2K/5}\right), \quad (6.47)$$

for K chosen sufficiently large with respect to r (once \mathcal{F}' , u and \mathcal{G}'' are chosen, by (B.3) there are $\log^{O(1)}(d)$ choices for e^*).

So putting (6.45), (6.46) and (6.47) together, the number of ways to choose \mathcal{F}' , u , \mathcal{G}' and e^* is contained in

$$(r - i - j) (1 \pm 5\epsilon(t)) y_{i,j-1}(t) y_{r-2,1}(t) d^{1-\frac{i-1}{r-1}}. \quad (6.48)$$

Combining (6.48) and (6.40) gives that the expected number of copies of $Y^{i,j}$ rooted at v created in the second way is contained in

$$\frac{\alpha(r - i - j)(1 \pm 20\epsilon(t_\ell)) y_{i,j-1}(t_\ell) y_{r-2,1}(t_\ell) d^{1-\frac{i}{r-1}}}{\gamma(t_\ell) N}. \quad (6.49)$$

Therefore, by (6.43) and (6.49) we have:

$$\mathbb{E}(C_{\ell+1} | \mathcal{F}_\ell) \in \frac{\alpha(1 \pm 20\epsilon(t_\ell)) d^{1-\frac{i}{r-1}} ((j+1) y_{i-1,j+1}(t_\ell) + (r-i-j) y_{i,j-1}(t_\ell) y_{r-2,1}(t_\ell))}{\gamma(t_\ell) N}. \quad (6.50)$$

By Definition 2.17, we see that the only way in which a copy of $Y^{i,j}$ rooted at v can be destroyed in the $(\ell+1)$ th step is if one of its hyperedges is sampled. Let $D_{\ell+1}$ be the number of such destroyed copies. As discussed above, since $\omega \notin \mathcal{B}_\ell$, all but at most $\log^{O(1)}(d) \cdot d^{1-\frac{i+1}{r-1}}$ copies of $Y^{i,j}$ in $\mathcal{H}(\ell)$ rooted at v have exactly i infected vertices in the central hyperedge. A copy whose central hyperedge is open is destroyed with probability $(j+1)q/Q(\ell)$. However,

for the vast majority of copies of $Y^{i,j}$ rooted at v , the central hyperedge is not open. Using (6.40), the probability that any such copy is destroyed is precisely

$$\frac{jq}{Q(\ell)} \in \frac{\alpha j}{(1 \mp 4\epsilon(t_\ell))\gamma(t_\ell)N}.$$

Therefore we have

$$\mathbb{E}(D_{\ell+1} | \mathcal{F}_\ell) \in \frac{\alpha j (Y_v^{i,j}(\ell) - \log^{O(1)}(d) \cdot d^{1-\frac{i+1}{r-1}}) \pm \alpha(j+1) \log^{O(1)}(d) \cdot d^{1-\frac{i+1}{r-1}}}{(1 \mp 4\epsilon(t_\ell))\gamma(t_\ell)N}.$$

As $\omega \notin \mathcal{B}_\ell$, by (B.1) we have

$$\mathbb{E}(D_{\ell+1} | \mathcal{F}_\ell) \in \frac{\alpha j (1 \pm 5\epsilon(t_\ell)) y_{i,j}(t_\ell) d^{1-\frac{i}{r-1}}}{(1 \mp 4\epsilon(t_\ell))\gamma(t_\ell)N} \subseteq \frac{\alpha j (1 \pm 20\epsilon(t_\ell)) y_{i,j}(t_\ell) d^{1-\frac{i}{r-1}}}{\gamma(t_\ell)N}. \quad (6.51)$$

As $Y_v^{i,j}(\ell+1) - Y_v^{i,j}(\ell) = C_{\ell+1} - D_{\ell+1}$, by linearity of expectation, (6.50) and (6.51) we get that

$$\mathbb{E}(Y_v^{i,j}(\ell+1) - Y_v^{i,j}(\ell) | \mathcal{F}_\ell) \in \frac{(1 \pm 20\epsilon(t_\ell)) y'_{i,j}(t_\ell)}{N} d^{1-\frac{i}{r-1}}, \quad (6.52)$$

by (6.38). Therefore the expectation of $A_v^+(\ell)$ given \mathcal{F}_ℓ is positive and the expectation of $A_v^-(\ell)$ given \mathcal{F}_ℓ is negative. So we have that $B_v^+(0), \dots, B_v^+(M)$ is a submartingale and $B_v^-(0), \dots, B_v^-(M)$ is a supermartingale, as required for Subclaim 6.39. \square

Now, let us bound the maximum and minimum possible values of $A_v^\pm(\ell)$.

Subclaim 6.53. *Both $B_v^+(0), \dots, B_v^+(M)$ and $B_v^-(0), \dots, B_v^-(M)$ are η -bounded, for*

$$\eta := d^{1-\frac{i}{r-1}} \log^{-K/2}(d).$$

Proof. The maximum value of $Y_v^{i,j}(\ell+1) - Y_v^{i,j}(\ell)$ occurs when the $(\ell+1)$ th sampling is successful and the newly infected vertex creates many new copies of $Y^{i,j}$ rooted at v in either the first or second way (see above). For each $u \in D$ let \tilde{Y}^u denote the configuration $(\mathcal{F}, R \cup \{u\}, D \setminus \{u\})$. If some vertex, say x , becomes infected in the $(\ell+1)$ th step, the number of copies of $Y^{i,j}$ created is $\sum_{u \in D} \tilde{Y}_{S \cup x}^u$. Each configuration \tilde{Y}^u consists of a copy \mathcal{F}' of firstly a secondary configuration with $(r-1-i)$ neutral vertices, and secondly some additional copies of W^1 rooted at vertices of \mathcal{F}' . So as $\omega \notin \mathcal{B}_\ell$ the maximum number of copies of $Y^{i,j}$ rooted at v that can be created in a time step is

$$\log^{O(1)}(d) \cdot d^{1-\frac{i}{r-1}} \log^{-3K/5}(d) = o(\eta). \quad (6.54)$$

As mentioned above, a copy of $Y^{i,j}$ rooted at v can only be destroyed in the $(\ell+1)$ th step if it contains an open hyperedge which is sampled. Given that some vertex x becomes infected, the maximum number of copies that can be destroyed is at most the number of

hyperedges containing $\{v, x\}$ and i (other) infected vertices, multiplied by $\log^{O(1)}(d)$ (for the other hyperedges, which are copies of W^1). A hyperedge containing $\{v, x\}$ and i infected vertices is either a copy in $\mathcal{H}(\ell)$ of W^2 rooted at $\{v, x\}$ (when $i = r - 2$) or a copy in $\mathcal{H}(\ell)$ of a secondary configuration $X = (\mathcal{F}, R, D)$ where \mathcal{F} is a single hyperedge, $|R| = 2$ and $|D| = i$ (when $i < r - 2$). So as $\omega \notin \mathcal{B}_\ell$, by (B.2) and (B.3), the maximum number of copies that can be destroyed in a time step is

$$\max\{\log^{O(1)}(d), \log^{O(1)}(d) \cdot d^{1-\frac{i+1}{r-1}} \log^{-3K/5}(d)\} = o(\eta). \quad (6.55)$$

As $t = O(\log d)$, $N \geq d^{\frac{1}{r-1}} \log^{bK}(d)$ and K is chosen to be large with respect to r, c and α ,

$$a_\ell^\pm = O\left(\frac{\log^{O(1)}(d) d^{1-\frac{i}{r-1}}}{N}\right) = o(\eta). \quad (6.56)$$

So by (6.54), (6.55) and (6.56), both $B_v^+(0), \dots, B_v^+(M)$ and $B_v^-(0), \dots, B_v^-(M)$ are η -bounded. \square

Now we bound the variance of $A_v^\pm(\ell)$ given \mathcal{F}_ℓ .

Subclaim 6.57.

$$\sum_{\ell=0}^{m-1} \text{Var}(A_S(\ell) \mid \mathcal{F}_\ell) \leq d^{2-\frac{2i}{r-1}} \log^{-9K/20}(d).$$

Proof. The calculation follows similarly to the calculations for Claim 6.17 and Claim 6.27. When $\omega \in \mathcal{B}_\ell$, we have $\text{Var}(A_S(\ell) \mid \mathcal{F}_\ell) = 0$. So now consider when $\omega \notin \mathcal{B}_\ell$. We have

$$\begin{aligned} \text{Var}(A_S^\pm(\ell) \mid \mathcal{F}_\ell) &\leq \mathbb{E}\left(\left(Y_v^{i,j}(\ell+1) - Y_v^{i,j}(\ell)\right)^2 \mid \mathcal{F}_\ell\right) \\ &\leq \mathbb{E}\left(\left|Y_v^{i,j}(\ell+1) - Y_v^{i,j}(\ell)\right| \mid \mathcal{F}_\ell\right) \cdot \sup |Y_v^{i,j}(\ell+1) - Y_v^{i,j}(\ell)|. \end{aligned} \quad (6.58)$$

In the proof of Subclaim 6.53 we saw that the maximum value of $|Y_v^{i,j}(\ell+1) - Y_v^{i,j}(\ell)|$ is η . So we have

$$\text{Var}(A_S^\pm(\ell) \mid \mathcal{F}_\ell) \leq \eta \cdot \mathbb{E}\left(\left|Y_v^{i,j}(\ell+1) - Y_v^{i,j}(\ell)\right| \mid \mathcal{F}_\ell\right).$$

By expressing $Y_v^{i,j}(\ell+1) - Y_v^{i,j}(\ell)$ as $C_{\ell+1} - D_{\ell+1}$ as in the proof of Subclaim 6.39, we have

$$\mathbb{E}\left(\left|Y_v^{i,j}(\ell+1) - Y_v^{i,j}(\ell)\right| \mid \mathcal{F}_\ell\right) \leq \mathbb{E}(C_{\ell+1} + D_{\ell+1} \mid \mathcal{F}_\ell).$$

By (6.50), (6.51) and definition of η , this is at most

$$\begin{aligned} \eta(1 + 20\epsilon(t_\ell)) d^{1-\frac{i}{r-1}} &\left(\frac{(j+1)\alpha y_{i-1,j+1}(t_\ell)}{\gamma(t_\ell)N} + \frac{(r-i-j)\alpha y_{i,j-1}(t_\ell) y_{r-2,1}(t_\ell)}{\gamma(t_\ell)N} + \frac{j\alpha y_{i,j}(t_\ell)}{\gamma(t_\ell)N} \right) \\ &= o\left(\frac{d^{2-\frac{2i}{r-1}}}{M \log^{9K/20}(d)}\right). \end{aligned}$$

So the sum of $\text{Var}(A_v^\pm(\ell) \mid \mathcal{F}_\ell)$ over $0 \leq \ell \leq M$ is at most $d^{2-\frac{2i}{r-1}} \log^{-9K/20}(d)$, as required. \square

Set

$$\nu := d^{2-\frac{2i}{r-1}} \log^{-9K/20}(d),$$

and set $a := \frac{1}{2}\epsilon(t_m)y_{i,j}(t_m)d^{1-\frac{i}{r-1}}$. Theorem 3.10 implies that

$$\mathbb{P}(B_v^-(m) \geq a) \leq \exp\left(-\frac{a^2}{2(\nu + a\eta)}\right) \ll N^{-5\sqrt{\log(N)}}$$

and

$$\mathbb{P}(B_v^+(m) \leq -a) \leq \exp\left(-\frac{a^2}{2(\nu + a\eta)}\right) \ll N^{-5\sqrt{\log(N)}}.$$

This completes the proof of Claim 6.34. \square

This claim was the final piece in the proof of Lemma 6.29. Proving Lemma 6.29 completes the proof of Theorem 6.2 and concludes our discussion of the first phase.

5.7 The Second Phase in the Subcritical Case

In this section we will complete the proof of Theorem 1.6 in the subcritical case. It may be helpful to recall the definition of the processes we run in the second phase from Subsection 5.2.2 and the definition of \mathcal{F}'_m from Subsection 5.3.4. Remember that in the second phase we ‘restart the clock’ again from time zero letting $\mathcal{H}(0) := \mathcal{H}(M)$, $I(0) := I(M)$.

We begin by presenting some technical definitions of various constants that will be used throughout this section. These constants are required to state our main lemmas and are useful to simplify future calculations.

First we present some further analysis of the function $\gamma(t)$ (defined in (2.8)). We have

$$\gamma'(t) = \alpha(r-1)(c + \alpha t)^{r-2} - 1.$$

Recall that t_{\min} was chosen to be the unique positive root of $\gamma'(t)$ and that $T_0 < t_{\min} < T_1$ where T_0, T_1 are the only two positive roots of $\gamma(t)$. As $\gamma(0) > 0$, $\gamma(T_0) = 0$ and $T_0 < t_{\min}$, the function $\gamma(t)$ is strictly decreasing in the interval $[0, t_{\min}]$ and so $\gamma'(T_0) < 0$. Therefore $\alpha(r-1)(c + \alpha T_0)^{r-2} - 1 < 0$, which implies that

$$(r-1)(c + \alpha T_0)^{r-2} < \frac{1}{\alpha}.$$

So there exists a constant $0 < \lambda < 1/8$ such that

$$(r-1)(c + \alpha T_0)^{r-2} < \frac{1-4\lambda}{\alpha}. \tag{7.1}$$

Recall from (2.10) and (2.11) that we run the first phase process until time T , where

$$T := \frac{1}{N} \min\{m \geq 0 : (1 + \epsilon(t_m))\gamma(t) < \zeta\}$$

and ζ will be chosen in a moment (in Definition 7.3). Using the fact that $T < T_0$ and that $(r-1)(c + \alpha t)^{r-2}$ is increasing for positive t , we obtain the following from (7.1),

$$(r-1)(c + \alpha T)^{r-2} < \frac{1 - 4\lambda}{\alpha}. \quad (7.2)$$

We are now ready to define ζ and some other useful constants.

Definition 7.3. For $0 \leq i \leq r-2$, define

$$\chi(i, 0) := \binom{r-1}{i} \left(\frac{1-4\lambda}{\alpha(r-1)} \right)^{\frac{i}{r-2}} + \frac{\lambda}{\alpha}.$$

Also, define

$$\begin{aligned} \chi_{\max} &:= \max_{0 \leq i \leq r-2} \chi(i, 0), \\ \chi_{\min} &:= \min_{0 \leq i \leq r-2} \chi(i, 0), \\ \zeta &:= \frac{\lambda^2 \chi_{\min}}{r^{6r+1} (1 + \alpha + \alpha^r)^2 (1 + \chi_{\max})^3}, \end{aligned}$$

and

$$\chi(r-2, 1) := (1 + \chi_{\max}) \cdot \zeta.$$

For $0 \leq i \leq r-3$ and $1 \leq j \leq r-i-1$, define

$$\chi(i, j) := r^{3r} \chi(i, 0) (1 + \alpha^j) \chi(r-2, 1)^j.$$

Note that as $\lambda < 1$ and $\chi_{\min} < 1 + \chi_{\max}$,

$$\chi(r-2, 1) := (1 + \chi_{\max}) \cdot \zeta = \frac{\lambda^2 \chi_{\min}}{r^{6r+1} (1 + \alpha + \alpha^r)^2 (1 + \chi_{\max})^2} < 1. \quad (7.4)$$

Note in addition that

$$\chi(r-2, 0) = \frac{1-3\lambda}{\alpha}. \quad (7.5)$$

Also observe that as λ is a function of r, c and α , so is ζ . Before presenting the main results and proofs of the section, let us deduce a simple fact about $\chi(i, j)$ which will be used a number of times below.

Proposition 7.6. For $0 \leq i \leq r-3$ and $1 \leq j \leq r-i-1$,

$$\chi(i, j) < \frac{\lambda^2}{r^{3r+1} \alpha^{j+1}}.$$

Proof. Note that, for any $\alpha > 0$ and $1 \leq j \leq r-1$, we have that $1 + \alpha + \alpha^r$ is greater than both $1 + \alpha^j$ and α^{j+1} . Thus, for $0 \leq i \leq r-3$ and $1 \leq j \leq r-1-i$, we have

$$\begin{aligned}
\chi(i, j) &= [r^{3r} \chi(i, 0)(1 + \alpha^j) \chi(r-2, 1)] \chi(r-2, 1)^{j-1} \\
&= [r^{3r} \chi(i, 0)(1 + \alpha^j)(1 + \chi_{\max}) \cdot \zeta] \chi(r-2, 1)^{j-1} \\
&= \left(\frac{r^{3r} \chi(i, 0)(1 + \alpha^j) \lambda^2 \chi_{\min}}{r^{6r+1} (1 + \alpha + \alpha^r)^2 (1 + \chi_{\max})^2} \right) \chi(r-2, 1)^{j-1} \\
&< \left(\frac{\lambda^2 \chi_{\min}}{r^{3r+1} (1 + \alpha + \alpha^r) (1 + \chi_{\max})} \right) \chi(r-2, 1)^{j-1} \\
&< \frac{\lambda^2}{r^{3r+1} \alpha^{j+1}},
\end{aligned} \tag{7.7}$$

as by (7.4), $\chi(r-2, 1) < 1$. □

Throughout the section we will also use the following function.

$$\psi(m) := \max \{ (1 - \lambda)^m, \log^{-10}(N) \}. \tag{7.8}$$

We have now completed the tedious technicalities and are able to present the real meat of the section. Let us briefly outline how the proof will proceed. For each time m , we will define \mathcal{A}_m (similarly to how we defined \mathcal{B}_m for the first phase) to be the event that some bound on the number of copies of a particular configuration in $\mathcal{H}(m)$ fails to hold. \mathcal{A}_0 will be formally defined in Definition 7.13 and, for $0 < m \leq M_2$, \mathcal{A}_m will be defined in Definition 7.14.

We will use the following lemma. (Recall Definition 2.33 for the definition of $Z^{i,j}$.)

Lemma 7.9. *Let $0 \leq m \leq M_2$. When $\omega \notin \mathcal{A}_m$ and K is sufficiently large, the following statement holds. For all S and v contained in $V(\mathcal{H})$:*

(L.1) For $0 \leq i \leq r-2$,

$$Y_v^{i,0}(m) \leq \chi(i, 0) d^{1 - \frac{i}{r-1}}.$$

(L.2) For $0 \leq i \leq r-2$ and $1 \leq j \leq r-1-i$,

$$Y_v^{i,j}(m) \leq \psi(m) \chi(i, j) d^{1 - \frac{i}{r-1}}.$$

(L.3) For $1 \leq i \leq r$,

$$W_S^i(m) \leq (m+1) \log^{r^3(r-i)}(d).$$

(L.4) For every secondary configuration $X = (\mathcal{F}, R, D)$,

$$X_S(m) \leq (m+1) \log^{2|D|r^7}(d) \cdot d^{\frac{|V(\mathcal{F})| - |R| - |D|}{r-1}} \log^{-3K/5}(d).$$

(L.5) For $0 \leq i \leq r-2$ and $1 \leq j \leq r-1-i$,

$$Z_v^{i,j}(m) \leq \left(1 + \frac{\lambda^2}{r^{3r}}\right) \psi(m) \chi(i, j) d^{1 - \frac{i}{r-1}}.$$

We will also prove that with high probability \mathcal{A}_{M_2} does not occur.

Theorem 7.10.

$$\mathbb{P}(\mathcal{A}_{M_2}) \leq N^{-\sqrt{\log(N)}}.$$

As we will see in Definition 7.14, the event \mathcal{A}_{M_2} contains the events \mathcal{A}_m for $0 \leq m < M_2$. Thus Theorem 7.10 implies that with high probability for all S and v contained in $V(\mathcal{H})$ and all $0 \leq m \leq M_2$ the bounds (L.1)-(L.5) hold.

Given Theorem 7.10 and Lemma 7.9, we now prove Theorem 2.31, which completes the proof of Theorem 1.6 in the subcritical case. After this, we will focus on the proof of the theorem and the lemma.

Proof of Theorem 2.31. Our goal is to show that, with probability $1 - o(1)$, we have

- (i) $Q(M_2) = \emptyset$, and
- (ii) $|I(M_2)| = \frac{N \cdot \log^{O(1)}(N)}{d^{1/(r-1)}}.$

From this, it follows easily that the probability of percolation is at most ε .

For (i), by Markov's Inequality (Theorem 3.1), we have

$$\mathbb{P}(Q(M_2) > 0) \leq \mathbb{E}(Q(M_2)).$$

Thus, it suffices to show that the right side is $o(1)$. We will prove the following.

Claim 7.11. For $0 \leq m \leq M_2$,

$$\mathbb{E}(Q(m)) \leq 2(1 - \lambda)^m \zeta \cdot N.$$

Proof. The proof proceeds by induction on m . First consider the base case $m = 0$. By definition of T , Theorem 6.2 and Lemma 4.1, we have $Q(0) \leq \zeta \cdot N$ with probability at least $1 - N^{-2\sqrt{\log(N)}}$. As \mathcal{H} has maximum degree d , we have $Q(0) \leq N \cdot d$ and so

$$\mathbb{E}(Q(0)) \leq \zeta \cdot N + N \cdot d \cdot N^{-2\sqrt{\log N}} \leq 2\zeta \cdot N,$$

as required.

Now suppose that the statement of the claim is proved for all $0 \leq \ell \leq m$ and we wish to show it holds for $m+1$. If $Q(m) = \emptyset$ we are done, so we may assume that $Q(m) \neq \emptyset$. Recall that in each round of the second phase, every open hyperedge is sampled.

Thus, $Q(m+1) \cap Q(m) = \emptyset$. This implies that, for each $e \in Q(m+1)$, there must be at least one vertex x of e such that $x \in I(m+1) \setminus I(m)$; that is, $x \notin I(m)$ and there is a hyperedge $e^* \in Q(m)$ containing x which was successfully sampled in the m th step. Therefore, conditioned on \mathcal{F}'_m , the expectation of $Q(m+1)$ is at most the product of

- the number of ways to choose a hyperedge $e^* \in Q(m)$ containing a vertex $x \notin I(m)$, and
- the sum of $Z_x^{r-2-j,j}(m)q^{j+1}$ over all $0 \leq j \leq r-2$.

When $\omega \notin \mathcal{A}_m$ (and K is sufficiently large), by Lemma 7.9 for every $w \in V(\mathcal{H})$ and $1 \leq j \leq r-2$, we have the following two bounds. First, using (L.5) we have

$$\begin{aligned} Z_w^{r-2-j,j}(m)q^{j+1} &\leq \left(1 + \frac{\lambda^2}{r^{3r}}\right) \psi(m)\chi(r-2-j, j)d^{\frac{j+1}{r-1}}q^{j+1} \leq 2\chi(r-2-j, j)\alpha^{j+1} \\ &\leq 2\left(\frac{\lambda^2}{r^{3r+1}\alpha^{j+1}}\right)\alpha^{j+1} < \frac{\lambda}{2r}. \end{aligned}$$

and by (7.5) and (L.1) we have

$$Z_w^{r-2,0}(m)q = Y_w^{r-2,0}(m)q \leq \chi(r-2, 0)\alpha = 1 - 3\lambda.$$

So when $\omega \notin \mathcal{A}_m$, we have

$$\begin{aligned} \mathbb{E}(Q(m+1) \mid \mathcal{F}'_m) &\leq \sum_{Q \in Q(m)} \sum_{w \in Q \setminus I(m)} \sum_{j=0}^{r-2} Z_w^{r-2-j,j}(m)q^{j+1} \\ &\leq Q(m) \left((1 - 3\lambda) + \sum_{j=1}^{r-2} \frac{\lambda}{2r} \right) \\ &\leq Q(m) (1 - 2\lambda). \end{aligned}$$

As \mathcal{H} has maximum degree d , we have $Q(m) \leq N \cdot d$. So, letting $\mathbb{1}_{\mathcal{E}}$ be the indicator function of the event \mathcal{E} occurring, we have

$$\begin{aligned} \mathbb{E}(Q(m+1) \mid \mathcal{F}'_m) &\leq Q(m) (1 - 2\lambda) \cdot \mathbb{1}_{\mathcal{A}_m^c} + N \cdot d \cdot \mathbb{1}_{\mathcal{A}_m} \\ &\leq Q(m) (1 - 2\lambda) + N \cdot d \cdot \mathbb{1}_{\mathcal{A}_m}. \end{aligned} \tag{7.12}$$

So using Theorem 7.10 to bound $\mathbb{P}(\mathcal{A}_m)$, by (7.12) and the law of iterated expectation we have

$$\begin{aligned} \mathbb{E}(Q(m+1)) &= \mathbb{E}(\mathbb{E}(Q(m+1) \mid \mathcal{F}'_m)) \\ &\leq (1 - 2\lambda) \mathbb{E}(Q(m)) + N^{-\sqrt{\log N}} \cdot N \cdot d \end{aligned}$$

Applying our induction hypothesis gives that this is at most

$$\begin{aligned} (1 - 2\lambda) 2(1 - \lambda)^m \zeta \cdot N + o(1) &\leq (1 - \lambda)^{m+1} 2\zeta \cdot N - \lambda(1 - \lambda)^m \zeta N + o(1) \\ &\leq (1 - \lambda)^{m+1} 2\zeta \cdot N. \end{aligned}$$

This completes the proof of the claim. □

So by Claim 7.11 and our choice of M_2 (see 2.30),

$$\mathbb{E}(Q(M_2)) \leq (1 - \lambda)^{M_2} 2\zeta N \leq \frac{2\zeta N}{N^2} = o(1).$$

So, (i) holds with probability $1 - o(1)$.

To complete the proof we show that (ii) also holds with probability $1 - o(1)$. By Proposition 2.6, with probability $1 - N^{-\Omega(\sqrt{\log N})}$ the number of vertices infected during the first phase is $O\left(\frac{N \cdot \log(N)}{d^{1/(r-1)}}\right)$.

By (L.3) of Lemma 7.9 and Theorem 7.10 and letting K be large, with probability at least $1 - N^{-\sqrt{\log(N)}}$ for every $0 \leq m \leq M_2$ and $v \in V(\mathcal{H}) \setminus I(m)$, we have $Q_v(m) = \log^{O(1)}(N)$. Using this, with high probability the number of vertices sampled in each round is $N \cdot \log^{O(1)}(N)$. So as $M_2 = \log^{O(1)}(d)$ by (2.30), the expected number of vertices infected during phase two is $\frac{N \cdot \log^{O(1)}(N)}{d^{1/(r-1)}}$. Applying the Chernoff bound, by (2.1) we get that with probability at least $1 - N^{-100}$ there are at most $O\left(\frac{N \cdot \log^{O(1)}(N)}{d^{1/(r-1)}}\right)$ vertices infected during phase two. So overall, with high probability there are at most $O\left(\frac{N \cdot \log^{O(1)}(N)}{d^{1/(r-1)}}\right) = o(N)$ infected vertices when the process terminates. This completes the proof. □

The remainder of the subsection is devoted to proving Lemma 7.9 and Theorem 7.10.

5.7.1 Defining \mathcal{A}_m and proof of Lemma 7.9

We begin by formally defining the events \mathcal{A}_m , for $0 \leq m \leq M_2$. It will be convenient for the proof of Theorem 7.10 to define \mathcal{A}_0 separately. Recall the definitions of λ and $\chi(i, j)$ from (7.1) and Definition 7.3, respectively.

Definition 7.13. Let \mathcal{A}_0 be the event (in Ω' , which was defined in Subsection 5.3.4) that for some $v \in V(\mathcal{H})$ or $S \subseteq V(\mathcal{H})$ one of the following statements fails to hold.

(A0.1) For all $0 \leq i \leq r - 2$,

$$Y_v^{i,0}(0) \leq \left(\chi(i, 0) - \frac{\lambda}{\alpha}\right) d^{1 - \frac{i}{r-1}}.$$

(A0.2) For $0 \leq i \leq r - 2$ and $1 \leq j \leq r - 1 - i$,

$$Y_v^{i,j}(0) < \chi(i, j) d^{1 - \frac{i}{r-1}}.$$

(A0.3) For $1 \leq i \leq r$,

$$W_S^i(0) \leq \log^{r^3(r-i)}(d).$$

(A0.4) For any secondary configuration $X = (\mathcal{F}, R, D)$,

$$X_S(0) \leq \log^{2|D|r^4}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d).$$

We remark that the case $i = 0$ of (L.1) in Lemma 7.9 always holds trivially as $\Delta(\mathcal{H}) \leq d$. This is reflected below in our definition of \mathcal{A}_m .

Definition 7.14. For $1 \leq m \leq M_2$, let \mathcal{A}_m be the event (in Ω' , which was defined in Subsection 5.3.4) that either \mathcal{A}_0 occurs, or there exists $1 \leq \ell \leq m$ such that, for some $v \in V(\mathcal{H})$ or $S \subseteq V(\mathcal{H})$ one of the following statements fails to hold:

(A.1) For $1 \leq i \leq r - 2$,

$$Y_v^{i,0}(\ell) - Y_v^{i,0}(\ell - 1) \leq \psi(\ell - 1) \left(\frac{\lambda^2}{2\alpha} \right) d^{1 - \frac{i}{r-1}}.$$

(A.2) For $0 \leq i \leq r - 2$ and $1 \leq j \leq r - 1 - i$,

$$Y_v^{i,j}(\ell) \leq \psi(\ell) \chi(i, j) d^{1 - \frac{i}{r-1}}.$$

(A.3) For $1 \leq i \leq r$,

$$W_S^i(\ell) - W_S^i(\ell - 1) \leq \log^{r^3(r-i)}(d).$$

(A.4) For every secondary configuration $X = (\mathcal{F}, R, D)$,

$$X_S(\ell) - X_S(\ell - 1) \leq \log^{2|D|r^7}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d).$$

We now present the proof of Lemma 7.9.

Proof of Lemma 7.9. We first show that if $\omega \notin \mathcal{A}_m$, then (L.1)-(L.4) hold. When $m = 0$ this follows by definition of \mathcal{A}_0 as $\psi(0) = 1$. Now consider $m > 0$.

Note that (A.2) is precisely the same statement as property (L.2). Observe in addition that if $\omega \notin \mathcal{A}_m$, then clearly (L.3) and (L.4) hold. Also, if $\omega \notin \mathcal{A}_m$, then using the definition

of ψ (see (7.8)),

$$\begin{aligned} Y_v^{i,0}(m) &= Y_v^{i,0}(0) + \sum_{\ell=1}^m (Y_v^{i,0}(\ell) - Y_v^{i,0}(\ell - 1)) \\ &\leq \left(\chi(i, 0) - \frac{\lambda}{\alpha} + \left(\frac{\lambda^2}{2\alpha} \right) \sum_{\ell=1}^{M_2} \psi(\ell - 1) \right) d^{1-\frac{i}{r-1}} \\ &\leq \left(\chi(i, 0) - \frac{\lambda}{\alpha} + \left(\frac{\lambda^2}{2\alpha} \right) \left(\sum_{\ell=0}^{\infty} (1 - \lambda)^\ell + \sum_{\ell=1}^{M_2} \log^{-10}(N) \right) \right) d^{1-\frac{i}{r-1}} \\ &\leq \left(\chi(i, 0) - \frac{\lambda}{\alpha} + \left(\frac{\lambda^2}{2\alpha} \right) \left(\frac{1}{\lambda} + o(1) \right) \right) \\ &\leq \chi(i, 0) d^{1-\frac{i}{r-1}}, \end{aligned}$$

since $M_2 = O(\log N)$. Hence, $\omega \notin \mathcal{A}_m$ implies that (L.1) holds.

We now show that if (L.1)-(L.4) hold at time m , then so does (L.5). By Observation 2.34, each member of $Z_v^{i,j}(m)$ is either a member of $Y_v^{i,j}(m)$ or consists of a copy \mathcal{F}' of a secondary configuration with $r - 1 - i$ neutral vertices and a set of copies of W^1 rooted at vertices of \mathcal{F}' . By (L.3), (L.4) and the fact that $m \leq M_2 = O(\log N)$ (from (2.2) and (2.30)), the number of members of this type is at most

$$\log^{O(1)}(d) \cdot d^{1-\frac{i}{r-1}} \cdot \log^{-3K/5}(d),$$

This is $o\left(\psi(m)d^{1-\frac{i}{r-1}}\right)$, provided that K is sufficiently large. The result now follows by the bound on $Y_v^{i,j}(m)$ given by (L.2). \square

5.7.2 Proof of Theorem 7.10

Given a point $\omega \in \mathcal{A}_{M_2} \setminus \mathcal{A}_0$, let

$$\mathcal{J} = \mathcal{J}(\omega) := \min\{i : \omega \in \mathcal{A}_i\}.$$

Define \mathcal{Y}^0 to be the set of $\omega \in \mathcal{A}_{M_2} \setminus \mathcal{A}_0$ such that (A.1) is violated at time $\mathcal{J}(\omega)$ for some $v \in V(\mathcal{H})$ (the superscript 0 represents that $j = 0$ in the configurations we consider in (A.1)). Similarly, define $\mathcal{Y}^{>0}$, \mathcal{W} and \mathcal{X} respectively to be the events that (A.2), (A.3) and (A.4) are violated at time $\mathcal{J}(\omega)$.

So the event \mathcal{A}_{M_2} is contained within the events \mathcal{A}_0 , and $\mathcal{Y}^0 \cup \mathcal{Y}^{>0} \cup \mathcal{W} \cup \mathcal{X}$. First we bound the probability of \mathcal{A}_0 occurring. The following Lemma is an easy consequence of Theorem 6.2.

Lemma 7.15.

$$\mathbb{P}(\mathcal{A}_0) \leq N^{-2\sqrt{\log(N)}}.$$

Proof. First suppose $\omega \notin \mathcal{A}_0$. By definition of \mathcal{B}_M (see Definition 6.1), with probability at least $1 - N^{-2\sqrt{\log N}}$ the bounds (B.1), (B.2) and (B.3) hold for all $v \in V(\mathcal{H})$ and $S \subseteq V(\mathcal{H})$ when $\ell = M$.

When (B.1) holds, for $0 \leq i \leq r - 2$ and $0 \leq j \leq r - 1 - i$, we have

$$Y_v^{i,j}(0) \leq (1 + \epsilon(T)) \binom{r-1}{i} \binom{r-1-i}{j} (c + \alpha T)^i \gamma(t)^j d^{1-\frac{i}{r-1}}$$

By (7.2) and the fact that $(1 + \epsilon(T))\gamma(t) < \zeta$ (by (2.30)), the above inequality implies that

$$\begin{aligned} Y_v^{i,j}(0) &\leq \binom{r-1}{i} \binom{r-1-i}{j} \left(\frac{1-4\lambda}{\alpha(r-1)} \right)^{\frac{i}{r-2}} \zeta^j d^{1-\frac{i}{r-1}} \\ &\leq \left(\chi(i,0) - \frac{\lambda}{\alpha} \right) \binom{r-1-i}{j} \zeta^j d^{1-\frac{i}{r-1}}. \end{aligned} \quad (7.16)$$

For $0 \leq i \leq r - 2$ and $j = 0$, (7.16) implies that

$$Y_v^{i,0}(0) \leq \left(\chi(i,0) - \frac{\lambda}{\alpha} \right) d^{1-\frac{i}{r-1}},$$

as required for (A0.1). Also, if $i = r - 2$ and $j = 1$, (7.16) implies that

$$Y_v^{r-2,1}(0) \leq \chi(r-2,0) \cdot \zeta \cdot d^{\frac{1}{r-1}} < \chi(r-2,1) d^{\frac{1}{r-1}}, \quad (7.17)$$

thus proving (A0.2) when $i = r - 2$. Now, for $0 \leq i \leq r - 3$ and $1 \leq j \leq r - 1 - i$, using (7.16) and the fact that $\zeta < \chi(r-2,1)$, gives

$$Y_v^{i,j}(0) \leq \chi(i,0) \binom{r-1-i}{j} \zeta^j \cdot d^{1-\frac{i}{r-1}} < \chi(i,j) d^{1-\frac{i}{r-1}},$$

completing the proof of (A0.2). For $1 \leq i \leq r$ and $S \subseteq V(\mathcal{H})$, when (B.3) holds at $\ell = M$,

$$W_S^i(0) \leq \log^{r^3(r-i)}(d),$$

thus proving (A0.3). Finally, for any secondary configuration $X = (\mathcal{F}, R, D)$ and $S \subseteq V(\mathcal{H})$, when (B.2) holds at $\ell = M$,

$$X_S(0) \leq \log^{2|D|r^4}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d),$$

as required for (A0.4). □

Our goal now is to show that the probability of each of \mathcal{Y}^0 , $\mathcal{Y}^{>0}$, \mathcal{W} and \mathcal{X} occurring is $N^{-10\sqrt{\log N}}$, from which Theorem 7.10 will follow. This will be done in Propositions 7.25, 7.31, 7.33 and 7.36 to come.

Given a configuration X that we care about controlling, we wish to apply Corollary 3.6 to bound the probability of $X_S(m+1) - X_S(m)$ being too large for each $S \subseteq V(\mathcal{H})$. It

will be helpful to introduce some general framework, which will aid us in our application of Corollary 3.6. Before stating the next lemma, we require quite a few technical definitions. Let us briefly motivate the definitions before stating them formally.

For a fixed configuration $X = (\mathcal{F}, R, D)$ and non-empty $U \subseteq D$, we will define a family X^U of configurations, such that each configuration in the family is created by changing the set U of marked vertices of X into neutral vertices and, for each $u \in U$, adding a hyperedge e_u containing u such that the vertices in $e_u \setminus \{u\}$ are all marked. The family X^U will be helpful when bounding how many new copies of X are made at some time step. Indeed, a new copy of X is made from a copy of a configuration in X^U when the open hyperedges rooted at vertices of U are all successfully sampled in some time step.

We will also define a configuration $X^{U,U'}$ created by taking some set U of marked vertices of X , turning some subset $U' \subseteq U$ into roots and turning $U \setminus U'$ into neutral vertices. This will be used to bound $\partial_A f$ (where f will be an upper bound on $X_S(m+1) - X_S(m)$ and $A \subseteq V(\mathcal{H})$) in the application of Corollary 3.6. As it turns out, we will always be able to express bounds on the number of copies of some $X' \in X^U$ or $X^{U,U'}$ in terms of copies of configurations we are keeping control over.

For a configuration $X = (\mathcal{F}, R, D)$, call a hyperedge of \mathcal{F} with $r - 1$ vertices in D *unstable*. If \mathcal{F} contains an unstable hyperedge then every copy of X in $\mathcal{H}(m)$ is destroyed in the $(m + 1)$ th time step, as at each time step every open hyperedge is sampled and deleted from our hypergraph. So in particular, a new copy of X can only be made from some X^U such that U intersects every unstable hyperedge of \mathcal{F} . Call such a set U *fruitful*.

We apologise that the following set of definitions are fairly technical. But the introduction of these concepts and Lemma 7.20 will greatly simplify and clarify the calculations that are to come in the proof of Lemma 7.9.

Definition 7.18. Let $X = (\mathcal{F}, R, D)$ be a configuration and let $\emptyset \neq U \subseteq D$ be fruitful and $U' \subseteq U$. Define $X^{U,U'}$ to be the configuration $X = (\mathcal{F}, R \cup U', D \setminus U)$.

Say that $U_1, U_2 \subseteq D$ are *non-isomorphic* if the configurations $(\mathcal{F}, R, D \setminus U_1)$ and $(\mathcal{F}, R, D \setminus U_2)$ are not isomorphic. Define $\mathcal{U}(X)$ to be the set of all non-isomorphic non-empty fruitful subsets of D .

For $k, a \geq 0$, define

$$X^{k,a} := \bigcup_{\substack{U \in \mathcal{U}(X) \\ U' \subseteq U \\ |U| - |U'| = k \\ |U'| = a}} X^{U,U'}$$

Also, define X^U to be the set of all configurations $X' = (\mathcal{F}', R', D')$ such that

- (a) $\mathcal{F}' \supseteq \mathcal{F}$, $R' = R$ and $D' \cap D = D \setminus U$ (so the vertices of U are neutral here);

- (b) $E(\mathcal{F}) \setminus E(\mathcal{G}) = \{e_u : u \in U\}$, where for each $u \in U$, e_u is an unstable hyperedge of \mathcal{F} containing u . (So $D' := D \setminus U \cup \bigcup_{u \in U} e_u \setminus \{u\}$.)

See Figure 5.13 for an example of a configuration in X^U and a configuration $X^{U,U'}$.

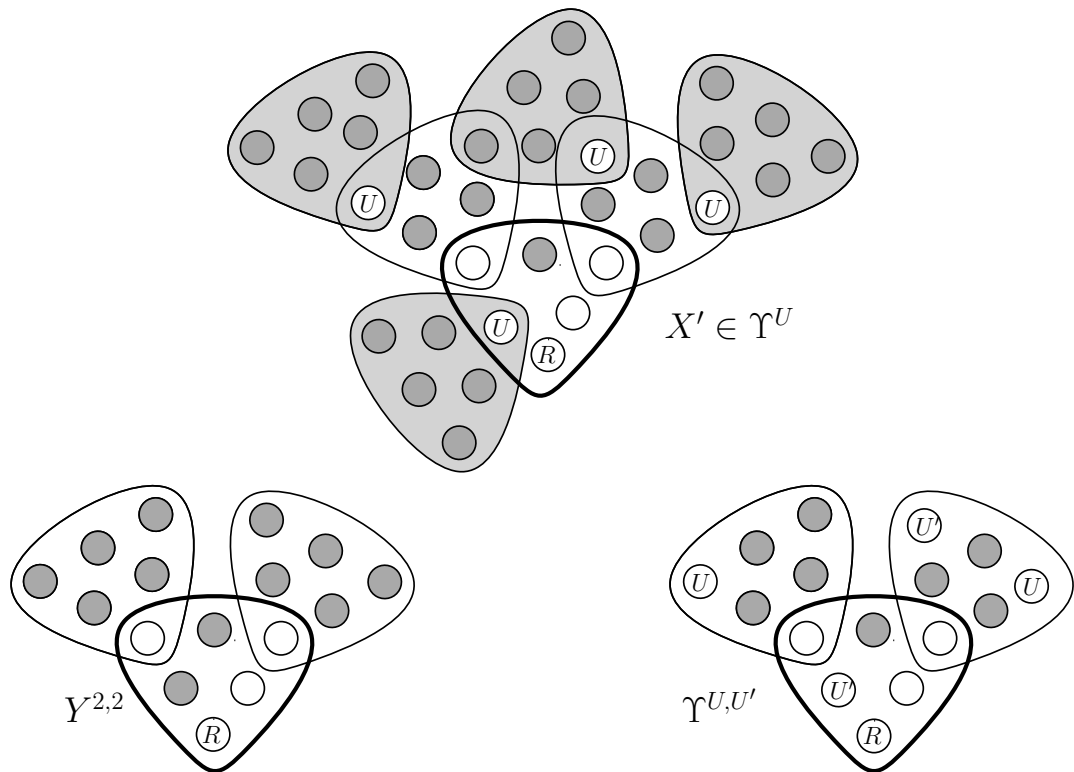


Figure 5.13: Here $\Upsilon := Y^{2,2}$ and $r = 6$. For a particular choice of fruitful $U \subseteq D$ and $U' \subseteq U$, we have an example of some configuration $X' \in \Upsilon^U$ and a configuration $\Upsilon^{U,U'}$. The vertices of U' are labelled with a U' and the vertices of $U \setminus U'$ are labelled with a U . The root is labelled by R , the marked vertices are dark grey, the neutral vertices are white and the central hyperedge is drawn with a thick outline. The hyperedges $\bigcup_{u \in U} e_u$ are shaded light grey. Observe that if each of these hyperedges in a copy of X' is successfully sampled in the $(m + 1)$ th step, then this creates a new copy of $Y^{2,2}$ in $\mathcal{H}(m + 1)$.

Throughout the rest of the section we will condition on \mathcal{F}'_m and assume that $\omega \notin \mathcal{A}_m$. We may not always write explicitly that we are assuming this and conditioning on \mathcal{F}'_m . We will use the following random variables throughout. Given $w \in V(\mathcal{H})$, let ξ_w be the Bernoulli random variable which is equal to one if $w \in I(m + 1) \setminus I(m)$ and zero otherwise. Also, for $e \in Q(m)$, we let ξ_e be the Bernoulli random variable which is equal to one if e

is successfully sampled and zero otherwise. Clearly, for each $v \notin I(m)$, we have that ξ_v is equal to one if and only if $\xi_e = 1$ for some $e \in Q_v(m)$. From this (and as we are conditioning on \mathcal{F}'_m), we have that ξ_u is independent of ξ_w for u, w distinct vertices of $V(\mathcal{H}) \setminus I(m)$. Also, we have

$$\xi_v \leq \sum_{e \in Q_v(m-1)} \xi_e. \quad (7.19)$$

Lemma 7.20. *Let $X = (\mathcal{F}, R, D)$ be a configuration and let S be a subset of $V(\mathcal{H})$ such that $|S| = |R|$. If (L.3) holds at time m , then there exists a polynomial $f(x_w : w \in V(\mathcal{H}))$ of degree at most $|D|$ where no variable has an exponent greater than 1 such that:*

$$(i) \quad X_S(m+1) - X_S(m) \leq |X_S(m+1) \setminus X_S(m)| \leq f(\xi_w : w \in V(\mathcal{H}))$$

(ii) For any set $A \subseteq V(\mathcal{H})$,

$$\mathbb{E}(\partial_A f(x_w : w \in V(\mathcal{H})) \mid \mathcal{F}'_m) \leq \sum_{k=0}^{|D|-|A|} \sum_{X' \in X^{k, |A|}} X'_{S \cup A}(m) (\log^{r^4}(d)q)^k.$$

(iii)

$$\mathbb{E}(f(x_w : w \in V(\mathcal{H})) \mid \mathcal{F}'_m) \leq \sum_{k=1}^{|D|} \sum_{\substack{U \in \mathcal{U}(X) \\ |U|=k}} \sum_{X' \in X^U} X'_S(m) q^k.$$

We remark that (ii) does give a bound on everything we need to apply Corollary 3.6. However, when dealing with W^1 and the Y configurations we need a more careful bound on the expected change and this is why we have (iii).

Proof of Lemma 7.20. If (L.3) holds at time m , as $m = O(\log d)$ (by (2.30) and (2.2)), by (7.19) we have

$$\mathbb{E}(\xi_v \mid \mathcal{F}'_m) \leq q \cdot (m+1) \log^{r^3(r-1)}(d) \leq q \log^{r^4}(d). \quad (7.21)$$

Let us consider how a copy of X rooted at S which is present in the $(m+1)$ th step, but not the m th step is created. Such a copy can only come from a subhypergraph $\mathcal{F}' \subseteq \mathcal{H}$ containing S that is like a copy of X missing some infections. \mathcal{F}' will become a new copy of X when it gains these ‘missing’ infections.

More formally, a new copy comes from a subhypergraph $\mathcal{F}' \subseteq \mathcal{H}(m)$, a non-empty fruitful set $U \subseteq D$ and subsets $D', V' \subseteq V(\mathcal{F}')$ such that:

- (1) there exists an isomorphism ϕ from \mathcal{F} to \mathcal{F}' with $\phi(R) = S$, $\phi(U) = V'$ and $\phi(D \setminus U) = D'$,
- (2) $V' \cap (S \cup I(m)) = \emptyset$ and $D' \subseteq I(m)$,

(3) every vertex $v \in V'$ becomes infected at the $(m + 1)$ th step.

Let $Z(X)$ be the set of triples (\mathcal{F}', V', D') that satisfy (1), (2) and (3). Then the variable $X_S(m + 1) - X_S(m)$ is bounded above by $f(\xi_w : w \in V(\mathcal{H}))$, where

$$f(x_w : w \in V(\mathcal{H})) := \sum_{(\mathcal{F}', V', D') \in Z(X)} \left(\prod_{v \in V'} x_v \right). \quad (7.22)$$

As $V' \subseteq D$, f has degree at most $|D|$. Observing that no variable in f has an exponent greater than 1 completes the proof of (i).

By definition, V' is a subset of $V(\mathcal{F}') \setminus (S \cup I(m))$. So if A contains an element of $(S \cup I(m))$, then $\partial_A f = 0$. So suppose that A is a subset of $V(\mathcal{H}) \setminus (S \cup I(m))$, Then for any such A , we have

$$\partial_A f = \sum_{\substack{(\mathcal{F}', V', D') \in Z(X) \\ A \subseteq V'}} \left(\prod_{v \in V' \setminus A} x_v \right). \quad (7.23)$$

We can partition the set $Z(X)$ by the cardinality of each V' . So for $0 \leq k \leq |D|$, let

$$Z^k(X) := \{(\mathcal{F}', V', D') \in Z(X) : |V'| - |A| = k\}.$$

So we can rewrite (7.23) as follows.

$$\partial_A f = \sum_{k=0}^{|D|-|A|} \sum_{\substack{(\mathcal{F}', V', D') \in Z^k(X) \\ A \subseteq V'}} \left(\prod_{v \in V' \setminus A} x_v \right). \quad (7.24)$$

Now recall Definition 7.18 and observe that any $(\mathcal{F}', V', D') \in Z^k(X)$ such that $A \subseteq V'$ is in fact a copy of $X^{U, U'}$ rooted at $S \cup A$ for some $U' \subseteq U \subseteq D$ with $|U| = |V'|$, $|U'| = |A|$ and $|U| - |U'| = k$. So from (7.24) and (7.21), as the variables ξ_w are independent (because $V' \setminus A \subseteq V(\mathcal{H}) \setminus I(m)$ and we are conditioning on \mathcal{F}'_m) we have

$$\mathbb{E}(\partial_A f(\xi_w : w \in V(\mathcal{H})) \mid \mathcal{F}'_m) \leq \sum_{k=0}^{|D|-|A|} \sum_{X' \in X^{k, |A|}} X'_{S \cup A}(m) (\log^{r^4}(d)q)^k,$$

as required for (ii).

Now substituting (7.21) into (7.22) gives

$$f(x_w : w \in V(\mathcal{H})) \leq g(x_e : e \in E(\mathcal{H})) := \sum_{(\mathcal{F}', V', D') \in Z(X)} \left(\prod_{v \in V'} \sum_{e \in Q_v(m)} x_e \right).$$

Notice that the random variables on the right hand side are now associated to hyperedges, not vertices.

Given a fixed $(\mathcal{F}', V', D') \in Z(X)$, let \mathcal{G}' be a subhypergraph of \mathcal{H} containing \mathcal{F}' such that $E(\mathcal{G}') \setminus E(\mathcal{F}') = \{e_v : v \in V'\}$, where $e_v \in Q_v(m)$ for all $v \in V'$. Recall Definition 7.18 again. By definition of $Z(X)$, there exists some non-empty fruitful $U \subseteq D$ such that \mathcal{G}' is a copy of some $X' \in X^U$ rooted at S in $\mathcal{H}(m)$.

As, for e, e' distinct hyperedges in $\mathcal{H}(m)$, we have ξ_e is independent of $\xi_{e'}$ and

$$\mathbb{E}(g(\xi_e : e \in E(\mathcal{H})) \mid \mathcal{F}'_m) \leq \sum_{U \in \mathcal{U}(X)} \sum_{X' \in X^U} X'_S(m) q^{|U|}.$$

By considering the cardinality of U , this can be rewritten as

$$\mathbb{E}(g(\xi_e : e \in E(\mathcal{H})) \mid \mathcal{F}'_m) \leq \sum_{k=1}^{|D|} \sum_{\substack{U \in \mathcal{U}(X) \\ |U|=k}} \sum_{X' \in X^U} X'_S(m) q^k,$$

as required for (iii). □

Now we have developed our tools, we are all set to prove Theorem 7.14. As usual we begin with the proof of the bound on $W_S^i(m)$, as it is the simplest case.

Proposition 7.25.

$$\mathbb{P}(\mathcal{W}) \leq N^{-10\sqrt{\log N}}.$$

Proof. For $0 \leq m \leq M_2 - 1$ and $S \subseteq V(\mathcal{H})$, let $\mathcal{W}(S, m + 1)$ be the set of ω in \mathcal{A}_{M_2} such that $\mathcal{J}(\omega) = m + 1$ and

$$W_S^i(m + 1) - W_S^i(m) > \log^{r^3(r-i)}(d).$$

It follows that

$$\mathcal{W} = \bigcup_{\substack{0 \leq m \leq M_2 - 1 \\ S \subseteq V(\mathcal{H})}} \mathcal{W}(S, m + 1).$$

We will prove that for $0 \leq m \leq M_2 - 1$,

$$\mathbb{P}(\mathcal{W}(S, m + 1) \mid \mathcal{F}'_m) \leq N^{-20\sqrt{\log N}}.$$

The proof of the proposition will then follow by taking the union bound over all $0 \leq m \leq M_2 - 1$ and $S \subseteq V(\mathcal{H})$ of size at most $r - 1$.

Conditioning on \mathcal{F}'_m for $m \geq 0$, the events \mathcal{A}_m and $\mathcal{W}(S, m + 1)$ are disjoint and so we assume that $\omega \notin \mathcal{A}_m$. Fix $S \subseteq V(\mathcal{H})$. We wish to apply Lemma 7.20 along with Corollary 3.6 to obtain the required bound on $\mathbb{P}(\mathcal{W}(S, m + 1))$. As $\omega \notin \mathcal{A}_m$, (L.3) holds at time m and we may apply Lemma 7.20. So let $\Upsilon := W^i$ and let $\tilde{W} := f(\xi_w : w \in V(\mathcal{H}))$,

where $f(x_w : w \in V(\mathcal{H}))$ is the polynomial of degree at most $r - i$ obtained by applying Lemma 7.20 to Υ .

We wish to apply Corollary 3.6 with

$$\tau := 2M_2 \log^{r^3(r-i-1)}(d)$$

and $\mathcal{E}_0 := \log^{2r}(N)$ to obtain an upper bound on \tilde{W} which holds with high probability. In order to apply Corollary 3.6 we must bound $\mathbb{E}_j(\tilde{W} \mid \mathcal{F}'_m)$ for $0 \leq j \leq r - i$.

Claim 7.26. *Let $\omega \notin \mathcal{A}_m$. For $0 \leq a \leq r - i$, we have*

$$\mathbb{E}_a(\tilde{W} \mid \mathcal{F}'_m) \leq \tau.$$

Proof. By Lemma 7.20 (ii), for any set $A \subseteq V(\mathcal{H})$ with $|A| = a$,

$$\mathbb{E}(\partial_A f(\xi_w : w \in V(\mathcal{H})) \mid \mathcal{F}'_m) \leq \sum_{k=0}^{r-i-a} \sum_{\Upsilon' \in \Upsilon^{k,a}} \Upsilon'_{\text{SUA}}(m) (\log^{O(1)}(d)q)^k. \quad (7.27)$$

So let us evaluate this expression. Recalling Definition 7.18 we see that $\Upsilon^{k,a}$ contains only one configuration, the configuration $\bar{\Upsilon}^{k,a} = (\mathcal{F}', R', D')$ where \mathcal{F}' is a set of r vertices contained in a single hyperedge with $i + a$ roots and k neutral vertices (and therefore $r - i - (k + a)$ marked vertices).

We will break the cases up by the value of $a + i$. First suppose $a + i = r$. Here $k = 0$ and $\partial_A \mathbb{E}(\tilde{W}) \leq 1$ since \mathcal{H} contains no hyperedge with multiplicity greater than one.

Now suppose $a \geq 1$ and $a + i < r$, so $2 \leq a + i < r$. When $k \geq 1$, $\bar{\Upsilon}^{k,a}$ is a secondary configuration, and as $\omega \notin \mathcal{A}_m$, by (L.4) we have

$$\bar{\Upsilon}^{k,a}_{\text{SUA}}(m) \leq (m + 1) \log^{O(1)}(d) d^{\frac{k}{r-1}} \log^{-3K/5}(d). \quad (7.28)$$

When $k = 0$, as $\omega \notin \mathcal{A}_m$, by (L.3) we have

$$\bar{\Upsilon}^{0,a}_{\text{SUA}}(m) = W^{i+a}_{\text{SUA}}(m) \leq (m + 1) \log^{r^3(r-i-a)}(d). \quad (7.29)$$

Using (7.28) and (7.29) to evaluate (7.27) gives that, when $\omega \notin \mathcal{A}_m$ and $a \geq 1$, the expectation of $\partial_A f(\xi_v : v \in V(\mathcal{H}))$ conditioned on \mathcal{F}'_m is at most

$$\sum_{k=0}^{r-i-a} \bar{\Upsilon}^{k,a}_{\text{SUA}}(m) (\log^{O(1)}(d)q)^k \leq (m + 1) \log^{r^3(r-i-a)}(d) + \sum_{k=1}^{r-i-a} \log^{O(1)}(d) \cdot \log^{-\frac{3K}{5}}(d).$$

Provided that K is large enough, this is at most τ when $a \geq 1$. This concludes the argument in the case when $a \geq 1$.

It remains to bound $\mathbb{E}_a(\tilde{W} \mid \mathcal{F}'_m)$ when $a = 0$ and $i < r$. We need to be more careful here. By Lemma 7.20 (iii),

$$\mathbb{E}(\tilde{W} \mid \mathcal{F}'_m) \leq \sum_{k=1}^{r-i} \sum_{\substack{U \in \mathcal{U}(\Upsilon) \\ |U|=k}} \sum_{\Upsilon' \in \Upsilon^U} \Upsilon'_S(m) q^k. \quad (7.30)$$

Recall Definition 7.18 and observe that, for each $k \geq 1$, there is precisely one $U \in \mathcal{U}$ with $|U| = k$.

First consider when $i = 1$. We see that when $|U| = k$, we have $\Upsilon^U = Z^{r-1-k,k}$. So as $\omega \notin \mathcal{A}_m$, applying (L.5) to evaluate (7.30) gives that the expectation of \tilde{W} given \mathcal{F}'_m is at most

$$\sum_{k=1}^{r-1} Z_v^{r-1-k,k}(m) q^k \leq \left(1 + \frac{\lambda^2}{r \cdot 3^r}\right) \psi(m) \sum_{k=1}^{r-1} \chi(r-1-k, k) \alpha^k = O(1) \leq \tau.$$

When $1 < i < r$ and $|U| = k$, we see each configuration in Υ^U consists of:

- (i) a secondary configuration $X = (\mathcal{G}, R', D')$, where \mathcal{G} consists of one hyperedge containing i roots and k neutral vertices, and in addition
- (ii) a collection of k unstable hyperedges rooted at neutral vertices of \mathcal{G} that are not in the non-central hyperedge.

So in this case, using (L.3) and (L.4) as $\omega \notin \mathcal{A}_m$ gives

$$\begin{aligned} |\Upsilon^U| &\leq X_S(m) \left((m+1) \log^{r^3(r-1)}(d) \right)^k \\ &\leq (m+1)^{k+1} \log^{2|D|r^r}(d) \cdot d^{\frac{k}{r-1}} \log^{-3K/5}(d) \log^{kr^3(r-1)}(d) \\ &= \log^{O(1)}(d) \cdot \log^{-3K/5}(d) \cdot d^{\frac{k}{r-1}}. \end{aligned}$$

Using this to evaluate (7.30) gives that the expectation of \tilde{W} given \mathcal{F}'_m is at most

$$\sum_{k=1}^{r-i} \log^{O(1)}(d) \cdot \log^{-3K/5}(d) \cdot d^{\frac{k}{r-1}} q^k = o(1),$$

for large enough K . This concludes the proof of the claim. \square

Recall that $\mathcal{E}_0 := \log^{2r}(N)$. As $2r+1 < r^3$, by Claim 7.26 we have

$$\mathbb{E}(\tilde{W} \mid \mathcal{F}'_m) \leq \tau = o\left(\log^{r^3(r-i)}(d)\right).$$

We also have

$$\tau \log^{r-i}(N) \sqrt{\mathcal{E}_0} = o\left(\log^{r^3(r-i)}(d)\right),$$

Thus applying Corollary 3.6 gives that

$$\mathbb{P}(\mathcal{W}(S, m+1) \mid \mathcal{F}'_m) \leq N^{-20\sqrt{\log(N)}},$$

as required. \square

Proposition 7.31.

$$\mathbb{P}(\mathcal{X}) \leq N^{-10\sqrt{\log N}}.$$

Proof. For a secondary configuration $X = (\mathcal{F}, R, D)$, $S \subseteq V(\mathcal{H})$ with $|S| = |R|$ and $0 \leq m \leq M_2 - 1$, let $\mathcal{X}(X, S, m + 1)$ be the set of ω in \mathcal{A}_{M_2} such that $\mathcal{J}(\omega) = m + 1$ and

$$X_S(m + 1) - X_S(m) > \log^{2|D|r^7}(d) \cdot d^{\frac{|V(\mathcal{F})| - |R| - |D|}{r-1}} \log^{-3K/5}(d).$$

It follows that

$$\mathcal{X} = \bigcup_{X, S, m+1} \mathcal{X}(X, S, m + 1).$$

We will prove that for $0 \leq m \leq M_2 - 1$,

$$\mathbb{P}(\mathcal{X}(X, S, m + 1) \mid \mathcal{F}'_m) \leq N^{-20\sqrt{\log N}}.$$

The proof of the proposition will then follow by taking the union bound over all choices of X , S and m .

Conditioning on \mathcal{F}'_m for $m \geq 0$, the events \mathcal{A}_m and $\mathcal{X}(X, S, m + 1)$ are disjoint and so we assume that $\omega \notin \mathcal{A}_m$. Fix $S \subseteq V(\mathcal{H})$ and a secondary configuration $X = (\mathcal{F}, R, D)$. We wish to apply Lemma 7.20 along with Corollary 3.6 to obtain the required bound on $\mathbb{P}(\mathcal{X}(X, S, m + 1))$. As $\omega \notin \mathcal{A}_m$, (L.3) holds at time m and we may apply Lemma 7.20. So let $\tilde{X} := f(\xi_w : w \in V(\mathcal{H}))$, where $f(x_w : w \in V(\mathcal{H}))$ is the polynomial of degree at most $|D|$ obtained by applying Lemma 7.20 to X .

We wish to apply Corollary 3.6 with

$$\tau := \log^{2(|D|-1)r^7+3r^6+1}(d) \cdot d^{\frac{|V(\mathcal{F})| - |R| - |D|}{r-1}} \log^{-3K/5}(d)$$

and $\mathcal{E}_0 := \log^{2|D|+1}(d)$ to obtain an upper bound on \tilde{X} which holds with high probability. We will prove the following claim.

Claim 7.32. *Let $\omega \notin \mathcal{A}_m$. For $0 \leq j \leq |D|$,*

$$\mathbb{E}_j(\tilde{X} \mid \mathcal{F}'_m) = o(\tau).$$

Proof. By Lemma 7.20 (ii), for any set $A \subseteq V(\mathcal{H})$ with $|A| = a$,

$$\mathbb{E}(\partial_A f(\xi_w : w \in V(\mathcal{H})) \mid \mathcal{F}'_m) \leq \sum_{k=0}^{|D|-a} \sum_{X' \in X^{k,a}} X'_{S \cup A}(m) (\log^{r^4}(d)q)^k.$$

So let us evaluate this expression. By Remark 2.25, every $X' \in X^{k,a}$ is a secondary configuration (\mathcal{F}', R', D') with $|D'| \leq |D| - 1$ and $|V(\mathcal{F}')| - |R'| - |D'| = |V(\mathcal{F})| - |R| - |D| + k$.

By definition, $|X^{k,a}| = O(1)$. So as $\omega \notin \mathcal{A}_m$ and $|D| < 3r$, using (L.4) to bound each such $X'_{S \cup A}(m)$ gives that $\mathbb{E}(\partial_A f(\xi_w : w \in V(\mathcal{H})) \mid \mathcal{F}'_m)$ is at most

$$\begin{aligned} & O \left(\sum_{k=0}^{|D|-a} (m+1) \log^{2(|D|-1)r^7}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|+k}{r-1}} \log^{-3K/5}(d) \left(\log^{r^4}(d)q \right)^k \right) \\ & = O \left(\log^{2(|D|-1)r^7+3r^6}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d) \right) \\ & = o(\tau), \end{aligned}$$

as required. \square

As X is secondary, $|D| \leq 3r$ and

$$\tau \log^{|D|}(d) \sqrt{\mathcal{E}_0} = o \left(\log^{2|D|r^7}(d) \cdot d^{\frac{|V(\mathcal{F})|-|R|-|D|}{r-1}} \log^{-3K/5}(d) \right).$$

Using this and the fact that $\mathbb{E}(\tilde{X}) = o(\tau)$ (by Claim 7.32), applying Corollary 3.6 gives that

$$\mathbb{P}(\mathcal{X}(X, S, m+1) \mid \mathcal{F}'_m) \leq N^{-20\sqrt{\log N}},$$

as required. \square

Proposition 7.33.

$$\mathbb{P}(\mathcal{Y}^0) \leq N^{-10\sqrt{\log N}}.$$

Proof. For $v \in V(\mathcal{H})$, $1 \leq i \leq r-2$ and $0 \leq m \leq M_2 - 1$, let $\mathcal{Y}^0(v, i, m+1)$ be the set of ω in \mathcal{A}_{M_2} such that at time $\mathcal{J}(\omega) = m+1$,

$$Y_v^{i,0}(m+1) - Y_v^{i,0}(m) > \psi(m) \left(\frac{\lambda^2}{2\alpha} \right) d^{1-\frac{i}{r-1}}.$$

It follows that

$$\mathcal{Y}^0 = \bigcup_{\substack{0 \leq m \leq M_2-1 \\ v \in V(\mathcal{H}) \\ 0 \leq i \leq r-2}} \mathcal{Y}^0(v, i, m+1).$$

We will prove that for $0 \leq m \leq M_2 - 1$,

$$\mathbb{P}(\mathcal{Y}^0(v, i, m+1) \mid \mathcal{F}'_m) \leq N^{-20\sqrt{\log N}}.$$

The proof of the proposition will then follow by taking the union bound over all choices of v, i and m .

Conditioning on \mathcal{F}'_m for $m \geq 0$, the events \mathcal{A}_m and $\mathcal{Y}^0(v, i, m+1)$ are disjoint and so we assume that $\omega \notin \mathcal{A}_m$. Fix $v \in V(\mathcal{H})$, $1 \leq i \leq r-2$ and set $\Upsilon := Y^i = (\mathcal{F}, R, D)$.

We wish to apply Lemma 7.20 along with Corollary 3.6 to obtain the required bound on $\mathbb{P}(\mathcal{Y}^0(v, i, m + 1))$. As $\omega \notin \mathcal{A}_m$, (L.3) holds at time m and we may apply Lemma 7.20. So let $\tilde{Y} := f(\xi_w : w \in V(\mathcal{H}))$, where $f(x_w : w \in V(\mathcal{H}))$ is the polynomial of degree at most i obtained by applying Lemma 7.20 to Υ .

We wish to apply Corollary 3.6 with

$$\tau := \frac{\psi(m)d^{1-\frac{i}{r-1}}}{\log^{r^2}(N)}$$

and $\mathcal{E}_0 := \log^{r^2+1}(N)$ to obtain an upper bound on \tilde{Y} which holds with high probability. We will prove the following claim.

Claim 7.34. *Let $\omega \notin \mathcal{A}_m$. For $1 \leq a \leq i$,*

$$\mathbb{E}_a(\tilde{Y} \mid \mathcal{F}'_m) = o(\tau).$$

Also,

$$\mathbb{E}(\tilde{Y} \mid \mathcal{F}'_m) \leq \psi(m) \left(\frac{\lambda^2}{4\alpha} \right) d^{1-\frac{i}{r-1}}.$$

Proof. By Lemma 7.20 (ii), for any set $A \subseteq V(\mathcal{H})$ with $|A| = a$,

$$\mathbb{E}(\partial_A f(\xi_w : w \in V(\mathcal{H})) \mid \mathcal{F}'_m) \leq \sum_{k=0}^{i-a} \sum_{\Upsilon' \in \Upsilon^{k,a}} \Upsilon'_{\{v\} \cup A}(m) (\log^{O(1)}(d)q)^k.$$

So let us evaluate this expression. When $|A| \geq 1$, as $i \leq r - 2$ each $\Upsilon' \in \Upsilon^{k,a}$ is a secondary configuration (\mathcal{F}', R', D') with $|V(\mathcal{F}')| = r$, $|R'| = |A| + 1$, $|D'| = i - |A| - k$. So $|V(\mathcal{F}')| - |R'| - |D'| = r - 1 - i + k$. Also, $|\Upsilon^{k,a}| = O(1)$. So as $\omega \notin \mathcal{A}_m$ and $|D| < 3r$, using (L.4) to bound each such $\Upsilon'_{\{v\} \cup A}(m)$ gives that $\mathbb{E}(\partial_A f(\xi_w : w \in V(\mathcal{H})) \mid \mathcal{F}'_m)$ is at most

$$\begin{aligned} & O \left(\sum_{k=0}^{i-a} \left((m+1) \log^{2|D|r^7}(d) \cdot d^{\frac{r-1-i+k}{r-1}} \log^{-3K/5}(d) \right) \left(\log^{r^4}(d)q \right)^k \right) \\ & = O \left(\log^{O(1)}(d) \log^{-3K/5}(d) d^{1-\frac{i}{r-1}} \right) \\ & = o(\tau), \end{aligned}$$

when K is large as $m \leq M_2 = O(\log(d))$. This proves the first statement of the claim.

In the case $A = \emptyset$, we need to bound the expectation of $f(\xi_w : w \in V(\mathcal{H}))$ more carefully. By Lemma 7.20 (iii),

$$\mathbb{E}(\tilde{Y} \mid \mathcal{F}'_m) \leq \sum_{k=1}^i \sum_{\substack{U \in \mathcal{U}(\Upsilon) \\ |U|=k}} \sum_{\Upsilon' \in \Upsilon^U} \Upsilon'_v(m) q^k. \quad (7.35)$$

Recall Definition 7.18 and observe that in this case, for each $1 \leq k \leq i$, there is precisely one $U \in \mathcal{U}$ with $|U| = k$. Here, we see that $\Upsilon^U = Z^{i-k,k}$. So as $\omega \notin \mathcal{A}_m$, applying (L.5) and using (7.6) to bound $\chi(i-k, k)$ (as $i-k \leq r-3$) to evaluate (7.35) gives

$$\begin{aligned} \mathbb{E}(\tilde{Y} \mid \mathcal{F}'_m) &\leq \sum_{k=1}^i Z_v^{i-k,k}(m) q^k \\ &\leq \sum_{k=1}^i \left(1 + \frac{\lambda^2}{r^{3r}}\right) \psi(m) \chi(i-k, k) \alpha^k d^{1-\frac{i}{r-1}} \\ &\leq \sum_{k=1}^i 2\psi(m) \left(\frac{\lambda^2}{r^{3r+1}\alpha^{k+1}}\right) \alpha^k d^{1-\frac{i}{r-1}} \\ &\leq \psi(m) \left(\frac{\lambda^2}{4\alpha}\right) d^{1-\frac{i}{r-1}}. \end{aligned}$$

So

$$\mathbb{E}(\tilde{Y} \mid \mathcal{F}'_m) \leq \psi(m) \left(\frac{\lambda^2}{4\alpha}\right) d^{1-\frac{i}{r-1}}.$$

This proves the second statement of the claim. \square

As

$$\tau \log^i(N) \sqrt{\mathcal{E}_0} = o\left(\psi(m) d^{1-\frac{i}{r-1}}\right),$$

then applying Corollary 3.6 with the bound on $\mathbb{E}(\tilde{Y})$ obtained in Claim 7.34 gives that

$$\mathbb{P}(\mathcal{Y}^0(v, i, m+1) \mid \mathcal{F}'_m) \leq N^{-20\sqrt{\log N}},$$

as required. \square

Proposition 7.36.

$$\mathbb{P}(\mathcal{Y}^{>0}) \leq N^{-10\sqrt{\log N}}.$$

Proof. For $v \in V(\mathcal{H})$, $0 \leq i \leq r-2$, $1 \leq j \leq r-1$ and $0 \leq m \leq M_2-1$, let $\mathcal{Y}^{>0}(v, i, j, m+1)$ be the set of ω in \mathcal{A}_{M_2} such that at time $\mathcal{J}(\omega) = m+1$,

$$Y_v^{i,j}(m+1) > \psi(m+1) \chi(i, j) d^{1-\frac{i}{r-1}}.$$

It follows that

$$\mathcal{Y}^{>0} = \bigcup_{m,v,i,j} \mathcal{Y}^{>0}(v, i, j, m+1).$$

We will prove that for $0 \leq m \leq M_2-1$,

$$\mathbb{P}(\mathcal{Y}^{>0}(v, i, j, m+1) \mid \mathcal{F}'_m) \leq N^{-20\sqrt{\log N}}.$$

The proposition will then follow by taking the union bound over all choices of v, i, j and m .

Conditioning on \mathcal{F}'_m for $m \geq 0$, the events \mathcal{A}_m and $\mathcal{Y}^{>0}(v, i, m+1)$ are disjoint and so we assume that $\omega \notin \mathcal{A}_m$. Fix $v \in V(\mathcal{H})$ $0 \leq i \leq r-2$, $1 \leq j \leq r-1$ and set $\Upsilon := Y^{i,j} = (\mathcal{F}, R, D)$. We wish to apply Lemma 7.20 along with Corollary 3.6 to obtain the required bound on $\mathbb{P}(\mathcal{Y}^{>0}(v, i, j, m+1))$. As $\omega \notin \mathcal{A}_m$, (L.3) holds at time m and we may apply Lemma 7.20. So let $\tilde{Z} := f(\xi_w : w \in V(\mathcal{H}))$, where $f(x_w : w \in V(\mathcal{H}))$ is the polynomial of degree at most $i + (r-1)j$ obtained by applying Lemma 7.20 to Υ .

As $j \geq 1$, $Y^{i,j}$ contains an unstable hyperedge and so every copy of $Y^{i,j}$ in $\mathcal{H}(m)$ is destroyed when we sample every open hyperedge in $Q(m)$. So $Y_v^{i,j}(m+1) \cap Y_v^{i,j}(m) = \emptyset$.

We wish to apply Corollary 3.6 with

$$\tau := \frac{\psi(m)d^{1-\frac{i}{r-1}}}{\log^{r^6}(N)}$$

and $\mathcal{E}_0 := \log^{r^6+1}(N)$ to obtain an upper bound on $Y_v^{i,j}(m)$ which holds with high probability. We will use the following claim.

Claim 7.37. *Let $\omega \notin \mathcal{A}_m$. For $1 \leq a \leq i + (r-1)j$,*

$$\mathbb{E}_a(\tilde{Z} \mid \mathcal{F}'_m) = o(\tau).$$

Also,

$$\mathbb{E}(\tilde{Z} \mid \mathcal{F}'_m) \leq \left(1 - \frac{3\lambda}{2}\right) \psi(m)\chi(i, j)d^{1-\frac{i}{r-1}}.$$

Proof. By Lemma 7.20 (ii), for any set $A \subseteq V(\mathcal{H})$ with $|A| = a$,

$$\mathbb{E}(\partial_A f(\xi_w : w \in V(\mathcal{H})) \mid \mathcal{F}'_m) \leq \sum_{k=0}^{i+(r-1)j-a} \sum_{\Upsilon' \in \Upsilon^{k,a}} \Upsilon'_{\{v\} \cup A}(m) (\log^{r^4}(d)q)^k. \quad (7.38)$$

First let us evaluate this expression when $|A| \geq 1$. For some $0 \leq k \leq i+(r-1)j-a$, consider $\Upsilon' = (\mathcal{F}', R', D') \in \Upsilon^{k,a}$ (as $|A| \geq 1$, we have $|R'| \geq 2$). By definition, \mathcal{F}' is isomorphic to \mathcal{F} (ignoring which vertices are roots and marked). In particular, $\mathcal{F}' = e_0 \cup e_1 \cup \dots \cup e_j$, where e_0 is the central hyperedge, the hyperedges e_1, \dots, e_j are pairwise disjoint and each e_i intersects e in a neutral vertex v_i .

For each $0 \leq \ell \leq j$, define $R_\ell := e_\ell \cap R'$ and define $D_\ell := e_\ell \cap D'$. For $1 \leq \ell \leq j$ define the configuration $Z^\ell := (e_\ell, R_\ell \cup \{v_\ell\}, D_\ell)$. By Definition 7.18, Υ' is obtained from $Y^{i,j}$ by making a fruitful set $\emptyset \neq U \subseteq D$ neutral and turning some subset $U' \subseteq U$ into roots. So as U is fruitful, $|D_\ell| < r-1$ for each ℓ .

For each $1 \leq \ell \leq j$, the configuration Z^ℓ satisfies one of the following.

- (1) $R_\ell = \emptyset$ and $|D_\ell| = s$ for some $s \leq r-2$: in this case $Z^\ell = Y^{s,0}$.
- (2) $R_\ell \neq \emptyset$ and Z^ℓ contains a neutral vertex: in this case Z^ℓ is a secondary configuration.

(3) $R_\ell \neq \emptyset$ and Z^ℓ contains no neutral vertex: in this case $Z^\ell = W^{|R_\ell|+1}$.

Let $S \subseteq V(\mathcal{H})$ be a set of cardinality $|R_\ell| + 1$. If Z_ℓ satisfies (1) then as $\omega \notin \mathcal{A}_m$, using (L.1) gives

$$Z_S^\ell(m) \leq \chi(s, 0) d^{1 - \frac{s}{r-1}} = O\left(d^{\frac{r-1-|R_\ell|-|D_\ell|}{r-1}}\right) \quad (7.39)$$

If Z_ℓ satisfies (2), then using (L.4) (as $\omega \notin \mathcal{A}_m$) and the fact that $m = O(\log d)$ gives

$$Z_S^\ell(m) \leq \log^{O(1)}(d) \cdot d^{\frac{r-1-|R_\ell|-|D_\ell|}{r-1}} \log^{-3K/5}(d). \quad (7.40)$$

If Z_ℓ satisfies (3), then using (L.3) (as $\omega \notin \mathcal{A}_m$) and the fact that $m = O(\log d)$ gives

$$Z_S^\ell(m) \leq \log^{r^4}(d) = O\left(\log^{O(1)}(d) \cdot d^{\frac{r-1-|R_\ell|-|D_\ell|}{r-1}}\right) \quad (7.41)$$

We consider two cases. Firstly $|e_0 \cap R'| = 1$. Then as $|A| \geq 1$, without loss of generality $|e_1 \cap R'| \geq 1$. In this case, the configuration $\hat{Z} := (e_0 \cup e_1, R_0 \cup R_1, D_0 \cup D_1)$ is secondary.

Let \mathcal{S} be the set of all partitions $\mathcal{P} := (S_1, \dots, S_j)$ of $\{v\} \cup A$ such that $|S_1| = |R_0 \cup R_1|$ and for $2 \leq \ell \leq j$, $|S_\ell| = |R_\ell|$. Using this notation, we can bound the number of copies of $\Upsilon' = (\mathcal{F}', R', D')$ $\in \Upsilon^{k,a}$ rooted at $\{v\} \cup A$. We have

$$\Upsilon'_{\{v\} \cup A}(m) \leq \sum_{\mathcal{P} \in \mathcal{S}} \sum_{Z \in \hat{Z}_{S_1}(m)} \sum_{(u_1, \dots, u_j) \subseteq V(Z)} \prod_{\ell=2}^j Z_{S_\ell \cup \{u_\ell\}}^\ell(m) \quad (7.42)$$

where the sum is taken over all distinct ordered (u_1, \dots, u_j) .

For any $S_1 \subseteq V(\mathcal{H})$ with $|S_1| = |R_0 \cup R_1|$, as $\omega \notin \mathcal{A}_m$ we can use (L.4) to give

$$\hat{Z}_{S_1}(m) \leq \log^{O(1)}(d) \cdot d^{\frac{2r-1-|R_0 \cup R_1|-|D_0 \cup D_1|}{r-1}} \log^{-3K/5}(d).$$

As $|\mathcal{S}| = O(1)$, using this with (7.39), (7.40) and (7.41) to evaluate (7.42) gives

$$\Upsilon'_{\{v\} \cup A}(m) = \log^{O(1)}(d) d^{\frac{(j+1)r-j-\sum_{\ell=0}^j(|R_\ell|+|D_\ell|)}{r-1}} \log^{-3K/5}(d).$$

As $\Upsilon' \in \Upsilon^{k,a}$, we have $j(r-1) + i + 1 = |R| + |D| = |R'| + |D'| + k$, therefore $|R'| + |D'| = j(r-1) + i + 1 - k$ and

$$\begin{aligned} (j+1)r - j - \sum_{\ell=0}^j (|R_\ell| + |D_\ell|) &= (j+1)r - j - |R'| - |D'| \\ &= (j+1)r - j - j(r-1) - i - 1 + k \\ &= r - 1 - i + k. \end{aligned}$$

So in this case,

$$\Upsilon'_{\{v\} \cup A}(m) = \log^{O(1)}(d) d^{1 - \frac{i-k}{r-1}} \log^{-3K/5}(d).$$

Now consider $\Upsilon' = (\mathcal{F}', R', D') \in \Upsilon^{k,a}$ such that $|e_0 \cap R'| > 1$. This case follows very similarly to the previous case. As $j > 1$, there is a neutral vertex of e_0 . So the configuration $Z^0 := (e_0, R_0, D_0)$ is secondary.

This time, let \mathcal{S} be the set of all partitions $\mathcal{P} := (S_0, \dots, S_j)$ of $\{v\} \cup A$ such that $|S_0| = |R_0|$ and for $1 \leq \ell \leq j$, $|S_\ell| = |R_\ell|$. As in the previous case, $|\mathcal{S}| = O(1)$. Using this notation, we can bound the number of copies of $\Upsilon' = (\mathcal{F}', R', D')$ in $\Upsilon^{k,a}$ rooted at $\{v\} \cup A$. As $\omega \notin \mathcal{A}_m$, we can use (L.4) to bound $Z_{S_0}^0$ and, for $1 \leq \ell \leq j$, we can use (7.39), (7.40) and (7.41) as in the previous case to bound $Z_{S_\ell \cup \{u_\ell\}}^\ell$. This gives

$$\begin{aligned} \Upsilon'_{\{v\} \cup A}(m) &\leq \sum_{\mathcal{P} \in \mathcal{S}} \sum_{Z \in Z^0(m)} \sum_{(u_1, \dots, u_j) \subseteq V(Z)} \prod_{\ell=1}^j Z_{S_\ell \cup \{u_\ell\}}^\ell(m) \\ &= \log^{O(1)}(d) d^{\frac{(j+1)r-j-\sum_{\ell=0}^j(|R_\ell|+|D_\ell|)}{r-1}} \log^{-3K/5}(d) \\ &= \log^{O(1)}(d) d^{1-\frac{i-k}{r-1}} \log^{-3K/5}(d), \end{aligned}$$

where, as before, the sum is taken over all distinct ordered (u_1, \dots, u_j) .

Now we are able to evaluate (7.38) for $|A| \geq 1$. As $|\Upsilon^{k,a}| = O(1)$,

$$\begin{aligned} \mathbb{E}(\partial_A f(\xi_w : w \in V(\mathcal{H})) \mid \mathcal{F}'_m) &\leq \sum_{k=0}^{i+(r-1)j-a} \sum_{\Upsilon' \in \Upsilon^{k,a}} \Upsilon'_{\{v\} \cup A}(m) (\log^{r^4}(d)q)^k \\ &= \sum_{k=0}^{i+(r-1)j-a} \log^{O(1)}(d) d^{1-\frac{i-k}{r-1}} \log^{-3K/5}(d) (\log^{r^4}(d)q)^k \\ &= \log^{O(1)}(d) d^{1-\frac{i}{r-1}} \log^{-3K/5}(d) \\ &= o(\tau), \end{aligned}$$

when K is sufficiently large, as required for the first statement of the claim.

In the case $A = \emptyset$, we need to bound the expectation of $f(\xi_w : w \in V(\mathcal{H}))$ more carefully. By Lemma 7.20 (iii),

$$\mathbb{E}(\tilde{Z} \mid \mathcal{F}'_m) \leq \sum_{k=1}^{i+(r-1)j} \sum_{\substack{U \in \mathcal{U}(\Upsilon) \\ |U|=k}} \sum_{\Upsilon' \in \Upsilon^U} \Upsilon'_v(m) q^k. \quad (7.43)$$

Recall the definition of Υ^U from Definition 7.18. Consider $\Upsilon' := (\mathcal{F}', D', R') \in \Upsilon^U$ for some $\emptyset \neq U \subseteq D$. Recall that $\mathcal{F}' \supseteq \mathcal{F}$. Let e_0 be the central hyperedge of \mathcal{F} and let e_1, \dots, e_j be the non-central hyperedges of \mathcal{F} .

For $0 \leq \ell \leq j$, define $U_\ell := e_\ell \cap U$. By definition, each vertex u of U is the unique neutral vertex of an unstable hyperedge e_u in \mathcal{F}' . Without loss of generality, suppose that $|U_1| \geq |U_2| \geq \dots \geq |U_j|$. As (by definition) U is fruitful (i.e. it intersects every unstable

hyperedge of Υ') it must intersect every hyperedge e_1, \dots, e_j and hence $|U_\ell| \geq 1$ for all $1 \leq \ell \leq j$.

We will now define some configurations that will help us ‘break up’ Υ' in order to bound the number of copies of it in $\mathcal{H}(m)$. For each $0 \leq \ell \leq j$, define \mathcal{G}_ℓ to be the subhypergraph of \mathcal{F}' containing $e_\ell \cup \{e_u : u \in U_\ell\}$. Define $R_0 := R'$ ($= \{v\}$) and, for $1 \leq \ell \leq j$, define $R_\ell := e_\ell \cap e_0$ (so $|R_\ell| = 1$). For $0 \leq \ell \leq j$, define $D_\ell := \mathcal{G}_\ell \cap D'$ and let Y^ℓ be the configuration $(\mathcal{G}_\ell, R_\ell, D_\ell)$.

First consider Y^0 . We have $|R_0| = 1$ and $|D_0| \geq i - |U_0|$. If $|U_0| \geq 1$, we have $Y^0 = Z^{i-|U_0|, |U_0|}$. If $|U_0| = 0$, we have $Y^0 = Y^{i,0}$. Hence as $\omega \notin \mathcal{A}_m$, by (L.1) and (L.5) we have for any $v \in V(\mathcal{H})$,

$$Y_v^0(m) \leq \begin{cases} \chi(i, 0) d^{1-\frac{i}{r-1}} & \text{if } |U_0| = 0, \\ \left(1 + \frac{\lambda^2}{r^{3r}}\right) \psi(m) \chi(i - |U_0|, |U_0|) d^{1-\frac{i-|U_0|}{r-1}} & \text{otherwise.} \end{cases} \quad (7.44)$$

Note that when $|U_0| = 0$, we can write $\chi(i, 0) d^{1-\frac{i}{r-1}} = \chi(i - |U_0|, |U_0|) d^{1-\frac{i-|U_0|}{r-1}}$.

Now when $1 \leq \ell \leq j$, we have $|U_\ell| \geq 1$ (as U is fruitful), $|R_\ell| = 1$ and $|U_\ell| + |e_\ell \cap D'| = r - 1$. So in particular, for each $1 \leq \ell \leq j$, we have $Y^\ell = Z^{r-1-|U_\ell|, |U_\ell|}$.

For $Z \in Y_v^0(m)$, let $h(Z)$ denote the central hyperedge of Z . Putting this together and applying (L.5) (as $\omega \notin \mathcal{A}_m$) gives

$$\begin{aligned} \Upsilon'_v(m) &\leq \sum_{Z \in Y_v^0(m)} \sum_{(u_1, \dots, u_j) \in h(Z) \setminus (\{v\} \cup I(m))} \prod_{\ell=1}^j Z_{u_\ell}^{r-1-|U_\ell|, |U_\ell|}(m) \\ &\leq Y_v^0(m) \frac{(r-1-i)!}{(r-1-i-j)!} \prod_{\ell=1}^j \left(\left(1 + \frac{\lambda^2}{r^{3r}}\right) \psi(m) d^{\frac{|U_\ell|}{r-1}} \chi(r-1-|U_\ell|, |U_\ell|) \right). \end{aligned}$$

Applying (7.44) to this gives

$$\Upsilon'_v(m) \leq C_{i,j} \cdot \chi(i - |U_0|, |U_0|) d^{1-\frac{i-|U_0|}{r-1}} \prod_{\ell=1}^j \left(d^{\frac{|U_\ell|}{r-1}} \cdot \chi(r-1-|U_\ell|, |U_\ell|) \right), \quad (7.45)$$

where

$$C_{i,j} := \frac{(r-1-i)!}{(r-1-i-j)!} \left(1 + \frac{\lambda^2}{r^{3r}}\right)^j \psi^j(m) \cdot \max \left\{ 1, \left(1 + \frac{\lambda^2}{r^{3r}}\right) \psi(m) \right\}. \quad (7.46)$$

Rewriting the right hand side of (7.43) gives

$$\mathbb{E}(\tilde{Z} \mid \mathcal{F}'_m) \leq \sum_{b=0}^i \sum_{\substack{U \in \mathcal{U}(\Upsilon) \\ |U_0|=b}} \sum_{\Upsilon' \in \Upsilon^U} \Upsilon'_v(m) q^{|U|}.$$

So using (7.45) to bound $\Upsilon'_v(m)$ and using the fact that $|U| = \sum_{\ell=0}^j |U_\ell|$ gives

$$\mathbb{E}(\tilde{Z} \mid \mathcal{F}'_m) \leq C_{i,j} \cdot \sum_{b=0}^i q^b \sum_{\substack{U \in \mathcal{U}(\Upsilon) \\ |U_0|=b}} \chi(i-b, b) d^{1-\frac{i-b}{r-1}} \prod_{\ell=1}^j d^{\frac{|U_\ell|}{r-1}} \cdot \chi(r-1-|U_\ell|, |U_\ell|) q^{|U_\ell|}. \quad (7.47)$$

Now for each distinct ordered partition (k_1, \dots, k_j) of $[k]$ where $k_1 \geq \dots \geq k_j$ and for each b , by the definition of $\mathcal{U}(\Upsilon)$ (in Definition 7.18) and the definition of $Y^{i,j}$ (recall Definition 2.18), there is precisely one $U \in \mathcal{U}(\Upsilon)$ such that $(|U_1|, \dots, |U_\ell|) = (k_1, \dots, k_\ell)$ and $|U_0| = b$. Therefore, from (7.47) we have

$$\begin{aligned} \mathbb{E}(\tilde{Z} \mid \mathcal{F}'_m) &\leq C_{i,j} \cdot \sum_{b=0}^i q^b \chi(i-b, b) d^{1-\frac{i-b}{r-1}} \left(\sum_{\ell=1}^{r-1} q^\ell \chi(r-1-\ell, \ell) d^{\frac{\ell}{r-1}} \right)^j \\ &= C_{i,j} \cdot d^{1-\frac{i}{r-1}} \cdot \sum_{b=0}^i \alpha^b \chi(i-b, b) \left(\sum_{\ell=1}^{r-1} \alpha^\ell \chi(r-1-\ell, \ell) \right)^j. \end{aligned} \quad (7.48)$$

Consider the first summation in (7.48). We have

$$\begin{aligned} \sum_{b=0}^i \chi(i-b, b) \alpha^b &= \chi(i, 0) + \sum_{b=1}^i \chi(i-b, b) \alpha^b \\ &= \chi(i, 0) \left(1 + \sum_{b=1}^i \frac{\chi(i-b, b) \alpha^b}{\chi(i, 0)} \right) \\ &\leq \chi(i, 0) \left(1 + \sum_{b=1}^i \frac{\alpha^b}{\chi(i, 0)} \cdot \frac{\lambda^2 \chi_{\min}}{r^{3r+1}(1+\alpha+\alpha^r)} \cdot \chi(r-2, 1)^{b-1} \right), \end{aligned} \quad (7.49)$$

where in the final line (7.7) was used to bound $\chi(i-b, b)$ (and we ignored the factor of $(1 + \chi_{\max})$). Now using (in this order) the fact that $\alpha^s < 1 + \alpha + \alpha^r$ for $0 \leq s \leq r-1$, the fact that $\chi_{\min} \leq \chi(i, 0)$ (by definition) and that $\chi(r-2, 1) < 1$, we have

$$\begin{aligned} \frac{\alpha^b}{\chi(i, 0)} \cdot \frac{\lambda^2 \chi_{\min}}{r^{3r+1}(1+\alpha+\alpha^r)} \cdot \chi(r-2, 1)^{b-1} &< \frac{\chi_{\min}}{\chi(i, 0)} \cdot \frac{\lambda^2}{r^{3r+1}} \cdot \chi(r-2, 1)^{b-1} \\ &< \frac{\lambda^2}{r^{3r+1}} \cdot \chi(r-2, 1)^{b-1} \\ &\leq \frac{\lambda^2}{r^{3r+1}}. \end{aligned} \quad (7.50)$$

Using (7.49) and (7.50), as $i < r$ we obtain the bound

$$\sum_{b=0}^i \chi(i-b, b) \alpha^b < \chi(i, 0) \left(1 + \frac{\lambda^2}{r^{3r}} \right). \quad (7.51)$$

Now consider the second summation in (7.48). We have

$$\begin{aligned}
 \sum_{\ell=1}^{r-1} \chi(r-1-\ell, \ell) \alpha^\ell &= \chi(r-2, 1) \alpha + \sum_{\ell=2}^{r-1} \chi(r-1-\ell, \ell) \alpha^\ell \\
 &= \chi(r-2, 1) \alpha \left(1 + \sum_{\ell=2}^{r-1} \frac{\chi(r-1-\ell, \ell) \alpha^{\ell-1}}{\chi(r-2, 1)} \right) \\
 &= \chi(r-2, 1) \alpha \left(1 + \sum_{\ell=2}^{r-1} \frac{\alpha^{\ell-1}}{\chi(r-2, 1)} \cdot \frac{\lambda^2 \chi_{\min}}{r^{3r+1}(1+\alpha+\alpha^r)} \cdot \chi(r-2, 1)^{\ell-1} \right),
 \end{aligned}$$

where in the final line (7.7) was used to bound $\chi(r-1-\ell, \ell)$. As in (7.50) we can bound the second term within the bracket by $\frac{\lambda^2}{r^{3r}}$ to obtain

$$\sum_{\ell=1}^{r-1} \chi(r-1-\ell, \ell) \alpha^\ell < \chi(r-2, 1) \alpha \left(1 + \frac{\lambda^2}{r^{3r}} \right). \quad (7.52)$$

Putting together (7.48), (7.51) and (7.52) gives

$$\begin{aligned}
 \mathbb{E}(\tilde{Z} \mid \mathcal{F}'_m) &\leq C_{i,j} \cdot d^{1-\frac{i}{r-1}} \cdot \chi(i, 0) \left(1 + \frac{\lambda^2}{r^{3r}} \right) \left(\chi(r-2, 1) \alpha \left(1 + \frac{\lambda^2}{r^{3r}} \right) \right)^j \\
 &\leq C_{i,j} \left(1 + \frac{\lambda^2}{r^{3r}} \right)^{j+1} \chi(i, 0) \chi(r-2, 1)^j \alpha^j d^{1-\frac{i}{r-1}}
 \end{aligned} \quad (7.53)$$

First consider the case $(i, j) \neq (r-2, 1)$. By Definition 7.3 and (7.53) we have

$$\mathbb{E}(\tilde{Z} \mid \mathcal{F}'_m) \leq \frac{C_{i,j}}{r^{3r}} \left(1 + \frac{\lambda^2}{r^{3r}} \right)^{j+1} \chi(i, j) d^{1-\frac{i}{r-1}}. \quad (7.54)$$

Using the definition of $C_{i,j}$ (in 7.46), we have

$$\begin{aligned}
 \frac{C_{i,j}}{r^{3r}} \left(1 + \frac{\lambda^2}{r^{3r}} \right)^{j+1} &= \frac{1}{r^{3r}} \frac{(r-1-i)!}{(r-1-i-j)!} \left(1 + \frac{\lambda^2}{r^{3r}} \right)^{2j+1} \psi^j(m) \cdot \max \left\{ 1, \left(1 + \frac{\lambda^2}{r^{3r}} \right) \psi(m) \right\} \\
 &\leq \frac{1}{r^{2r}} \left(1 + \frac{\lambda^2}{r^{3r}} \right)^{2j+1} \psi(m) \cdot \max \left\{ 1, \left(1 + \frac{\lambda^2}{r^{3r}} \right) \psi(m) \right\} \\
 &\leq \frac{1}{r^{2r}} \left(1 + \frac{\lambda^2}{r^{3r}} \right)^{2j+2} \psi(m) \\
 &\leq \frac{1}{r^{2r}} (1 + \lambda^2) \psi(m) \\
 &\leq \left(1 - \frac{3\lambda}{2} \right) (1 - \lambda) \psi(m) \\
 &= \left(1 - \frac{3\lambda}{2} \right) \psi(m+1),
 \end{aligned} \quad (7.55)$$

where the final inequality follows from the fact that $\lambda < 1/8$ and $r \geq 3$. Combining (7.54) with (7.55) gives

$$\mathbb{E}(\tilde{Z} \mid \mathcal{F}'_m) \leq \left(1 - \frac{3\lambda}{2} \right) \psi(m+1) \chi(i, j) d^{1-\frac{i}{r-1}},$$

as required. This completes the case when $(i, j) \neq (r-2, 1)$.

In the case $i = r-2$ and $j = 1$, from (7.46) we have

$$C_{r-2,1} \leq \left(1 + \frac{\lambda^2}{r^{3r}}\right)^2 \psi(m). \quad (7.56)$$

So from (7.56), (7.53) and (7.5) we have

$$\begin{aligned} \mathbb{E}(\tilde{Z} \mid \mathcal{F}'_m) &\leq C_{r-2,1} \left(1 + \frac{\lambda^2}{r^{3r}}\right)^2 \chi(r-2, 0)\chi(r-2, 1)\alpha d^{1-\frac{r-2}{r-1}} \\ &\leq C_{r-2,1} \left(1 + \frac{\lambda^2}{r^{3r}}\right)^2 \frac{1-3\lambda}{\alpha} \chi(r-2, 1)\alpha d^{1-\frac{r-2}{r-1}} \\ &\leq \left(1 + \frac{\lambda^2}{r^{3r}}\right)^4 (1-3\lambda)\psi(m)\chi(r-2, 1)d^{1-\frac{r-2}{r-1}} \\ &\leq \left(1 - \frac{3\lambda}{2}\right) (1-\lambda)\psi(m)\chi(r-2, 1)d^{1-\frac{r-2}{r-1}} \\ &\leq \left(1 - \frac{3\lambda}{2}\right) \psi(m+1)\chi(r-2, 1)d^{1-\frac{r-2}{r-1}}, \end{aligned}$$

where the penultimate expression follows from the fact that $\lambda < 1/8$ and $r \geq 3$.

To summarise, when $\omega \notin \mathcal{A}_m$ we have

$$\mathbb{E}(\tilde{Z} \mid \mathcal{F}'_m) \leq \left(1 - \frac{3\lambda}{2}\right) \psi(m+1)\chi(i, j)d^{1-\frac{i}{r-1}},$$

as required. \square

As

$$\tau \log^{i+(r-1)j}(N) \sqrt{\mathcal{E}_0} = o\left(\psi(m)d^{1-\frac{i}{r-1}}\right),$$

then applying Corollary 3.6 with the bound on $\mathbb{E}(\tilde{Z})$ obtained in Claim 7.37 we get that

$$\mathbb{P}(\mathcal{Y}^{>0}(v, i, j, m+1) \mid \mathcal{F}'_m) \leq N^{-20\sqrt{\log N}},$$

as required. \square

This completes the proof of Lemma 7.9 and the proof of Theorem 1.6 in the subcritical case.

5.8 The Second Phase in the Supercritical Case

In this section we prove the ‘supercritical case’ of Theorem 1.6 (i.e. when $c^{r-2}\alpha > \frac{(r-2)^{r-2}}{(r-1)^{r-1}}$).

Recall, as in the previous section, that we have ‘restarted the clock’ and that in this case the first phase runs until time $T := \lfloor \frac{1}{\alpha} \log N \rfloor$. It may be helpful to recall the exact process we run in this phase, given in Subsection 5.2.4.

Theorem 1.6 in this case is implied by the following lemma.

Lemma 8.1. *With probability at least $1 - N^{-7}$: For all $1 \leq m \leq M_2$ the bound*

$$Q_v(m) \geq (\log N)^{\left(\frac{3}{2}\right)^m}$$

holds for all $v \in V(\mathcal{H}) \setminus I(m)$.

We first show how the supercritical case of Theorem 1.6 follows from Lemma 8.1, before proving Lemma 8.1.

Proof of Theorem 1.6 in the supercritical case. By Lemma 8.1 and definition of M_2 (see (2.30)), with probability at least $1 - N^{-7}$ the bound

$$Q_v(M_2) \geq d^{\frac{1}{r-1} + \frac{1}{10}},$$

holds for all $v \in V(\mathcal{H}) \setminus I(M_2)$. We now sample every open hyperedge in $Q(M_2)$. As each hyperedge is independently in \mathcal{H}_q , the probability that some vertex v fails to be infected after this is

$$(1 - q)^{Q_v(M_2)} \leq (1 - q)^{d^{\frac{1}{r-1} + \frac{1}{10}}} \leq e^{-qd^{\frac{1}{r-1} + \frac{1}{10}}} \leq e^{-\alpha d^{\frac{1}{10}}} < N^{-10\sqrt{\log N}}. \quad (8.2)$$

So by the union bound, with probability at least $1 - N^{-5}$, every healthy vertex becomes infected in this round. This completes the proof of Theorem 1.6 in the supercritical case. \square

Recall from Subsection 5.2.4 that in the supercritical case each round consists of two steps. In the first step a subset $Q'(m) \subseteq Q(m)$ is chosen and every hyperedge in $Q'(m)$ is sampled. In the second step, for any healthy vertex contained in ‘enough’ open hyperedges, all hyperedges containing that vertex are sampled.

Definition 8.3. Define $\mathcal{A}_0 := \mathcal{B}_M$ (given in Definition 6.1). For $1 \leq m \leq M_2$, define \mathcal{A}_m to be the event (in Ω' , which was defined in Subsection 5.3.4) that either \mathcal{A}_0 occurs, or there exists some $v \in V(\mathcal{H}) \setminus I(m)$ such that either:

(S.1) $Q_v(m) < (\log N)^{\left(\frac{3}{2}\right)^m}$, or

(S.2) v was in an open hyperedge in some $Q_v^i(\ell)$, for some $\ell \leq m$.

So to prove Lemma 8.1, it suffices to prove that

$$\mathbb{P}(\mathcal{A}_{M_2}) \leq N^{-7}. \quad (8.4)$$

As mentioned above, to prove Lemma 8.1 we use a lower tail version of Janson’s Inequality, Theorem 3.9. In each of our applications of Theorem 3.9, we will simply set $\varepsilon = 1/2$ and use the fact that $\varphi(-1/2) = (1/2) \log(1/2) + (1/2) \geq 1/10$. We are now ready to complete the proof of Lemma 8.1.

Proof of Lemma 8.1. Given a point $\omega \in \mathcal{A}_{M_2}$, let

$$\mathcal{J} = \mathcal{J}(\omega) := \min\{i : \omega \in \mathcal{A}_i\}.$$

For $0 \leq m \leq M_2 - 1$ and $v \in V(\mathcal{H})$, let $\mathcal{S}(v, m+1)$ be the set of $\omega \in \mathcal{A}_{M_2}$ such that $\omega \notin \mathcal{A}_0$, $\mathcal{J}(\omega) = m+1$, $v \notin I(m+1)$ and either:

- (i) $Q_v(m+1) < (\log N)^{\left(\frac{3}{2}\right)^{m+1}}$, or
- (ii) v was in an open hyperedge in some $Q_v^i(m+1)$.

We remark that as $\mathcal{J}(\omega) = m+1$, v cannot have been in an open hyperedge in some $Q_v^i(\ell)$ for $\ell < m+1$.

It follows that

$$\mathcal{A}_{M_2} = \mathcal{A}_0 \cup \bigcup_{\substack{0 \leq m \leq M_2 - 1 \\ v \in V(\mathcal{H})}} \mathcal{S}(v, m+1). \quad (8.5)$$

We will prove that for all $0 \leq m \leq M_2 - 1$ and $v \in V(\mathcal{H})$,

$$\mathbb{P}(\mathcal{S}(v, m+1)) \leq N^{-10}. \quad (8.6)$$

By Theorem 6.2, $\mathbb{P}(\mathcal{A}_0) \leq N^{-2\sqrt{\log N}}$. Lemma 8.1 will follow from (8.6) by taking the union bound over all m and $v \in V(\mathcal{H})$ (as $M_2 = O(\log \log d)$ by (2.2) and (2.30)).

So for $0 \leq m \leq M_2 - 1$, assume $\omega \notin \mathcal{A}_m$. Fix $v \in V(\mathcal{H}) \setminus I(m)$ and consider $\mathcal{S}(v, m+1)$. First let us bound the probability that $v \notin I(m+1)$ but (ii) holds.

Suppose $v \in Q_v^i(m+1)$. By the argument above (culminating in (8.2)), with probability at least $1 - N^{-10\sqrt{\log N}}$ the vertex v becomes infected (and is hence not present in $I(m+1)$) when every hyperedge of $Q_v^i(m+1)$ is sampled. So the probability that $v \notin I(m+1)$ but (ii) holds is at most $N^{-10\sqrt{\log N}}$.

We will now show that when $\omega \notin \mathcal{A}_m$,

$$\mathbb{P}\left(Q_v(m+1) < (\log N)^{\left(\frac{3}{2}\right)^{m+1}} \mid \mathcal{F}'_m\right) \leq N^{-15}. \quad (8.7)$$

Then (8.6) will follow from this and the argument of the previous paragraph.

As $\omega \notin \mathcal{A}_m$, for all $u \in V(\mathcal{H}) \setminus I(m)$ we have

$$Q_u(m) \leq d^{\frac{1}{r-1} + \frac{1}{10}}. \quad (8.8)$$

(By definition of the process, if $Q_u(m) > d^{\frac{1}{r-1} + \frac{1}{10}}$ then we would have sampled these open hyperedges.) Recall from Subsection 5.2.4 that $Q'(0) := Q(0)$ and for $m > 0$, for each

$v \in V(\mathcal{H}) \setminus I(m)$ we choose $Q'_v(m)$ to be a subset of $Q_v(m)$ with cardinality $(\log N)^{\left(\frac{3}{2}\right)^m}$ (which is possible here by definition of \mathcal{A}_m) and set

$$Q'(m) := \bigcup_{v \in V(\mathcal{H}) \setminus I(m)} Q'_v(m).$$

Fix $v \in V(\mathcal{H}) \setminus I(m)$. Define $s := s(m)$, to be the integer in $[0, r-2]$ such that the cardinality of $Y_v^s(m) := Y_v^{s,0}(m) - Y_v^{s+1,0}(m)$ is maximised ($Y_v^s(m)$ is the number of hyperedges of $\mathcal{H}(m)$ containing v and exactly s infected vertices) As $\omega \notin \mathcal{A}_0$, by (B.1) from Definition 6.1 we have $s(0) = 0$.

Define $\mathcal{G} := \mathcal{G}(m)$ to be the $(r-1-s)$ -uniform hypergraph on vertex set $Q'(m)$, where $S := \{e_1, \dots, e_{r-1-s}\} \in E(\mathcal{G})$ if and only if there exists some $\mathcal{F} \in Y_v^s(m)$ such that $V(\mathcal{F}) \setminus (\{v\} \cup I(m)) = V(S) \setminus I(m)$. In other words, for each healthy vertex x of $\mathcal{F} \setminus \{v\}$ there is a hyperedge in S containing x . \mathcal{F} will become a member of $Q_v(m)$ if each of e_1, \dots, e_{r-1-s} is successfully sampled from $Q'(m)$.

Recall the definition of $Q_v^0(m+1)$ from the description of the second phase process in the supercritical case, given in Subsection 5.2.4. If ξ_e is the Bernoulli random variable which is equal to 1 if and only if $e \subseteq V(\mathcal{G})_q$, we have $Q_v^0(m+1) \geq X$, where

$$X := \sum_{e \in E(\mathcal{G}(m))} \xi_e.$$

We will apply Theorem 3.9 to prove a bound on X and hence to prove (8.6).

To bound $\mathbb{E}(X)$, we must first bound $|E(\mathcal{G})|$.

Claim 8.9. *If $\omega \notin \mathcal{A}_0$:*

$$\frac{d}{4} \cdot (\log N)^{(r-1)^2} \leq |E(\mathcal{G}(0))| \leq 4d \cdot (\log N)^{(r-1)^2}$$

and, for $m \geq 1$, if $\omega \notin \mathcal{A}_m$:

$$\frac{d}{2(r-1)} (\log N)^{(r-1-s)\left(\frac{3}{2}\right)^m} \leq |E(\mathcal{G}(m))| \leq d \cdot (\log N)^{(r-1-s)\left(\frac{3}{2}\right)^m}.$$

Proof. First consider $E(\mathcal{G}(0))$. As mentioned above, as $\omega \notin \mathcal{A}_0$ we have that $s(0) = 0$ and so $\mathcal{G}(0)$ is a $(r-1)$ -uniform hypergraph on vertex set $Q'(0) = Q(0)$. For an open hyperedge e , let $h(e)$ be the healthy vertex of e . By definition, the set $\{e_1, \dots, e_{r-1}\}$ is a hyperedge of $\mathcal{G}(0)$ if and only if there exists a hyperedge $\{v, h(e_1), \dots, h(e_{r-1})\} \in E(\mathcal{H}(m))$.

Say that a member of $Z_v^{0,r-1}(m)$ is *centrally healthy* if its central hyperedge contains no vertex of $I(m)$ and let $\tilde{Z}_v^{0,r-1}(0)$ be the set of centrally healthy members of $Z_v^{0,r-1}(0)$. Consider the map ϕ from $E(\mathcal{G}(0))$ to $\tilde{Z}_v^{0,r-1}(0)$ which maps $\{e_1, \dots, e_{r-1}\} \in E(\mathcal{G}(0))$ to the unique member of $\tilde{Z}_v^{0,r-1}(0)$ with central hyperedge $\{v, h(e_1), \dots, h(e_{r-1})\}$ and non central hyperedges e_1, \dots, e_{r-1} . By definition of $\mathcal{G}(0)$, we have that ϕ is a bijection.

So to give a lower bound on $|E(\mathcal{G}(0))|$, it suffices to prove a lower bound on the number of centrally healthy members of $Z_v^{0,r-1}(m)$. By definition of $Z^{0,r-1}$, a lower bound is given by the number of centrally healthy members of $Y_v^{0,r-1}(0)$. We will bound this number from below.

First let us bound the number of Y in $Y_v^{0,r-1}(0)$ that are not centrally healthy but that do contain a healthy vertex x other than v in their central hyperedge. Each such Y can be thought of as the union of some Y' in $Y_v^{1,r-2}(0)$ and an additional open hyperedge rooted at x . So in particular, as $\omega \notin \mathcal{A}_0$, using (B.1) and (B.3) from Definition 6.1 we can bound the number of such Y above by $\log^{O(1)}(d) \cdot d^{\frac{r-2}{r-1}}$.

Now consider a Y in $Y_v^{0,r-1}(0)$ that contains no healthy vertex other than v . The central hyperedge e^* of Y is in $Q_v(0)$ and every non-central hyperedge is a copy of W^1 , rooted at a vertex of e^* (whose root is actually in $I(0)$). So, as $\omega \notin \mathcal{A}_0$, again using (B.1) and (B.3) from Definition 6.1 we can bound the number of such Y above by $\log^{O(1)}(d)$.

So by the previous two paragraphs, the number of members of $Y_v^{0,r-1}(0)$ that are not centrally healthy is $\log^{O(1)}(d) \cdot d^{\frac{r-2}{r-1}}$. As $\omega \notin \mathcal{A}_0$, using (B.1) from Definition 6.1 we have

$$Y_v^{0,r-1}(0) \in (1 \pm o(1))d((c + \alpha T)^{r-1} - T)^{r-1}. \tag{8.10}$$

So

$$\frac{1}{2}d((c + \alpha T)^{r-1} - T)^{r-1} \leq |E(\mathcal{G}(0))|. \tag{8.11}$$

Now let us think about an upper bound for $|E(\mathcal{G}(0))|$. By the above discussion we have $|E(\mathcal{G}(0))| \leq Z_v^{0,r-1}(0)$. We bound $Z_v^{0,r-1}(0)$ analogously to in the proof of Lemma 7.9. The variable $Z_v^{0,r-1}(m)$ counts the number of ways to choose

- (a) a hyperedge $e = \{v, w_1, \dots, w_{r-1}\} \in \mathcal{H}(m)$ containing v , and
- (b) hyperedges e_1, \dots, e_{r-1} such that $e_\ell \in W_{w_\ell}^1(m)$ for $1 \leq \ell \leq r-1$.

The number of choices in which each of the hyperedges e_1, \dots, e_{r-1} intersects e on only one vertex and no pair of them intersect one another is precisely $Y_v^{0,r-1}(m)$. If one of the hyperedges e_1, \dots, e_{r-1} intersects e on more than one vertex or two of them intersect one another, then the union of e and e_1, \dots, e_{r-1} consists of a copy \mathcal{F}' of a secondary configuration with $r-1$ neutral vertices and a set of copies of W^1 rooted at vertices of \mathcal{F}' . As $\omega \notin \mathcal{A}_0$, by (B.2) and (B.3) from Definition 6.1 the number of choices in this case is at most

$$\log^{O(1)}(d) \cdot d \cdot \log^{-3K/5}(d),$$

This is $o(d)$, provided that K is sufficiently large.

As $\omega \notin \mathcal{A}_0$, by (B.1) from Definition 6.1 and the previous argument we have

$$Z_v^{0,r-1}(0) \leq (1 + o(1))Y_v^{0,r-1}(0).$$

So $|E(\mathcal{G}(0))| \leq 2 \cdot Y_v^{0,r-1}(0)$ and combining this with (8.10) and (8.11) gives

$$\frac{1}{2}d((c + \alpha T)^{r-1} - T)^{r-1} \leq |E(\mathcal{G}(0))| \leq 2d((c + \alpha T)^{r-1} - T)^{r-1}. \quad (8.12)$$

By definition of T (see 2.11), for N sufficiently large we have

$$\frac{1}{2}(\log(N))^{(r-1)^2} \leq ((c + \alpha T)^{r-1} - T)^{r-1} \leq 2(\log(N))^{(r-1)^2}. \quad (8.13)$$

Combining (8.12) and (8.13) gives

$$\frac{d}{4} \cdot (\log N)^{(r-1)^2} \leq |E(\mathcal{G}(0))| \leq 4d \cdot (\log N)^{(r-1)^2},$$

as required for the first part of the claim.

Now consider $E(\mathcal{G}(m))$ for $m \geq 1$. First let us bound $Y_v^s(m)$ for any $v \in V(\mathcal{H}(m)) \setminus I(m)$. As $\omega \notin \mathcal{A}_m$, v was never in an open hyperedge sampled in the second step of any round. So by (8.8), definition of the process and definition of M_2 (see 2.30):

- (1) at most $m \cdot d^{\frac{1}{r-1} + \frac{1}{10}} = o(d)$ open hyperedges containing v have been sampled during the second phase up until now,
- (2) $Q_v(m) \leq d^{\frac{1}{r-1} + \frac{1}{10}}$.

As $\omega \notin \mathcal{A}_m$, by (B.1) from Definition 6.1 we have $Y_v^0(0) \geq (1 - o(1))d$. So putting this together with (1) and (2) gives that there are at least $(1 - o(1))d \geq d/2$ hyperedges containing v in $\mathcal{H}(m) \setminus Q_v(m)$.

So by the pigeonhole principle,

$$\frac{d}{2(r-1)} \leq Y_v^s(m) \leq d.$$

By definition of $\mathcal{G}(m)$, we have

$$|E(\mathcal{G}(m))| = \sum_{Y \in Y_v^s(m)} \prod_{u \in V(Y) \setminus (\{v\} \cup I(m))} Q'_u(m).$$

Using the definition of $Q'_u(m)$ with the previous two expressions gives

$$\frac{d}{2(r-1)} (\log N)^{(r-1-s)\left(\frac{3}{2}\right)^m} \leq |E(\mathcal{G}(m))| \leq d \cdot (\log N)^{(r-1-s)\left(\frac{3}{2}\right)^m},$$

completing the proof of the claim. □

As $\mathbb{E}(\xi_e) = q^{r-1-s}$, Claim 8.9 implies that when $m = 0$ we have

$$\frac{\alpha^{r-1}}{4} \cdot (\log N)^{(r-1)^2} \leq \mathbb{E}(X | \mathcal{F}'_m) \leq 2\alpha^{r-1} \cdot (\log N)^{(r-1)^2}, \quad (8.14)$$

and when $m > 0$, we have

$$\frac{\alpha^{r-1-s}}{2(r-1)} d^{\frac{s}{r-1}} \cdot (\log N)^{(r-1-s)(\frac{3}{2})^m} \leq \mathbb{E}(X | \mathcal{F}'_m) \leq \alpha^{r-1-s} d^{\frac{s}{r-1}} \cdot (\log N)^{(r-1-s)(\frac{3}{2})^m}. \quad (8.15)$$

So for any $m \geq 0$, as $0 \leq s \leq r-2$ the lower bounds of (8.14) and (8.15) give

$$\mathbb{E}(X | \mathcal{F}'_m) \geq 2 \cdot (\log N)^{(\frac{3}{2})^m} \quad (8.16)$$

We require one more claim before we can apply Theorem 3.9.

Claim 8.17. *For $0 \leq m \leq M_2 - 1$, when $\omega \notin \mathcal{A}_m$,*

$$\sum_{\substack{e, e' \in E(\mathcal{G}(m)) \\ e \cap e' \neq \emptyset}} \mathbb{E}(\xi_e \xi_{e'} | \mathcal{F}'_m) = O\left(\mathbb{E}(X | \mathcal{F}'_m)^2 \log^{-\frac{3}{2}} N\right).$$

Proof. We have

$$\sum_{\substack{e, e' \in E(\mathcal{G}(m)) \\ e \cap e' \neq \emptyset}} \mathbb{E}(\xi_e \xi_{e'} | \mathcal{F}'_m) = \sum_{e \in E(\mathcal{G}(m))} \sum_{j=1}^{r-2} \sum_{\substack{e' \in E(\mathcal{G}(m)) \\ |e \cap e'|=j}} \mathbb{E}(\xi_e \xi_{e'} | \mathcal{F}'_m).$$

For fixed $e \in E(\mathcal{G}(m))$, recall that $e = \{e_1, \dots, e_{r-s-1}\}$ where each $e_i \in Q'(m)$. The quantity $|\{e' \in E(\mathcal{G}) : |e \cap e'| = j\}|$ is bounded above by the number of copies Z of $Z^{s, r-1-s}$ rooted at v in $\mathcal{H}(m)$ such that Z contains some $\{e'_1, \dots, e'_j\} \subseteq \{e_1, \dots, e_{r-s-1}\}$. There are $O(\Delta_{j+1})$ ways to choose the hyperedge f of Z containing v . Given the choice of f , for each u that is not contained in a hyperedge of $\{e'_1, \dots, e'_j\}$ there are precisely $Q'_u(m)$ ways to choose the open hyperedge rooted at u .

So when $m = 0$, as $\omega \notin \mathcal{A}_0$, by (B.3) from Definition 6.1 for each $v \in V(\mathcal{H}) \setminus I(0)$ we have $Q_v(0) \leq \log^{r^4}(d)$ and so

$$\begin{aligned} \sum_{\substack{e, e' \in E(\mathcal{G}(0)) \\ e \cap e' \neq \emptyset}} \mathbb{E}(\xi_e \xi_{e'} | \mathcal{F}'_m) &\leq \frac{1}{2} \sum_{e \in E(\mathcal{G}(0))} \sum_{j=1}^{r-2} \Delta_{j+1} (\log N)^{r^4(r-1-j)} q^{2(r-1)-j} \\ &= O\left((\log N)^{r^4(r-2)+(r-1)^2-K}\right), \end{aligned}$$

where in the last line we used the upper bound on $|E(\mathcal{G}(0))|$ given by Claim 8.9. Using the upper bound of (8.14), we see the required bound for $m = 0$ follows when K is taken to be large with respect to r .

When $m \geq 1$, again using Claim 8.9 we have

$$\begin{aligned} \sum_{\substack{e, e' \in E(\mathcal{G}(m)) \\ e \cap e' \neq \emptyset}} \mathbb{E}(\xi_e \xi_{e'} \mid \mathcal{F}'_m) &\leq \sum_{e \in E(\mathcal{G}(m))} \sum_{j=1}^{r-2} \Delta_{j+1} \left((\log N) \left(\frac{3}{2} \right)^m \right)^{(r-1-s-j)} q^{2(r-1-s)-j} \\ &= O \left(d^{\frac{2s}{r-1}} (\log N)^{(2(r-1-s)-1) \left(\frac{3}{2} \right)^m - K} \right). \end{aligned}$$

Using the upper bound in (8.15), we see that the required bound holds for $m > 0$. \square

So we can apply Theorem 3.9 with $\epsilon = 1/2$ to give

$$\mathbb{P} \left(X \leq \frac{1}{2} \mathbb{E}(X) \mid \mathcal{F}'_m \right) \leq e^{-\Omega \left(\log^{\frac{3}{2}} N \right)}.$$

So using the lower bounds given in (8.14) and (8.15) we have

$$\mathbb{P} \left(Q_v^0(m+1) < (\log N) \left(\frac{3}{2} \right)^{m+1} \mid \mathcal{F}'_m \right) \leq N^{-20}.$$

If v is contained in an open hyperedge sampled in the second step of this round, it becomes infected with probability at least $1 - N^{-10\sqrt{\log n}}$ (by the argument culminating in (8.2)). As $Q_v^0(m) = Q_v(m)$ for any healthy vertex v not contained in an open hyperedge sampled in the second step of this round, this completes the proof of (8.7) and hence the proof of Lemma 8.1. \square

This concludes our discussion of the second phase and therefore the proof of Theorem 1.6.

5.9 Strictly k -Balanced Hypergraphs

In this final section, using Theorem 1.4 we prove Theorem 1.9 as well as a generalisation of it to strictly k -balanced k -uniform hypergraphs, which we define now. The following two definitions are a generalisation of Definitions 1.7 and 1.8.

Definition 9.1. Given a k -uniform hypergraph \mathcal{F} with at least two hyperedges, the *k -density* of \mathcal{F} is defined by

$$d_k(\mathcal{F}) := \frac{|E(\mathcal{F})| - 1}{|V(\mathcal{F})| - k}.$$

Definition 9.2. Say that a k -uniform hypergraph \mathcal{F} with at least two hyperedges is *k -balanced* if $d_k(\mathcal{F}') \leq d_k(\mathcal{F})$ for every proper subhypergraph \mathcal{F}' of \mathcal{F} with at least two hyperedges. Say that \mathcal{F} is *strictly k -balanced* if this inequality is strict for all such subhypergraphs.

For hypergraphs \mathcal{G} and \mathcal{F} , define $\mathcal{H}_{\mathcal{G},\mathcal{F}}$ to be the hypergraph whose vertices are the hyperedges of \mathcal{G} and whose hyperedges are the sets of hyperedges which form copies of \mathcal{F} in \mathcal{G} . Let $K_n^{(k)}$ denote the complete k -uniform hypergraph on n vertices. For any k -uniform hypergraph \mathcal{F} , we have that $\mathcal{H}_{K_n^{(k)},\mathcal{F}}$ is $|E(\mathcal{F})|$ -uniform and $d(n,\mathcal{F})$ -regular for some integer $d(n,\mathcal{F})$ such that $d(n,\mathcal{F}) = \Theta(n^{|V(\mathcal{F})|-k})$ (the constant factor is related to the number of automorphisms of \mathcal{F} which fix a hyperedge).

Now we state a generalisation of Theorem 1.9 to strictly k -balanced k -uniform hypergraphs.

Theorem 9.3. *Let \mathcal{F} be a strictly k -balanced k -uniform hypergraph with at least three hyperedges satisfying $\delta(\mathcal{F}) \geq 2$. Define $r := |E(\mathcal{F})|$, $\mathcal{H} := \mathcal{H}_{K_n^{(k)},\mathcal{F}}$ and $d := d(n,\mathcal{F})$. If $q := \alpha d^{-1/(r-1)}$ for some fixed $\alpha > 0$, then*

$$p_c(\mathcal{H}_q) = \left(\frac{r-2}{\alpha^{1/(r-2)}(r-1)^{(r-1)/(r-2)} + o(1)} \right) \cdot d^{-1/(r-1)}.$$

Below, we show that Theorem 1.9 follows from Theorem 1.4 and the case $k = 2$ of the following proposition. Theorem 9.3 follows from the same proof.

Proposition 9.4. *For $k \geq 2$, let \mathcal{F} be a k -uniform hypergraph with at least 3 hyperedges and let $n \geq |V(\mathcal{F})|$. Define $r := |E(\mathcal{F})|$, $\mathcal{H} := \mathcal{H}_{K_n^{(k)},\mathcal{F}}$ and $d := d(n,\mathcal{F})$.*

- (i) *If \mathcal{F} is strictly k -balanced and $\delta(\mathcal{F}) \geq 2$, then \mathcal{H} is (d, ρ, ν) -well behaved where $\rho = O\left(d^{-\frac{1}{(|E(\mathcal{F})|-1)(|V(\mathcal{F})|-k)}}\right)$ and $\nu = O\left(d^{\frac{k}{|V(\mathcal{F})|-k}}\right)$.*
- (ii) *If \mathcal{F} is not strictly k -balanced, then for every hyperedge of \mathcal{H} , there exists some $2 \leq \ell \leq r-1$ and a set S of ℓ vertices, such that $\deg(S) = \Omega\left(d^{1-\frac{\ell-1}{|E(\mathcal{F})|-1}}\right)$.*
- (iii) *If $\delta(\mathcal{F}) = 1$, then for every $e_1 \in V(\mathcal{H})$ there exists some $e_2 \in V(\mathcal{H})$ such that $|N_{\mathcal{H}}(e_1) \cap N_{\mathcal{H}}(e_2)| = \Omega(d)$.*

The purpose of parts (ii) and (iii) of Proposition 9.4 is to show that both hypotheses in part (i) of Proposition 9.4 are necessary (in a strong sense). That is, in order for $\mathcal{H}_{K_n^{(k)},\mathcal{F}}$ to be $(d(n,\mathcal{F}), \rho, \nu)$ -well behaved where ρ and ν are functions of n such that ρ tends to zero as $n \rightarrow \infty$, one requires both that \mathcal{F} is strictly k -balanced and that $\delta(\mathcal{F}) \geq 2$.

As it turns out, every strictly 2-balanced graph F with at least three edges satisfies $\delta(F) \geq 2$ (see the proof of Theorem 1.9 below). Thus, in this case, the assumption that $\delta(F) \geq 2$ in part (i) of Proposition 9.4 is redundant. However, for k -uniform hypergraphs with $k \geq 3$, this is no longer the case. For example, a ‘loose k -uniform cycle’ is strictly k -balanced and contains vertices of degree one; see Figure 5.14 for an example.

We first show that Proposition 9.4 implies Theorem 1.9 before proving the proposition itself.

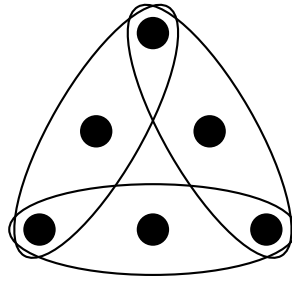


Figure 5.14: A loose 3-uniform triangle. The black circles represent vertices and the ellipses represent hyperedges.

Proof of Theorem 1.9. First, we show that $d_2(F) \geq 1$. If not, then F contains at most $|V(F)| - 2$ edges, and so F is disconnected. If F contains a connected component F' with at least two edges, then $d_2(F') \geq 1 > d_2(F)$, contradicting the fact that F is strictly 2-balanced. Otherwise, every component of F contains at most one edge, which implies that $d_2(F) \leq 1/2$. In this case, we let F' be a subgraph of F consisting of two edges and four vertices. Since F has at least three edges, we have that F' is a proper subgraph of F . Also, $d_2(F') = 1/2 \geq d_2(F)$, which contradicts the fact that F is strictly 2-balanced.

By the previous paragraph, we have $d_2(F) \geq 1$. In particular, this implies that F cannot contain a vertex v of degree one; otherwise, let $F' := F \setminus \{v\}$ and observe that F' has at least two edges and that $d_2(F') \geq d_2(F)$. Thus, $\delta(F) \geq 2$ and so we can apply part (i) of Proposition 9.4 to get that there exists $\beta, s > 0$ such that $\mathcal{H}_{K_n^{(k)}, \mathcal{F}}$ is $(d(n, F), d(n, F)^{-s}, d(n, F)^\beta)$ -well behaved for n sufficiently large. The result now follows by applying Theorem 1.4 to $\mathcal{H}_{K_n^{(k)}, \mathcal{F}}$. \square

We now present the proof of Proposition 9.4.

Proof of Proposition 9.4. First suppose that \mathcal{F} is strictly k -balanced and that $\delta(\mathcal{F}) \geq 2$. As \mathcal{H} is d -regular, conditions (a) and (b) of Definition 1.5 hold for \mathcal{H} . Also, $|V(\mathcal{H})| = \binom{n}{k}$ and so, as $d = \Theta(n^{|V(\mathcal{F})|-k})$, we have

$$|V(\mathcal{H})| = O(n^k) = O\left(n^{\binom{|V(\mathcal{F})|-k}{|V(\mathcal{F})|-k}}\right) = O\left(d^{\frac{k}{|V(\mathcal{F})|-k}}\right).$$

Therefore, condition (e) of Definition 1.5 holds.

Next, we show that condition (c) of Definition 1.5 is satisfied. For $2 \leq \ell \leq r - 1$, let S be a set of ℓ vertices of \mathcal{H} and let \mathcal{F}' be the subhypergraph of $K_n^{(k)}$ induced by the hyperedges corresponding to elements of S . If \mathcal{F}' is not isomorphic to a subhypergraph of \mathcal{F} , then $\deg(S) = 0$ and we are done. Otherwise, we have

$$\deg(S) = \Theta\left(n^{|V(\mathcal{F})|-|V(\mathcal{F}')|}\right).$$

Now, since \mathcal{F} is strictly k -balanced,

$$\begin{aligned} \frac{|E(\mathcal{F}')| - 1}{|V(\mathcal{F}')| - k} &< \frac{|E(\mathcal{F})| - 1}{|V(\mathcal{F})| - k} \Rightarrow \frac{(|E(\mathcal{F}')| - 1)(|V(\mathcal{F})| - k)}{|E(\mathcal{F})| - 1} < |V(\mathcal{F}')| - k \\ \Rightarrow |V(\mathcal{F})| - |V(\mathcal{F}')| &< (|V(\mathcal{F})| - k) \left(1 - \frac{|E(\mathcal{F}')| - 1}{|E(\mathcal{F})| - 1}\right) = (|V(\mathcal{F})| - k) \left(1 - \frac{\ell - 1}{|E(\mathcal{F})| - 1}\right). \end{aligned}$$

Since the above inequality is strict, we get

$$\begin{aligned} \deg(S) &= \Theta \left(n^{|V(\mathcal{F})| - |V(\mathcal{F}')|} \right) = O \left(n^{(|V(\mathcal{F})| - k) \left(1 - \frac{\ell - 1}{|E(\mathcal{F})| - 1}\right)} n^{-\frac{1}{|E(\mathcal{F})| - 1}} \right) \\ &= O \left(d^{1 - \frac{\ell - 1}{|E(\mathcal{F})| - 1}} n^{-\frac{1}{|E(\mathcal{F})| - 1}} \right) \\ &= O \left(d^{1 - \frac{\ell - 1}{|E(\mathcal{F})| - 1}} d^{-\frac{1}{(|E(\mathcal{F})| - 1)(|V(\mathcal{F})| - k)}} \right) \end{aligned}$$

which implies condition (c) of Definition 1.5 is satisfied.

Next we show that if $\delta(\mathcal{F}) \geq 2$, then condition (d) of Definition 1.5 holds. Let e_1 and e_2 be two distinct hyperedges of $K_n^{(k)}$ and let v_1 be a vertex of e_1 which is not contained in e_2 . Suppose that \mathcal{F}' is a copy of \mathcal{F} in $K_n^{(k)}$ containing e_2 such that $(\mathcal{F}' \setminus \{e_2\}) \cup \{e_1\}$ is also a copy of \mathcal{F} . Then, as $\delta(\mathcal{F}) \geq 2$, we have that \mathcal{F}' also contains v_1 . However, the number of copies of \mathcal{F} in $K_n^{(k)}$ containing e_2 and v_1 is

$$O \left(n^{|V(\mathcal{F})| - k - 1} \right) = O \left(d^{1 - \frac{1}{|V(\mathcal{F})| - k}} \right) = o \left(d^{1 - \frac{1}{(|E(\mathcal{F})| - 1)(|V(\mathcal{F})| - k)}} \right)$$

since \mathcal{F} has at least two hyperedges. Therefore, \mathcal{H} satisfies condition (d) of Definition 1.5. This completes the proof of part (i).

Now suppose that \mathcal{F} is not strictly k -balanced and let \mathcal{F}' be a subgraph of \mathcal{F} with at least 2 hyperedges such that $d_k(\mathcal{F}') \geq d_k(\mathcal{F})$ and define $\ell = |E(\mathcal{F}')|$. Let e be a hyperedge of \mathcal{H} (by definition the corresponding hyperedges of $K_n^{(k)}$ induce a copy of \mathcal{F}'). So we can pick S be a set of ℓ vertices of e such that the corresponding hyperedges of $K_n^{(k)}$ induce a copy of \mathcal{F}' . Once again, we get that $\deg(S) = \Theta \left(n^{|V(\mathcal{F})| - |V(\mathcal{F}')|} \right)$. This time, though,

$$|V(\mathcal{F})| - |V(\mathcal{F}')| \geq (|V(\mathcal{F})| - k) \left(1 - \frac{\ell - 1}{|E(\mathcal{F})| - 1}\right).$$

Therefore, $\deg(S) = \Omega \left(d^{1 - \frac{\ell - 1}{|E(\mathcal{F})| - 1}} \right)$. This completes the proof of part (ii).

To prove (iii), suppose that \mathcal{F} contains a vertex of degree one and let e be a hyperedge of \mathcal{F} containing such a vertex. Let t denote the number of vertices of degree one contained in e . Fix $e_1 \in V(\mathcal{H})$ and let e_2 be any hyperedge of $K_n^{(k)}$ which intersects e_1 on $k - t$ vertices. The number of copies of \mathcal{F} in $K_n^{(k)}$ containing e_1 but not containing any vertex of $e_2 \setminus e_1$ is at least a constant multiple of $\binom{n - |e_2 \setminus e_1|}{|V(\mathcal{F})| - k} = \Omega \left(n^{|V(\mathcal{F})| - k} \right)$. For any such copy, say \mathcal{F}' , we have that $\mathcal{F}' \setminus \{e_1\} \cup \{e_2\}$ is also a copy of \mathcal{F} . So

$$|N_{\mathcal{H}}(e_1) \cap N_{\mathcal{H}}(e_2)| = \Omega \left(n^{|V(\mathcal{F})| - k} \right) = \Omega(d),$$

as $d = \Theta(n^{|V(\mathcal{F})|-k})$. This completes the proof of part (iii), and of the proposition. \square

As we have proved, we can apply Theorem 1.4 to $\mathcal{H}_{K_n^{(k)}, \mathcal{F}}$ if and only if \mathcal{F} is strictly k -balanced and $\delta(\mathcal{F}) \geq 2$. However, not only are these properties of \mathcal{F} required to apply the theorem, but if they are violated, then the critical probability may take on a different value; in particular it may be lower than the result stated by our theorem.

Let us heuristically discuss why this is so. First consider the hypothesis that \mathcal{F} is strictly k -balanced. By Proposition 9.4(ii), if \mathcal{F} is not strictly k -balanced, then there exists $2 \leq \ell \leq r-1$ such that every hyperedge contains a set S of ℓ vertices such that $\deg(S) = \Omega\left(d^{1-\frac{\ell-1}{r-1}}\right)$. In particular, the bounds on the ℓ -codegrees are not sufficient to ensure that the secondary configurations remain a lower order term to the Y configurations. For example, consider the secondary configuration $X = (\mathcal{G}, R, D)$ where $|V(\mathcal{G})| = 2r - \ell$, $E(\mathcal{G}) = \{e_1, e_2\}$, $e_1 \cap e_2 = \{v_1, v_2, \dots, v_\ell\}$, $|R| = 1$, $R \subseteq e_1 \setminus e_2$ and $D = V(\mathcal{G}) \setminus R \cup \{v_1\}$. Letting ℓ be as in Proposition 9.4(ii), we would expect

$$X_v(0) = \Theta\left(d \cdot \Delta_\ell \cdot p^{2r-\ell-2}\right) = \Omega\left(d^{\frac{1}{r-1}}\right),$$

which is the same order as $Y_v^{r-2,1}(0)$. So right from the start of the process the Y configurations and hence the number of open hyperedges will grow faster than they would if the infection were uniform. Given this, we would expect the critical probability to be lower.

Similarly, if $\delta(\mathcal{F}) = 1$, then by Proposition 9.4(iii), for every vertex $e_1 \in V(\mathcal{H})$ there exists some e_2 such that $|N_{\mathcal{H}}(e_1) \cap N_{\mathcal{H}}(e_2)|$ is large. We needed condition (d) of Definition 1.5 to bound the secondary configurations with no infected vertices. As in the previous paragraph, if this bound is relaxed, the secondary configurations are no longer forced to be a lower order term compared to the Y configurations and we expect this to cause the open hyperedges to grow faster than they would if the infection was uniform.

References

- [1] J. Adler and U. Lev. Bootstrap percolation: Visualizations and applications. *Braz. J. Phys.*, 33:641–644, 2003.
- [2] M. Aizenman and J. L. Lebowitz. Metastability effects in bootstrap percolation. *J. Phys. A*, 21(19):3801–3813, 1988.
- [3] N. Alon. An extremal problem for sets with applications to graph theory. *J. Combin. Theory Ser. A*, 40(1):82–89, 1985.
- [4] H. Amini. Bootstrap percolation in living neural networks. *J. Stat. Phys.*, 141(3):459–475, 2010.
- [5] H. Amini, R. Cont, and A. Minca. Resilience to contagion in financial networks. *Math. Finance*, 26(2):329–365, 2016.
- [6] H. Amini and N. Fountoulakis. Bootstrap percolation in power-law random graphs. *J. Stat. Phys.*, 155(1):72–92, 2014.
- [7] H. Amini, N. Fountoulakis, and K. Panagiotou. Bootstrap percolation in inhomogeneous random graphs. arXiv:1402.2815v2, preprint, June 2015.
- [8] O. Angel and B. Kolesnik. Thresholds for contagious sets in random graphs. arXiv:1611.10167v1, preprint, November 2016.
- [9] R. Baber. Turán densities of hypercubes. arXiv:1201.3587v2, preprint, November 2012.
- [10] J. Balogh and B. Bollobás. Bootstrap percolation on the hypercube. *Probab. Theory Related Fields*, 134(4):624–648, 2006.
- [11] J. Balogh, B. Bollobás, H. Duminil-Copin, and R. Morris. The sharp threshold for bootstrap percolation in all dimensions. *Trans. Amer. Math. Soc.*, 364(5):2667–2701, 2012.

- [12] J. Balogh, B. Bollobás, and R. Morris. Bootstrap percolation in three dimensions. *Ann. Probab.*, 37(4):1329–1380, 2009.
- [13] J. Balogh, B. Bollobás, and R. Morris. Majority bootstrap percolation on the hypercube. *Combin. Probab. Comput.*, 18(1-2):17–51, 2009.
- [14] J. Balogh, B. Bollobás, and R. Morris. Bootstrap percolation in high dimensions. *Combin. Probab. Comput.*, 19(5-6):643–692, 2010.
- [15] J. Balogh, B. Bollobás, and R. Morris. Graph bootstrap percolation. *Random Structures Algorithms*, 41(4):413–440, 2012.
- [16] J. Balogh, B. Bollobás, R. Morris, and O. Riordan. Linear algebra and bootstrap percolation. *J. Combin. Theory Ser. A*, 119(6):1328–1335, 2012.
- [17] J. Balogh, P. Hu, B. Lidický, and H. Liu. Upper bounds on the size of 4- and 6-cycle-free subgraphs of the hypercube. *European J. Combin.*, 35:75–85, 2014.
- [18] J. Balogh, P. Hu, B. Lidický, and F. Pfender. Maximum density of induced 5-cycle is achieved by an iterated blow-up of 5-cycle. *European J. Combin.*, 52(part A):47–58, 2016.
- [19] J. Balogh and B. G. Pittel. Bootstrap percolation on the random regular graph. *Random Structures Algorithms*, 30(1-2):257–286, 2007.
- [20] C. A. Barefoot, L. H. Clark, R. C. Entringer, T. D. Porter, L. A. Székely, and Zs. Tuza. Cycle-saturated graphs of minimum size. *Discrete Math.*, 150(1-3):31–48, 1996. Selected papers in honour of Paul Erdős on the occasion of his 80th birthday (Keszthely, 1993).
- [21] F. Benevides and M. Przykucki. On slowly percolating sets of minimal size in bootstrap percolation. *Electron. J. Combin.*, 20(2):Paper 46, 20, 2013.
- [22] F. Benevides and M. Przykucki. Maximum percolation time in two-dimensional bootstrap percolation. *SIAM J. Discrete Math.*, 29(1):224–251, 2015.
- [23] P. Bennett and T. Bohman. A natural barrier in random greedy hypergraph matching. arXiv:1210.3581v1, preprint, October 2012. Accepted to *Combin. Probab. Comput.*
- [24] P. Bennett and T. Bohman. A note on the random greedy independent set algorithm. *Random Structures Algorithms*, 49(3):479–502, 2016.
- [25] T. Bohman. The triangle-free process. *Adv. Math.*, 221(5):1653–1677, 2009.

- [26] T. Bohman and P. Keevash. The early evolution of the H -free process. *Invent. Math.*, 181(2):291–336, 2010.
- [27] T. Bohman and P. Keevash. Dynamic concentration of the triangle-free process. In *The Seventh European Conference on Combinatorics, Graph Theory and Applications*, volume 16 of *CRM Series*, pages 489–495. Ed. Norm., Pisa, 2013.
- [28] B. Bollobás. On generalized graphs. *Acta Math. Acad. Sci. Hungar.*, 16:447–452, 1965.
- [29] B. Bollobás. Weakly k -saturated graphs. In *Beiträge zur Graphentheorie (Kolloquium, Manebach, 1967)*, pages 25–31. Teubner, Leipzig, 1968.
- [30] B. Bollobás. *Extremal graph theory*, volume 11 of *London Mathematical Society Monographs*. Academic Press, London-New York, 1978.
- [31] B. Bollobás, Y. Egawa, A. Harris, and G. P. Jin. The maximal number of induced r -partite subgraphs. *Graphs Combin.*, 11(1):1–19, 1995.
- [32] B. Bollobás, K. Gunderson, C. Holmgren, S. Janson, and M. Przykucki. Bootstrap percolation on Galton-Watson trees. *Electron. J. Probab.*, 19:no. 13, 27, 2014.
- [33] B. Bollobás, C. Nara, and S. Tachibana. The maximal number of induced complete bipartite graphs. *Discrete Math.*, 62(3):271–275, 1986.
- [34] P. Brass, H. Harborth, and H. Nienborg. On the maximum number of edges in a C_4 -free subgraph of Q_n . *J. Graph Theory*, 19(1):17–23, 1995.
- [35] J. I. Brown and A. Sidorenko. The inducibility of complete bipartite graphs. *J. Graph Theory*, 18(6):629–645, 1994.
- [36] R. Cerf and E. N. M. Cirillo. Finite size scaling in three-dimensional bootstrap percolation. *Ann. Probab.*, 27(4):1837–1850, 1999.
- [37] R. Cerf and F. Manzo. The threshold regime of finite volume bootstrap percolation. *Stochastic Process. Appl.*, 101(1):69–82, 2002.
- [38] J. Chalupa, P. L. Leath, and G. R Reich. Bootstrap percolation on a Bethe lattice. *J. Phys. C*, 12:L31–L35, 1979.
- [39] Y.-C. Chen. Minimum C_5 -saturated graphs. *J. Graph Theory*, 61(2):111–126, 2009.
- [40] Y.-C. Chen. All minimum C_5 -saturated graphs. *J. Graph Theory*, 67(1):9–26, 2011.

- [41] S. Choi and P. Guan. Minimum critical squarefree subgraph of a hypercube. In *Proceedings of the Thirty-Ninth Southeastern International Conference on Combinatorics, Graph Theory and Computing*, volume 189, pages 57–64, 2008.
- [42] F. R. K. Chung. Subgraphs of a hypercube containing no small even cycles. *J. Graph Theory*, 16(3):273–286, 1992.
- [43] M. Conder. Hexagon-free subgraphs of hypercubes. *J. Graph Theory*, 17(4):477–479, 1993.
- [44] D. Conlon. An extremal theorem in the hypercube. *Electron. J. Combin.*, 17(1):Research Paper 111, 7 pages, 2010.
- [45] A. N. Day. Saturated graphs of prescribed minimum degree. *Combin. Probab. Comput.*, 26(2):201–207, 2017.
- [46] D. P. Dubhashi and A. Panconesi. *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, Cambridge, 2009.
- [47] P. Erdős. On some problems in graph theory, combinatorial analysis and combinatorial number theory. In *Graph theory and combinatorics (Cambridge, 1983)*, pages 1–17. Academic Press, London, 1984.
- [48] P. Erdős and T. Gallai. On the minimal number of vertices representing the edges of a graph. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 6:181–203, 1961.
- [49] P. Erdős, A. Hajnal, and J. W. Moon. A problem in graph theory. *Amer. Math. Monthly*, 71:1107–1110, 1964.
- [50] P. Erdős and M. Simonovits. A limit theorem in graph theory. *Studia Sci. Math. Hungar.*, 1:51–57, 1966.
- [51] C. Even-Zohar and N. Linial. A note on the inducibility of 4-vertex graphs. *Graphs Combin.*, 31(5):1367–1380, 2015.
- [52] G. Exoo. Dense packings of induced subgraphs. *Ars Combin.*, 22:5–10, 1986.
- [53] J. R. Faudree, R. J. Faudree, and J. R. Schmitt. A survey of minimum saturated graphs. *Electron. J. Combin.*, 18:Dynamic Survey 19, 36 pp. (electronic), 2011.
- [54] R. J. Faudree and R. J. Gould. Saturation numbers for nearly complete graphs. *Graphs Combin.*, 29(3):429–448, 2013.

- [55] G. Fiz Pontiveros, S. Griffiths, and R. Morris. The triangle-free process and $R(3, k)$. arXiv:1302.6279v1, preprint, February 2013.
- [56] L. R. Fontes, R. H. Schonmann, and V. Sidoravicius. Stretched exponential fixation in stochastic Ising models at zero temperature. *Comm. Math. Phys.*, 228(3):495–518, 2002.
- [57] N. Fountoulakis, M. Kang, C. Koch, and T. Makai. A phase transition on the evolution of bootstrap processes in inhomogeneous random graphs. arXiv:1609.08892v1, preprint, September 2016.
- [58] P. Frankl. An extremal problem for two families of sets. *European J. Combin.*, 3(2):125–127, 1982.
- [59] D. A. Freedman. On tail probabilities for martingales. *Ann. Probability*, 3:100–118, 1975.
- [60] Z. Füredi and Y. Kim. Cycle-saturated graphs with minimum number of edges. *J. Graph Theory*, 73(2):203–215, 2013.
- [61] Z. Füredi and L. Özkahya. On 14-cycle-free subgraphs of the hypercube. *Combin. Probab. Comput.*, 18(5):725–729, 2009.
- [62] Z. Füredi and L. Özkahya. On even-cycle-free subgraphs of the hypercube. *J. Combin. Theory Ser. A*, 118(6):1816–1819, 2011.
- [63] Z. Füredi and M. Simonovits. The history of degenerate (bipartite) extremal graph problems. In *Erdős centennial*, volume 25 of *Bolyai Soc. Math. Stud.*, pages 169–264. János Bolyai Math. Soc., Budapest, 2013.
- [64] R. Gould, T. Łuczak, and J. Schmitt. Constructive upper bounds for cycle-saturated graphs of minimum size. *Electron. J. Combin.*, 13(1):Research Paper 29, 19 pages, 2006.
- [65] M. Granovetter. Threshold models of collective behavior. *American Journal of Sociology*, 8:1420–1443, 1978.
- [66] J. Gravner, A. E. Holroyd, and R. Morris. A sharper threshold for bootstrap percolation in two dimensions. *Probab. Theory Related Fields*, 153(1-2):1–23, 2012.
- [67] A. Grzesik. On the maximum number of five-cycles in a triangle-free graph. *J. Combin. Theory Ser. B*, 102(5):1061–1066, 2012.

- [68] A. Grzesik. *Flag Algebras in Extremal Graph Theory*. PhD thesis, Jagiellonian University, 2014.
- [69] K. Gunderson, S. Koch, and M. Przykucki. The time of graph bootstrap percolation. arXiv:1503.01454v2, preprint, February 2016. To appear in *Random Structures Algorithms*.
- [70] K. Gunderson and M. Przykucki. Lower bounds for bootstrap percolation on Galton-Watson trees. *Electron. Commun. Probab.*, 19:no. 42, 7, 2014.
- [71] R. W. Hamming. Error detecting and error correcting codes. *Bell System Tech. J.*, 29:147–160, 1950.
- [72] H. Hatami, J. Hirst, and S. Norine. The inducibility of blow-up graphs. *J. Combin. Theory Ser. B*, 109:196–212, 2014.
- [73] H. Hatami, J. Hladký, D. Král, S. Norine, and A. Razborov. On the number of pentagons in triangle-free graphs. *J. Combin. Theory Ser. A*, 120(3):722–732, 2013.
- [74] D. Hefetz and M. Tyomkyn. On the inducibility of cycles. arXiv:1702.0734v1, preprint, February 2017.
- [75] J. Hirst. The inducibility of graphs on four vertices. *J. Graph Theory*, 75(3):231–243, 2014.
- [76] A. E. Holroyd. Sharp metastability threshold for two-dimensional bootstrap percolation. *Probab. Theory Related Fields*, 125(2):195–224, 2003.
- [77] S. Janson. Poisson approximation for large deviations. *Random Structures Algorithms*, 1(2):221–229, 1990.
- [78] S. Janson. On percolation in random graphs with given vertex degrees. *Electron. J. Probab.*, 14:no. 5, 87–118, 2009.
- [79] S. Janson, T. Łuczak, T. Turova, and T. Vallier. Bootstrap percolation on the random graph $G_{n,p}$. *Ann. Appl. Probab.*, 22(5):1989–2047, 2012.
- [80] J. R. Johnson and Trevor Pinto. Saturated subgraphs of the hypercube. *Combin. Probab. Comput.*, 26(1):52–67, 2017.
- [81] G. Kalai. Weakly saturated graphs are rigid. In *Convexity and graph theory (Jerusalem, 1981)*, volume 87 of *North-Holland Math. Stud.*, pages 189–190. North-Holland, Amsterdam, 1984.

- [82] G. Kalai. Hyperconnectivity of graphs. *Graphs Combin.*, 1(1):65–79, 1985.
- [83] M. Kang and T. Makai. A simple proof of almost percolation on $G(n, p)$. arXiv:1608.00800v1, preprint, August 2016.
- [84] L. Kászonyi and Zs. Tuza. Saturated graphs with minimal number of edges. *J. Graph Theory*, 10(2):203–210, 1986.
- [85] P. Keevash. Counting designs. arXiv:1504.02909v1, preprint, April 2015. To appear in *J. Eur. Math. Soc.*
- [86] P. Keevash. Hypergraph Turán problems. In *Surveys in combinatorics 2011*, volume 392 of *London Math. Soc. Lecture Note Ser.*, pages 83–139. Cambridge Univ. Press, Cambridge, 2011.
- [87] P. Keevash. The existence of designs. arXiv:1401.3665v1, preprint, January 2014.
- [88] J. H. Kim and V. H. Vu. Concentration of multivariate polynomials and its applications. *Combinatorica*, 20(3):417–434, 2000.
- [89] B. Kolesnik. Sharp threshold for K_4 -percolation. arXiv:1705.08882, preprint, May 2017.
- [90] J. Komlós, A. Shokoufandeh, M. Simonovits, and E. Szemerédi. The regularity lemma and its applications in graph theory. In *Theoretical aspects of computer science (Tehran, 2000)*, volume 2292 of *Lecture Notes in Comput. Sci.*, pages 84–112. Springer, Berlin, 2002.
- [91] J. Komlós and M. Simonovits. Szemerédi’s regularity lemma and its applications in graph theory, 1996.
- [92] D. Korándi, Y. Peled, and B. Sudakov. A random triadic process. *SIAM J. Discrete Math.*, 30(1):1–19, 2016.
- [93] T. Kövari, V. T. Sós, and P. Turán. On a problem of K. Zarankiewicz. *Colloquium Math.*, 3:50–57, 1954.
- [94] L. Lovász. Flats in matroids and geometric graphs. In *Combinatorial surveys (Proc. Sixth British Combinatorial Conf., Royal Holloway Coll., Egham, 1977)*, pages 45–86. Academic Press, London, 1977.

- [95] F. J. MacWilliams and N. J. A. Sloane. *The theory of error-correcting codes. I*. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1977. North-Holland Mathematical Library, Vol. 16.
- [96] F. J. MacWilliams and N. J. A. Sloane. *The theory of error-correcting codes. II*. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1977. North-Holland Mathematical Library, Vol. 16.
- [97] W. Mantel. Problem 28. *Wiskundige Opgaven*, 10:60–61, 1907.
- [98] C. McDiarmid. On the method of bounded differences. In *Surveys in combinatorics, 1989 (Norwich, 1989)*, volume 141 of *London Math. Soc. Lecture Note Ser.*, pages 148–188. Cambridge Univ. Press, Cambridge, 1989.
- [99] B. McKay. Personal communication, 2015.
- [100] R. Morris. Minimal percolating sets in bootstrap percolation. *Electron. J. Combin.*, 16(1):Research Paper 2, 20, 2009.
- [101] R. Morris. Zero-temperature Glauber dynamics on \mathbb{Z}^d . *Probab. Theory Related Fields*, 149(3-4):417–434, 2011.
- [102] N. Morrison and J. A. Noel. Extremal bounds for bootstrap percolation in the hypercube. arXiv:1506.04686v1, preprint, June 2015.
- [103] N. Morrison, J. A. Noel, and A. Scott. On saturated k -Sperner systems. *Electron. J. Combin.*, 21(3):Paper 3.22, 17 pages, 2014.
- [104] N. Morrison, J. A. Noel, and A. Scott. Saturation in the hypercube and bootstrap percolation. *Combin. Probab. Comput.*, 26(1):78–98, 2017.
- [105] N. Morrison and A. Scott. Maximising the number of induced cycles in a graph. *J. Combin. Theory Ser. B*, 2017. <http://dx.doi.org/10.1016/j.jctb.2017.03.007>.
- [106] G. Moshkovitz and A. Shapira. Exact bounds for some hypergraph saturation problems. *J. Combin. Theory Ser. B*, 111:242–248, 2015.
- [107] A. G. Moshonkin. Concerning Hall’s theorem. In *Mathematics in St. Petersburg*, volume 174 of *Amer. Math. Soc. Transl. Ser. 2*, pages 73–77. Amer. Math. Soc., Providence, RI, 1996.

- [108] L. T. Ollmann. $K_{2,2}$ saturated graphs with a minimal number of edges. In *Proceedings of the Third Southeastern Conference on Combinatorics, Graph Theory, and Computing (Florida Atlantic Univ., Boca Raton, Fla., 1972)*, pages 367–392. Florida Atlantic Univ., Boca Raton, Fla., 1972.
- [109] G. Pete. How to make the cube weedy? *Polygon*, VII(1):69–80, 1997. in Hungarian.
- [110] O. Pikhurko. *Extremal Hypergraphs*. PhD thesis, University of Cambridge, October 1999.
- [111] O. Pikhurko. Weakly saturated hypergraphs and exterior algebra. *Combin. Probab. Comput.*, 10(5):435–451, 2001.
- [112] O. Pikhurko. Results and open problems on minimum saturated hypergraphs. *Ars Combin.*, 72:111–127, 2004.
- [113] O. Pikhurko and A. Razborov. Asymptotic structure of graphs with the minimum number of triangles. *Combin. Probab. Comput.*, 26(1):138–160, 2017.
- [114] N. Pippenger and M. Golumbic. The inducibility of graphs. *J. Combinatorial Theory Ser. B*, 19(3):189–203, 1975.
- [115] M. Przykucki. Maximal percolation time in hypercubes under 2-bootstrap percolation. *Electron. J. Combin.*, 19(2):Paper 41, 13, 2012.
- [116] A. Razborov. Flag algebras. *J. Symbolic Logic*, 72(4):1239–1282, 2007.
- [117] E. Riedl. Largest minimal percolating sets in hypercubes under 2-bootstrap percolation. *Electron. J. Combin.*, 17(1):Research Paper 80, 13, 2010.
- [118] M. Robson. personal communication, 2015.
- [119] V. Rödl. On a packing and covering problem. *European J. Combin.*, 6(1):69–78, 1985.
- [120] G.-C. Rota. *Indiscrete thoughts*. Birkhäuser Boston, Inc., Boston, MA, 1997. With forewords by Reuben Hersh and Robert Sokolowski, Edited with notes and an epilogue by Fabrizio Palombi.
- [121] L. Shapiro and A. B. Stephens. Bootstrap percolation, the Schröder numbers, and the N -kings problem. *SIAM J. Discrete Math.*, 4(2):275–280, 1991.
- [122] A. Thomason and P. Wagner. Bounding the size of square-free subgraphs of the hypercube. *Discrete Math.*, 309(6):1730–1735, 2009.

- [123] M. Truszczyński and Zs. Tuza. Asymptotic results on saturated graphs. *Discrete Math.*, 87(3):309–314, 1991.
- [124] P. Turán. On an extremal problem in graph theory. *Matematikai és Fizikai Lapok*, 48:436–452, 1941.
- [125] Zs. Tuza. A generalization of saturated graphs for finite languages. In *Proceedings of the 4th international meeting of young computer scientists, IMYCS '86 (Smolenice Castle, 1986)*, number 185, 1986.
- [126] Zs. Tuza. Extremal problems on saturated graphs and hypergraphs. *Ars Combin.*, 25(B):105–113, 1988. Eleventh British Combinatorial Conference (London, 1987).
- [127] Zs. Tuza. Asymptotic growth of sparse saturated structures is locally determined. *Discrete Math.*, 108(1-3):397–402, 1992. Topological, algebraical and combinatorial structures. Frolík's memorial volume.
- [128] Zs. Tuza. Unsolved combinatorial problems. *BRICS Lecture Series, LS-01-1*, 2001. (available from <http://www.brics.dk/publications/>).
- [129] Zs. Tuza. Problems on cycles and colorings. *Discrete Math.*, 313(19):2007–2013, 2013.
- [130] S. Ulam. Random processes and transformations. In *Proceedings of the International Congress of Mathematicians, Vol. 2, Cambridge, Mass., 1950*, pages 264–275. Amer. Math. Soc., Providence, R. I., 1952.
- [131] J. von Neumann. *Theory of Self-Reproducing Automata*. University of Illinois Press, 1966. A. W. Burks (ed.).
- [132] V. H. Vu. Concentration of non-Lipschitz functions and applications. *Random Structures Algorithms*, 20(3):262–316, 2002. Probabilistic methods in combinatorial optimization.
- [133] L. Warnke. On the method of typical bounded differences. *Combin. Probab. Comput.*, 25(2):269–299, 2016.
- [134] D. J. Watts. A simple model of global cascades on random networks. *Proc. Natl. Acad. Sci. USA*, 99(9):5766–5771, 2002.
- [135] N. C. Wormald. Differential equations for random processes and random graphs. *Ann. Appl. Probab.*, 5(4):1217–1235, 1995.

- [136] N.C. Wormald. The differential equation method for random graph processes and greedy algorithms. In *Lectures on Approximation and Randomized Algorithms*, pages 73–155. PWN, Warsaw, 1999.
- [137] J. T. Yu. An extremal problem for sets: a new approach via Bezoutians. *J. Combin. Theory Ser. A*, 62(1):170–175, 1993.
- [138] A. A. Zykov. On some properties of linear complexes. *Mat. Sbornik N.S.*, 24(66):163–188, 1949.