LOCAL APPROACHES TO GLOBAL PROBLEMS IN EXTREMAL COMBINATORICS

by

Matthew John Coulson

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Mathematics
College of Engineering and Physical Sciences
The University of Birmingham
September 2019

Abstract

In this thesis we consider five problems in extremal combinatorics all of which which are all amenable to approaches based on local structure.

The first part of this thesis looks at rainbow subgraphs at extremal thresholds. We show that as soon as they appear, we can also find rainbow copies of Perfect Matchings, H-factors and Hamilton cycles in large graphs.

We then look to random digraphs and consider the $D(n,p)$ model in which each edge is present independently with probability p. We find tail bounds on the size of the largest strongly connected component in the critical window around $p = 1/n$.

Finally, we consider the partition function of the ferromagnetic Potts model on graphs of bounded maximum degree. We show that there exists an open set in $\mathbb{C}$ containing an interval $[1, w]$ inside which the partition function has no zeros.

ACKNOWLEDGEMENTS

I would like to thank Guillem Perarnau for being a fantastic supervisor during the course of my PhD even after moving to Barcelona. Thanks also to Richard Mycroft for taking over duties as my supervisor after Guillem left.

Thanks are also due to all of my collaborators without whom I would certainly not have worked on as many interesting problems in combinatorics. In particular thanks to Viresh Patel and Guus Regts for hosting me in Amsterdam for two months while we worked together.

Finally, thanks to my office mates in “the dungeon” for providing great company and helping make my time in Birmingham better.

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Extremal Combinatorics

Extremal combinatorics is concerned with problems related to the containment of certain substructures. In particular one can ask how many edges a graph must contain before containing a given subgraph. Or alternatively, what minimum degree is required to contain a given spanning subgraph.

Dirac's theorem [33] states that if a graph with n vertices has minimum degree at least $n/2$, then it has a Hamilton cycle. We call such graphs *Dirac graphs*. One can deduce that such a graph has a perfect matching if it has an even number of vertices by picking alternate edges of this cycle. A bipartite graph with parts of size n has a perfect matching precisely when it satisfies Hall's condition [52]. Note that such a graph of minimum degree $n/2$ can easily be seen to satisfy Hall's condition. We call bipartite graphs with parts of size n and minimum degree $n/2$ *Dirac bipartite graphs*.

There are many further results on when one can find certain spanning subgraphs of a large graph based upon the minimum degree. The Hajnal-Szemerédi [51] theorem states that one can find a K_k-factor of any graph with n vertices and minimum degree at least $(1 - 1/k)n$ (provided the obvious divisibility condition is satisfied). This was generalised to H-factors for arbitrary H by Alon and Yuster [6] and subsequently improved by Kühn and Osthus [75, 77] who gave an optimal threshold up to an additive constant factor.

In general the question of when one can find graphs of bounded maximum degree is still open and the following was conjectured by Bollobás Eldridge and Catlin [11, 19].

Conjecture 1.1 (Bollobás-Eldridge-Catlin Conjecture). *Let G and H be graphs with n vertices. If H has maximum degree at most Δ and G has minimum degree greater than $(1 - \frac{1}{\Delta+1})n$, then H is a subgraph of G.*

This conjecture is however known to be true for graphs of sublinear bandwidth [15]:

Theorem 1.2 (Bandwidth Theorem). *For all $r, \Delta \in \mathbb{N}$ and $\gamma > 0$, there exist constants $\beta > 0$ and $n_0 \in \mathbb{N}$ such that for every $n \geq n_0$ the following holds.*

If H is an r-colourable graph on n vertices with $\Delta(H) \leq \Delta$, and bandwidth at most βn and if G is a graph on n vertices with minimum degree at least $(1 - 1/r + \gamma)n$, then G contains a copy of H.

Note that the issue of the substitution of the maximum degree of H for its chromatic number may be rectified by appealing to Brooks' theorem [16].

1.2 Local Approaches

Methods used in the study of extremal problems can broadly be grouped into two types: local and global. Global approaches use holistic properties of the graph or structure in question to draw conclusions. Examples of such methods include Alon's combinatorial nullstellensatz [4], Szemerédi's regularity lemma [113] and container-based arguments [66, 67].

In contrast, local approaches look at or edit very small pieces of the graph at a time. In this thesis we use a number of local methods to study large scale structure and properties of graphs, hypergraphs and digraphs. In particular we use switchings, exploration processes and vertex neighbourhood based methods. We describe these methods in the following sections.

1.2.1 Switchings

Informally, a switching is a small local change to a graph or other structure which we can control well. For example, we might say that two matchings M_1 and M_2 are related by switching over a 4-cycle if the graph $M_1 \Delta M_2$ is a 4-cycle. The benefit of switchings is that they are often able to give good bounds on the probability a random subgraph of a graph which is difficult to sample from contains some collection of edges. For example, consider the following bound on the probability that a uniformly random perfect matching in a graph G of minimum degree at least $(1/2+\varepsilon)n$ contains a given edge e.

The proof proceeds by a double counting argument. Let $\mathcal{F}$ be the collection of perfect matchings in G and $\mathcal{F}_e$ be the set of those which contain the edge e. We construct an auxiliary bipartite graph $\mathcal{G}$ between $\mathcal{F}_e$ and $\mathcal{F} \backslash \mathcal{F}_e$ where M_1 and M_2 are connected if they are related by a 4-cycle switching. For a given $M_1 \in \mathcal{F}_e$, by looking at neighbourhoods, there are at least $2\varepsilon n - 2$ 4-cycles containing e and another edge of M_1. Given $M_2 \in \mathcal{F} \backslash \mathcal{F}_e$, there is at most one 4-cycle containing two edges of M_2 as well as e. So, $\delta(\mathcal{F}_e) \geq 2\varepsilon n - 2$ and $\Delta(\mathcal{F} \setminus \mathcal{F}_e) \leq 1$. Counting the edges of $\mathcal{G}$ using these bounds we discover $(2\varepsilon n - 2)|\mathcal{F}_e| \leq |E(\mathcal{G})| \leq |\mathcal{F} \setminus \mathcal{F}_e|$. As we choose a perfect matching uniformly at random from all possible perfect matchings of G, the probability that this perfect matching contains e is precisely $|\mathcal{F}_e|/|\mathcal{F}|$ which can be bounded as follows,

$$\frac{|\mathcal{F}_e|}{|\mathcal{F}|} \leq \frac{|\mathcal{F}_e|}{|\mathcal{F} \setminus \mathcal{F}_e|} \leq \frac{1}{2\varepsilon n - 2}.$$

That is, a random perfect matching in G contains e with probability at most $\frac{1}{2\varepsilon n - 2}$, which is clearly of the correct order of magnitude.

We will use switchings in Chapters 4, 5 and 6 to bound the probabilities of containing a pair of edges with the same colour in given subgraphs.

1.2.2 Exploration Processes

Exploration processes are commonly used to find connected components of a graph. We run an exploration process as follows. Partition the vertex set of the graph into three sets: active, neutral and explored vertices. At each step, we pick an active vertex v and look at its neighbours and add any neutral neighbours into the active set. After this is done we move v into the explored set.

The method of selecting which active vertex to explore next is often important in algorithmic applications. Two common paradigms are Breadth-first search (BFS) and Depth-first search (DFS). In BFS vertices are explored in the order which they are found. This gives rise to a search tree in which all but the lowest two layers are completely explored at any point. This makes BFS an excellent choice for shortest path finding and related algorithms. DFS algorithms explore the most recently added active vertices first. This means that the algorithm gives rise to search trees that are commonly very deep. As such DFS search procedures are often used to find long paths in random graphs or for maze solving.

In chapter 7 we use an exploration process to find the descendants of a subset of a random digraph. In this case we are only interested in the number of descendants so the way in which we choose the next active vertex to explore does not matter as any procedure will find all of the descendants.

1.2.3 Neighbourhoods

In chapter 8 we study zero-free regions of the ferromagnetic Potts model. The proof of the zero-free region is by showing that changing the spin of one vertex does not have a large effect on the partition function. To do this we look at one vertex at a time, replacing it with one copy of itself connected to each of its neighbours so that we only need to look at vertices of degree 1 in a similar way to previous work by Bencs et al. [9]. This could also be viewed as another type of switching.

1.3 Our Work

We study a number of problems related to extremal combinatorics with very local methods. A short introduction to each topic and a statement of the main results can be found in this section. In addition, a more in depth introduction is provided at the start of each of the relevant chapters.

1.3.1 Rainbow Subgraphs

Given a graph G and an edge colouring $\chi : E(G) \to \mathbb{N}$ of G, the subgraph $H \subseteq G$ is *rainbow* if for every $c \in \mathbb{N}$, $|\chi^{-1}(c) \cap E(H)| \leq 1$. That is, H is rainbow if and only if each of its edges has a different colour.

The study of rainbow substructures originated with the study of Latin squares. A *Latin square* of order n is an $n \times n$ array of symbols where each symbol occurs precisely one in each row and column. A *partial transversal* of size k in a Latin square is a set of cells, including at most one from each row and each column that contains k distinct symbols. The question of finding the largest transversal in an arbitrary Latin square has attracted considerable attention in particular it has been conjectured that it is almost possible to find a full transversal.

Conjecture 1.3 (Ryser, Brualdi, Stein [17, 103, 109])**.** *Every Latin square of order n contains a partial transversal of size at least $n-1$.*

The link with rainbow subgraphs is due to a natural bijection between Latin squares of order n and proper edge colourings of the complete balanced bipartite graph on $2n$ vertices such that a partial transversal of size k in the Latin square is mapped to to a rainbow matching of size k.

One can extend the problem to edge colourings of $K_{n,n}$ that satisfy a milder condition. An edge colouring is *k-bounded* if $|\chi^{-1}(c)| \leq k$ for every $c \in \mathbb{N}$. Stein [109] conjectured that Conjecture 1.3 still holds for n-bounded edge colourings. This is true if the size of each colour class is small enough.

Theorem 1.4 (Erdős, Spencer [39]). *Let $K_{n,n}$ be the complete bipartite graph on $2n$ vertices, then any $(n-1)/16$-bounded edge colouring of $K_{n,n}$ contains a rainbow perfect matching.*

This motivates further study of rainbow substructures of large graphs and the properties which they have. In particular we look at three structures whose extremal thresholds have been well studied: Perfect matchings, H-factors and Hamilton cycles. We show that as soon as these structures appear in a graph, a rainbow copy can also be found in a μn-bounded edge colouring of the same graph.

Perfect Matchings

First we look at perfect matchings. It is an easy corollary of Hall's theorem that a bipartite graph with n vertices in each part and minimum degree $n/2$ has a perfect matching. We will call such graphs Dirac bipartite graphs. It turns out that this is all we need to find rainbow perfect matchings.

Theorem 1.5. *There exist $\mu > 0$ and $n_0 \in \mathbb{N}$ such that if $n \geq n_0$ and G is a Dirac bipartite graph on $2n$ vertices, then any μn-bounded edge colouring of G contains a rainbow perfect matching.*

H-factors

An H-factor in a graph G is a collection of $|G|/|H|$ vertex disjoint copies of the graph H (note that we require that $|H|$ divides $|G|$ here). In graphs the threshold for their existence is well understood [77] however no similar result exists for hypergraphs. Our result shows that once an H-factor exists, so does a rainbow H-factor and only requires the knowledge that a threshold for the existence of an H-factor exists. We will denote this threshold as $\delta^*_\ell(H)$ and refer the reader to (5.1) for the definition.

Theorem 1.3.1. *Let $1/n \ll \mu \ll \varepsilon \ll 1/h \leq 1/r < 1/\ell \leq 1$ with $h|n$ and $\ell, r, h, n \in \mathbb{N}$. Let H be an r-graph on h vertices and G be an r-graph on n vertices with $\delta_\ell(G) \geq$*

$(\delta_\ell^*(H)+\varepsilon)n^{r-\ell}$. *Then any μ-bounded edge-colouring of G admits a rainbow H-factor.*

Hamilton cycles

A Dirac graph on n vertices is any graph with minimum degree $n/2$. Dirac's theorem tells us that this is the threshold for the existence of a Hamilton cycle. We generalise this theorem and show further that this threshold also suffices for the existence of a rainbow Hamilton cycle.

Theorem 1.6. *There exist $\mu > 0$ and $n_0 \in \mathbb{N}$ such that if $n \geq n_0$ and G is a Dirac graph on n vertices, then any μn-bounded colouring of $E(G)$ contains a rainbow Hamilton cycle.*

1.3.2 Random Digraphs

The binomial random digraph $D(n,p)$ is defined analogously to the Erdős-Renyi random graph $G(n,p)$; Each arc is present independently with probability p. A strong component in a digraph D is a maximal subset S of the vertices such that for each pair $u, v \in S$ there exists a directed $u-v$ path and $v-u$ path.

The threshold for the existence of a giant strong component was shown to be when $p = 1/n$ by Karp [60] and Łuckzak [84] independently. This was refined by Łuczak and Seierstad [86] who studied the strong components of $D(n,p)$ in the range $p = (1+\varepsilon)/n$ for $\varepsilon \to 0$ and $|\varepsilon|^3 n \to \infty$. We further refine this and study the critical window, $p = n^{-1} + \lambda n^{-4/3}$ for λ constant providing tail bounds on the size of the largest strong component in this range.

Theorem 1.7 (Lower Bound)**.** *Let $0 < \delta < 1/800$, $\lambda \in \mathbb{R}$ and $n \in \mathbb{N}$. Let $\mathcal{C}_1$ be the largest component of $D(n,p)$ for $p = n^{-1} + \lambda n^{-4/3}$. Then if n is sufficiently large with respect to δ, λ,*

$$\mathbb{P}(|\mathcal{C}_1| < \delta n^{1/3}) \leq 2e\delta^{1/4}$$

provided that $\delta \leq \frac{(\log 2)^2}{4|\lambda|^2}$.

Theorem 1.8 (Upper Bound). *Let $\mathcal{C}_1$ be the largest component of $D(n,p)$ for $p = n^{-1} + \lambda n^{-4/3}$. There exist constants, $\zeta, \eta > 0$ such that for any $A > 0, \lambda \in \mathbb{R}$the following holds. Provided n is sufficiently large with respect to A, λ, and defining $\lambda^+ := \max(\lambda, 0)$,*

$$\mathbb{P}(|\mathcal{C}_1| > An^{1/3}) \leq \zeta e^{-\eta A^{3/2} + \lambda^+ A}.$$

1.3.3 Zero-Free Regions

Given a graph G, $k \in \mathbb{N}$, and $w \in \mathbb{C}$ the partition function of the univariate Potts model is defined as

$$\mathbf{Z}(G;k,w) := \sum_{\phi: V \to [k]} \prod_{\substack{uv \in E \\ \phi(u)=\phi(v)}} w,$$

For $w > 0$ real, this can be viewed as the normalising constant for a family of probability distributions over functions, $\phi : V \to [k]$ where the mass of ϕ is given by

$$\mu_{G;w}(\phi) = \mathbf{Z}(G;k,w)^{-1} \prod_{\substack{uv \in E \\ \phi(u)=\phi(v)}} w$$

When $w < 1$, this is often referred to as the anti-ferromagnetic Potts model as connected particles with the same spin repel one another. This regime was studied by Bencs et al. [9] who found an open subset of $\mathbb{C}$ containing $[0, 1]$ (for w) on which $\mathbf{Z}(G;k,w) \neq 0$. Combining their results with a method of Patel and Regts [96] they find a fully polynomial time approximation scheme (FPTAS) to count k-colourings of bounded degree graphs whenever $k \geq e\Delta(G)$.

The case $w > 1$ is known as the ferromagnetic Potts model as particles with the same spin are attracted to each other. There are a number of barriers to finding zero-free regions containing any large interval $[1, x]$ for any $x > 1$. In particular, this would imply the existence of an FPTAS for #BIS [43, 45] if one can take $x > 1 + \frac{2\log k}{\Delta(G)}$ approximately. However, this would conflict with the commonly held belief that #BIS is an NP-intermediate problem i.e. neither in P nor NP. We show that one can get almost half way to this obstruction.

Theorem 1.9. *Let* $k, \Delta \geq 3$ *then there exists an open set* U *containing the interval* $[1, 1 + \frac{\log(k)-1}{\Delta}]$ *such that for any* $w \in U$ *and any graph* G *of maximum degree at most* Δ, $\mathbf{Z}(G; k, w) \neq 0$.

The work presented in this thesis is joint work with various subsets of Ewan Davies [26], Peter Keevash [27], Alexandra Kolla [26], Guus Regts [26], Viresh Patel [26], Guillem Perarnau [27, 28, 29] and Liana Yepremyan [27] as well as some solo work [25].

CHAPTER 2

NOTATION AND TERMINOLOGY

This chapter contains the major notation which we use for the remainder of the thesis. We shall repeat any definitions given in the introduction for convenience.

2.1 Graphs

A *graph* $G = (V, E)$ is an ordered pair where V is a set of *vertices* and E is a collection of unordered pairs of elements from V is the set of *edges*. If u and v are vertices of $G = (V, E)$ write uv for the edge $\{u, v\}$ we say u and v are the *ends* of uv. We write $V(G)$, $E(G)$ for the sets of vertices and edges of the graph G respectively. Furthermore define $v(G) = |V(G)|$ and $e(G) = |E(G)|$ for the numbers of vertices and edges of G.

Given a vertex v of a graph $G = (V, E)$, its *neighbourhood* is $N(v)$, the set of vertices $u \in V$ such that $uv \in E$. Its *closed neighbourhood* is $N(v) \cup \{v\}$ and its *degree* is $d(v) = |N(v)|$. If $X \subseteq V$, we let $N_X(v) = N(v) \cap X$ and define $d_X(v) = |N_X(v)|$. Furthermore the neighbourhood of X is defined as $N(X) = \cup_{x \in X} N(x)$. The *minimum degree* of G is $\delta(G) = \min_{v \in V(G)} d(v)$. Similarly the *maximum degree* of G is $\Delta(G) = \max_{v \in V(G)} d(v)$.

A *subgraph* of $G = (V, E)$ is a graph $G' = (V', E')$ such that $V' \subseteq V$ and $E' \subseteq E$. Write $G' \subseteq G$ to mean G' is a subgraph of G. We say a subgraph G' of G is *spanning* if $V(G') = V(G)$. An *induced subgraph* of G is a graph $G' = (V', E')$ such that $uv \in E'$ if and only if $u, v \in V'$ and $uv \in E$. We write $G[V']$ for the induced subgraph $G' = (V', E')$

and call it the subgraph induced by V'.

A *path* P is a graph such that there exists an ordering $v_0 v_1 \dots v_m$ of the vertices of P where the edges of P are precisely $v_{i-1}v_i$ for $i \in [m]$. A $u - v$ *path* in the graph G is a subgraph of G which is a path and such that $v_0 = u$ and $v_m = v$. A *connected component* of the graph G is a maximal induced subgraph C of G such that for every pair of vertices $u, v \in V(C)$ there exists a $u - v$ path in G. If G has an unique connected component then we say G is *connected*.

A *matching* in a graph G is a subset M of $E(G)$ consisting of vertex disjoint edges. A *perfect matching* in G is a matching M such that each $v \in V(G)$ is contained in some element of M.

A *cycle* C of length $m \geq 3$ is a graph such that there exists an ordering $v_1 v_2 \dots v_m$ of the vertices of C where the edges of C are precisely $v_i v_{i+1 \mod m}$ for $i \in [m]$. A *Hamilton cycle* in a graph G is a subgraph of G which is a cycle of length $v(G)$.

Let H be a fixed graph. An *H-factor* is a collection of vertex disjoint copies of H which covers all the vertices of a host graph G. If the latter condition is not satisfied we will use the phrase *partial H-factor*.

2.1.1 Bipartite Graphs

A graph $G = (V, E)$ is *bipartite* if there exists a partition (A, B) of the vertex set such that every edge of G has exactly one end in A. We will often write $G = (A, B; E)$ if G is a bipartite graph with partition (A, B). All definitions for graphs as stated above still hold for bipartite graphs. In addition we also define the maximum and minimum A-degree or B-degree as $\delta_A(G) = \min_{v \in A} d(v)$ and the other three definitions are analogous.

2.2 Digraphs

A *digraph* $D = (V, E)$ is an ordered pair where V is a set of *vertices* and E is a collection of ordered pairs of elements from V is the set of *edges*. If u and v are vertices of $D = (V, E)$ write uv for the edge (u, v) we say u and v are the *ends* of uv.

Given a vertex v of D its *out-neighbourhood*, $N^+(v)$ is the set of vertices u such that $vu \in E$. Its *in-neighbourhood*, $N^-(v)$ is the set of vertices u such that $uv \in E$. We may define the *in-degree* and *out-degree* as $d^-(v) = |N^-(v)|$ and $d^+(v) = |N^+(v)|$ respectively. We also define the minimum and maximum in or out-degrees in the obvious way.

A *subdigraph* of $D = (V, E)$ is a digraph $D' = (V', E')$ such that $V' \subseteq V$ and $E' \subseteq E$. An *induced subdigraph* of D is a digraph $D' = (V', E')$ such that $uv \in E'$ if and only if $u, v \in V'$ and $uv \in E$. We write $D[V']$ for the induced subgraph $D' = (V', E')$ and call it the subdigraph induced by V'. The *underlying graph* of a digraph $D = (V, E)$ is the graph, $G = (V, E')$ where $uv \in E'$ if and only if $uv \in E$ or $vu \in E$.

Furthermore, a path may be defined exactly the same as in graphs. However, note that these paths are directed and so a $u-v$ path is not the same as a $v-u$ path. A *weakly connected component* of the digraph D is a connected component C of the underlying graph G. The *in-component* of v in the digraph D is the set of vertices $u \in V(D)$ such that there exists a $u - v$ path in D. The *out-component* of v in the digraph D is the set of vertices $u \in V(D)$ such that there exists a $v - u$ path in D. A *strongly connected component* of the digraph D is a maximal induced subgraph C of G such that for every pair of vertices $u, v \in V(C)$ there exists a $u - v$ path and a $v - u$ path in G.

2.3 Hypergraphs

A *hypergraph* $H = (V, E)$ is an ordered pair where V is a set of *vertices* and E is a collection of subsets of V which we call the set of *edges*. A hypergraph is r-*uniform* or an r-*graph* if every edge contains r elements. So, in particular a graph is a 2-uniform hypergraph.

For a set L of ℓ vertices of H, its *ℓ-neighbourhood* $N_\ell(L)$ is the collection of all edges $e \in E$ such that $L \subseteq e$. Its *ℓ-degree* is $d_\ell(L) = |N_\ell(L)|$. The *minimum ℓ-degree* of H is $\delta_\ell(H) = \min_{L \subseteq V^{(\ell)}} d_\ell(L)$ and the *maximum ℓ-degree* is $\Delta_\ell(H) = \max_{L \subseteq V^{(\ell)}} d_\ell(L)$.

We define a *subgraph* of H analogously to subgraphs of a graph G. If H is a fixed hypergraph and $\mathcal{H}$ a hypergraph such that $v(H)|v(\mathcal{H})$, we say $F \subseteq \mathcal{H}$ is an H-factor if it consists of $v(\mathcal{H})/v(H)$ vertex disjoint copies of H.

2.4 Common Definitions

In this section we give some definitions which do not rely on whether we have a graph, digraph, hypergraph or any other combinatorial structure. We state these definitions for graphs, however they may be extended in the obvious way to digraphs or hypergraphs.

A colouring of the vertices (or edges) of a graph G is a function $\phi : V(G) \to C$ (or $\phi : E(G) \to C$) where C is a set of colours. Often we will take $C = [k]$, an initial segment of $\mathbb{N}$. A graph G together with a colouring ϕ of its vertices (edges) is called a vertex (edge) coloured graph.

A subgraph G' of an edge (vertex) coloured graph G is called rainbow if the restriction of ϕ to $E(G')$ $(V(G'))$ is injective.

Given a graph G, $k \in \mathbb{N}$, and $w \in \mathbb{C}$ the partition function of the univariate Potts model is defined as

$$\mathbf{Z}(G; k, w) := \sum_{\phi : V \to [k]} \prod_{\substack{uv \in E \\ \phi(u) = \phi(v)}} w,$$

In many of the results of this thesis we use hierarchies of the form

$$a \ll b \ll c$$

By this we mean that we may select the constants from right to left where we may pick

$b \leq f(c) < c$, and $a \leq g(b) < b$ etc. for some non-decreasing functions f, g of c, b respectively. For example, we could have $5\varepsilon^3 + 2\varepsilon e^2 \ll \varepsilon \ll \frac{-1}{\log(\varepsilon)} \ll 1$.

In principle we may be able to explicitly compute the constants in question, however this does not add anything to the theorems in question as the constants involved are rather small. (For example the proof of Theorem 1.5 can be made to work taking $\mu \approx 10^{-16}$.)

CHAPTER 3

THE LOPSIDED LOVÁSZ LOCAL LEMMA

The Lovász local lemma originated in a paper of Erdős and Lovász [38]. A lopsided generalisation due to Erdős and Spencer [39] has often been a key tool in the study of rainbow subgraph problems. We shall apply a similar generalisation a number of times in Chapters 4, 5 and 6 where we study rainbow subgraph problems at extremal thresholds.

In this chapter we will introduce the associated notation and prove the versions of the local lemma which we will require later.

The general setting we shall work in is as follows. Let $\mathcal{E} = \{E_1, \ldots, E_n\}$ be a finite family of events over some probability space. We would like to show that

$$\mathbb{P}\left(\bigcap_{i=1}^{n} E_i^c\right) > 0. \tag{3.1}$$

That is, we wish to show that with positive probability none of the events in $\mathcal{E}$ occur. Of course if $\mathcal{E}$ is a family of independent events, then provided no event holds with probability 1, the inequality (3.1) follows by the definition of independence. The Lovász local lemma and its generalisations in some sense gives a definition of the family $\mathcal{E}$ being "nearly independent" in the sense that it provides a sufficient condition such that inequality (3.1) still holds.

3.1 The Lovász Local Lemma

We begin with a definition. Let $\mathcal{E}$ be a finite family of events and suppose that D is a digraph with vertex set $V(D) = \mathcal{E}$. Suppose further that for each $E \in \mathcal{E}$, E is independent of the family $\mathcal{E} \setminus (\{E\} \cup N^+(E))$. Then we say that D is a dependency digraph for $\mathcal{E}$.

In this thesis we shall more commonly work with dependency graphs for simplicity. We say G with vertex set $\mathcal{E}$ is a dependency graph for $\mathcal{E}$ if for each $E \in \mathcal{E}$, E is independent of the family $\mathcal{E} \setminus (\{E\} \cup N(E))$. Note that a dependency digraph naturally gives rise to a dependency graph formed by undirecting each edge and removing any double edges that may be formed.

With the notion of dependency digraphs we may now state the local lemma.

Lemma 3.1 (Lovász Local Lemma). *Suppose that $\mathcal{E}$ is a finite family of events with dependency (di)graph D and that for each $E \in \mathcal{E}$ there exists $x_E \in (0,1)$ such that*

$$\mathbb{P}(E) \leq x_E \prod_{(E,F) \in E(G)} (1 - x_F). \tag{3.2}$$

Then,

$$\mathbb{P}\Bigg(\bigcap_{E \in \mathcal{E}} E^c\Bigg) \geq \prod_{E \in \mathcal{E}} (1 - x_E) > 0. \tag{3.3}$$

We defer the proof to section 3.3 where we prove a generalisation of this lemma.

3.2 The Lopsided Lovász Local Lemma

The independence condition in the local lemma is very strong and often difficult to check in practice. An observation by Erdős and Spencer [39] is that we do not require independence and in fact a negative correlation condition suffices to deduce the conclusion of Lemma 3.1. Thus we define the *lopsidependency digraph* as follows.

Let $\mathcal{E}$ be a finite family of events and suppose that D is a digraph with vertex set $V(D) = \mathcal{E}$. Suppose further that for each $E \in \mathcal{E}$, the following holds for all sets $\mathcal{S} \subseteq$

$\mathcal{E} \setminus (\{E\} \cup N^+(E))$ with $\mathbb{P}(\cap_{F \in \mathcal{S}} F^c) > 0$.

$$\mathbb{P}\left(E \middle| \bigcap_{F \in \mathcal{S}} F^c\right) \leq \mathbb{P}(E). \tag{3.4}$$

Then we say that D is a lopsidependency digraph for $\mathcal{E}$. Lopsidependency graphs may be defined analogously.

Now, the Local Lemma holds as before with lopsidependency (di)graphs in place of dependency (di)graphs.

Lemma 3.2 (Lopsided Lovász Local Lemma)**.** *Suppose that $\mathcal{E}$ is a finite family of events with lopsidependency (di)graph D and that for each $E \in \mathcal{E}$ there exists $x_E \in (0,1)$ such that*

$$\mathbb{P}(E) \leq x_E \prod_{(E,F) \in E(G)} (1 - x_F). \tag{3.5}$$

Then,

$$\mathbb{P}\left(\bigcap_{E \in \mathcal{E}} E^c\right) \geq \prod_{E \in \mathcal{E}} (1 - x_E) > 0. \tag{3.6}$$

3.3 The p-Lopsided Lovász Local Lemma

Using $\mathbb{P}(E)$ as a bound in equations (3.4) and (3.5) also turns out to be unnecessary. All we require here is any bound which does not depend on what we condition on. This motivates the definition of the **p**-dependency (di)graph.

So, let $\mathcal{E}$ be a finite family of events and suppose that D is a (di)graph with vertex set $V(D) = \mathcal{E}$. Suppose further that for each $E \in \mathcal{E}$, there exists $p_E \in (0,1)$ and that the following holds for all sets $\mathcal{S} \subseteq \mathcal{E} \setminus (\{E\} \cup N^+(E))$ with $\mathbb{P}(\cap_{F \in \mathcal{S}} F^c) > 0$.

$$\mathbb{P}\left(E \middle| \bigcap_{F \in \mathcal{S}} F^c\right) \leq p_E. \tag{3.7}$$

Then we say that D is a **p**-dependency (di)graph for $\mathcal{E}$ where $\mathbf{p} = (p_E : E \in \mathcal{E})$.

Then the **p**-lopsided Lovász local lemma is then essentially the same as the lopsided version with $\mathbb{P}(E)$ replaced by p_E. We shall give a proof following the proof of the Lovász local lemma by Spencer [108] with minor adaptations.

Lemma 3.3 (**p**-Lopsided Lovász Local Lemma)**.** *Suppose that $\mathcal{E}$ is a finite family of events and $\mathbf{p}$ a vector indexed by $\mathcal{E}$. Suppose further that $\mathcal{E}$ has $\mathbf{p}$-dependency (di)graph G and that for each $E \in \mathcal{E}$ there exists $x_E \in (0,1)$ such that*

$$p_E \leq x_E \prod_{(E,F)\in E(G)} (1 - x_F). \tag{3.8}$$

Then,

$$\mathbb{P}\bigg(\bigcap_{E\in\mathcal{E}} E^c\bigg) \geq \prod_{E\in\mathcal{E}}(1 - x_E) > 0. \tag{3.9}$$

Proof. For an arbitrary subset $\mathcal{S} \subseteq \mathcal{E}$ we define

$$B_{\mathcal{S}} = \bigcap_{F\in\mathcal{S}} F^c, \tag{3.10}$$

where we use the convention that the empty intersection is the entire probability space.

We will show by induction on $|\mathcal{S}|$ that for any $E \notin \mathcal{S}$,

$$\mathbb{P}\bigg(E\bigg|\bigcap_{F\in\mathcal{S}} F^c\bigg) \leq x_E \tag{3.11}$$

When $|\mathcal{S}| = 0$, we have

$$\mathbb{P}\bigg(E\bigg|\bigcap_{F\in\mathcal{S}} F^c\bigg) = \mathbb{P}(E) \leq p_E \leq x_E \prod_{(E,F)\in E(G)} (1 - x_F) \leq x_E$$

as required. Now, suppose that the claim holds whenever $|\mathcal{S}| < r$. Suppose further that $|\mathcal{S}| = r$ and $E \notin \mathcal{S}$. Define $\mathcal{T} = \{F \in \mathcal{S} : (E,F) \in E(D)\}$. Now,

$$\mathbb{P}(E|B_{\mathcal{S}}) = \frac{\mathbb{P}(E \cap B_{\mathcal{S}})}{\mathbb{P}(B_{\mathcal{S}})} = \frac{\mathbb{P}(E \cap B_{\mathcal{S}})}{\mathbb{P}(B_{\mathcal{S}\setminus\mathcal{T}})} \frac{\mathbb{P}(B_{\mathcal{S}\setminus\mathcal{T}})}{\mathbb{P}(B_{\mathcal{S}\setminus\mathcal{T}} \cap B_{\mathcal{T}})} = \frac{\mathbb{P}(E \cap B_{\mathcal{T}}|B_{\mathcal{S}\setminus\mathcal{T}})}{\mathbb{P}(B_{\mathcal{T}}|B_{\mathcal{S}\setminus\mathcal{T}})}. \tag{3.12}$$

We now bound the numerator and denominator of equation (3.12) separately. First for the numerator, note that $\mathbb{P}(E \cap B_{\mathcal{T}}|B_{\mathcal{S}\setminus\mathcal{T}}) \leq \mathbb{P}(E|B_{\mathcal{S}\setminus\mathcal{T}})$. By the definition of the **p**-dependency (di)graph and the fact the $\mathcal{S} \setminus \mathcal{T}$ is a subset of the non-neighbours of E in D,

$$\mathbb{P}(E|B_{\mathcal{S}\setminus\mathcal{T}}) \leq p_E \leq x_E \prod_{(E,F)\in E(G)} (1 - x_F). \tag{3.13}$$

Next, we look at the denominator. Suppose that $|\mathcal{T}| = t$, we let $\emptyset = \mathcal{R}_0 \subseteq \mathcal{R}_1 \subseteq \mathcal{R}_2 \subseteq \ldots \subseteq \mathcal{R}_t = \mathcal{T}$ be any sequence of subsets of $\mathcal{T}$ such that $|\mathcal{R}_i| = i$ for each $i \in [t]$. Define E_i to be the unique event in $\mathcal{R}_i \setminus \mathcal{R}_{i-1}$ for $i \in [t]$. Then, we have the telescoping product,

$$\begin{aligned}\mathbb{P}(B_{\mathcal{T}}|B_{\mathcal{S}\setminus\mathcal{T}}) = \frac{\mathbb{P}(B_{\mathcal{S}})}{\mathbb{P}(B_{\mathcal{S}\setminus\mathcal{T}})} &= \prod_{i=1}^{t} \frac{\mathbb{P}(B_{\mathcal{S}\setminus R_{i-1}})}{\mathbb{P}(B_{\mathcal{S}\setminus R_i})} \\ &= \prod_{i=1}^{t} \mathbb{P}(E_i^c|B_{\mathcal{S}\setminus R_i}) \geq \prod_{i=1}^{t}(1 - x_{E_i}) \geq \prod_{(E,F)\in E(G)} (1 - x_F) \end{aligned} \tag{3.14}$$

where the second last inequality follows by induction as $|\mathcal{S}\setminus\mathcal{R}_i| < r$ for each i. Combining equations (3.13) and (3.14) yields claim (3.11).

Finally, to deduce (3.9) we simply apply a telescoping product over $\mathcal{E}$ applying (3.11) to each term. □

3.4 Corollaries

The form of the **p**-lopsided Lovász local lemma as stated in Lemma 3.3 is often referred to as the general form of the local lemma. The conditions we need to check can be complicated and so it is often helpful to use simpler forms of the local lemma. In particular when we have some knowledge about the relations between the events it is usually possible to find a simpler condition to check than (3.8).

In this section we deduce a number of these simpler conditions. We will use these conditions in various places throughout this thesis when applying the local lemma. We start with the symmetric form which is designed for when all of the events "look the same".

Corollary 3.4 (Symmetric Form)**.** *Suppose that we are in the setting of Lemma 3.3 and that $p_E \leq p$ for each $E \in \mathcal{E}$ and furthermore that the dependency (di)graph has (out) degree at most d. If $4pd \leq 1$ then,*

$$\mathbb{P}\bigg(\bigcap_{E\in\mathcal{E}} E^c\bigg) \geq (1-2p)^{|\mathcal{E}|} > 0.$$

Proof. We take $x_E = 2p$ for each $E \in \mathcal{E}$. We must then show that

$$p \leq 2p(1-2p)^d. \tag{3.15}$$

By assumption, $4pd \leq 1$ so it suffices to show for all $d \in \mathbb{N}$ that

$$(1-\frac{1}{2d})^d \geq \frac{1}{2}. \tag{3.16}$$

However this holds by the standard inequality $(1-x)^d \geq 1 - xd$ thus condition (3.8) holds provided that $4pd \leq 1$. Therefore,

$$\mathbb{P}\bigg(\bigcap_{E\in\mathcal{E}} E^c\bigg) \geq (1-2p)^{|\mathcal{E}|} > 0$$

as required. □

The next form we look at is the bounded form which we will use when there are a small number of types of events to control. The proof of this and the subsequent weighted form are due to Molloy and Reed [90, Chapter 19].

Corollary 3.5 (Bounded Form)**.** *Suppose that we are in the setting of Lemma 3.3 and that for each $E \in \mathcal{E}$ we have*

$$\sum_{F:(E,F)\in E(G)} p_F \leq \frac{1}{4} \tag{3.17}$$

then,

$$\mathbb{P}\Bigg(\bigcap_{E\in\mathcal{E}} E^c\Bigg) > 0.$$

Proof. In this proof, we use the fact that $1-x \geq e^{-\alpha x}$ for $x \leq \frac{1}{2}$ and $\alpha = 2\log 2$. Let $x_E = 2p_E$ for each $E \in \mathcal{E}$. We now check that condition (3.8) holds with this choice. So, note that

$$\begin{aligned} x_E \prod_{F:(E,F)\in E(G)} (1-x_F) &= 2p_E \prod_{F:(E,F)\in E(G)} (1-2p_F) \\ &\geq 2p_E \exp\Bigg(-2\alpha \sum_{F:(E,F)\in E(G)} p_F\Bigg) \geq 2p_E e^{-2\alpha/4} = p_E. \end{aligned} \tag{3.18}$$

So, condition (3.8) holds and

$$\mathbb{P}\Bigg(\bigcap_{E\in\mathcal{E}} E^c\Bigg) > 0.$$

□

The final form we will use is the weighted form so called as we can assign weights to each event. This form is useful for dealing with events which are very unlikely but which depend on a large number of events which are more likely to occur.

Corollary 3.6 (Weighted Form)**.** *Suppose that we are in the setting of Lemma 3.3 and that each $E \in \mathcal{E}$ is assigned a weight $w_E \in \mathbb{N}$ and that there exists $p \in [0, 1/4]$ such that*

$$p_E \leq p^{w_E} \qquad \textit{and} \qquad \sum_{F:(E,F)\in E(G)} (2p)^{w_F} \leq \frac{w_E}{2} \tag{3.19}$$

for each $E \in \mathcal{E}$. Then,

$$\mathbb{P}\Bigg(\bigcap_{E\in\mathcal{E}} E^c\Bigg) > 0.$$

Proof. Again we will use the fact that $1-x \geq e^{-\alpha x}$ for $x \leq \frac{1}{2}$ and $\alpha = 2\log 2$. Let $x_E = (2p_E)^{w_E}$ for each $E \in \mathcal{E}$. To check that condition (3.8) holds with this choice observe,

$$x_E \prod_{F:(E,F)\in E(G)} (1 - x_F) = (2p)^{w_E} \prod_{F:(E,F)\in E(G)} (1 - (2p)^{w_F})$$
$$\geq (2p)^{w_E} \exp\Bigg(- \alpha \sum_{F:(E,F)\in E(G)} (2p)^{w_F} \Bigg) \geq (2p)^{w_E} e^{-\alpha w_E/2} = p^{w_E} \geq p_E. \quad (3.20)$$

So, condition (3.8) holds and

$$\mathbb{P}\Bigg(\bigcap_{E\in\mathcal{E}} E^c \Bigg) > 0.$$

□

CHAPTER 4

RAINBOW PERFECT MATCHINGS

4.1 Introduction

An $n \times n$ array of symbols where each symbol occurs precisely once in each row and column is called a *Latin square* of order n. A *partial transversal* of size k in a Latin square is a set of cells, including at most one from each row and each column that contains k distinct symbols. The question of finding the largest transversal in an arbitrary Latin square has attracted considerable attention. There exist Latin squares, such as the addition table of $\mathbb{Z}_n$ for even n, whose largest transversal has size $n-1$ [40, 116]. It has been conjectured that this is the worst case.

Conjecture 4.1 (Ryser, Brualdi, Stein [17, 103, 109]). *Every Latin square of order n contains a partial transversal of size at least $n-1$.*

The best known lower bound is due to Hatami and Shor [53], who showed that every Latin square of order n has a partial transversal of size $n - O(\log^2(n))$.

There is a one-to-one correspondence between Latin squares $L = (L_{ij})_{i,j\in[n]}$ of order n and proper edge colourings of the complete bipartite graph $K_{n,n}$ on $2n$ vertices; simply, assign colour L_{ij} to the edge a_ib_j, where $A = \{a_1, \dots, a_n\}$ and $B = \{b_1, \dots, b_n\}$ are the parts of $K_{n,n}$. Given a graph G and an edge colouring $\chi : E(G) \to \mathbb{N}$ of G, the subgraph $H \subseteq G$ is *rainbow* if for every $c \in \mathbb{N}$, $|\chi^{-1}(c) \cap E(H)| \leq 1$. Under the above correspondence, a partial transversal of size k in a Latin square is equivalent to a partial rainbow matching of size k.

One can extend the problem to edge colourings of $K_{n,n}$ that satisfy a milder condition. An edge colouring is *k-bounded* if $|\chi^{-1}(c)| \leq k$ for every $c \in \mathbb{N}$. Stein [109] conjectured that Conjecture 4.1 still holds for n-bounded edge colourings. This was very recently disproved by Pokrovskiy and Sudakov [100]. However, positive results can be obtained if the size of each colour class is small enough.

Theorem 4.2 (Erdős, Spencer [39])**.** *Let $K_{n,n}$ be the complete bipartite graph on $2n$ vertices, then any $(n-1)/16$-bounded edge colouring of $K_{n,n}$ contains a rainbow perfect matching.*

The goal of this chapter is to obtain a sparse version of Theorem 4.2. A balanced bipartite graph G contains a perfect matching if and only if G satisfies Hall's condition. However, it is easy to see that Hall's condition is not sufficient for the existence of a rainbow perfect matching if colour classes have linear size. For example, consider a graph consisting of a perfect matching which trivially satisfies Hall's Condition but has no rainbow perfect matching unless each colour class has size 1. Thus, we impose a stronger condition concerning the minimum degree of G. A *Dirac bipartite graph* on $2n$ vertices is a balanced bipartite graph with minimum degree at least $n/2$. The main result of this chapter shows the existence of rainbow perfect matchings in Dirac bipartite graphs.

Theorem 4.3. *There exist $\mu > 0$ and $n_0 \in \mathbb{N}$ such that if $n \geq n_0$ and G is a Dirac bipartite graph on $2n$ vertices, then any μn-bounded edge colouring of G contains a rainbow perfect matching.*

The proof of Theorem 4.3 combines probabilistic and extremal ingredients. The main tool used to provide the existence of a rainbow matching is a lopsided version of the Lovász Local Lemma, which is standard in this setting. One of the novelties of our approach is the estimation of conditional probabilities in the uniform distribution on the set of perfect matchings of a Dirac bipartite graph, via a switching argument (see

Section 4.2). However, this probability space often exhibits strong dependencies which limit the application of the local lemma.

In order to overcome this problem, in Section 4.3 we use a well-established dichotomy for Dirac bipartite graphs; either the graph has good expansion properties (*robust expander*) or the graph consists of two (possibly unbalanced) very dense bipartite graphs of order roughly n with few edges in-between (*extremal graph*). The notion of robust expanders was first introduced by Kühn et al. [79] in the context of Hamiltonian digraphs (see also [78]). A local lemma-based argument provides the existence of a rainbow perfect matching in robust expanders (Section 4.4). However, this argument cannot be applied directly to extremal graphs. In Section 4.5, we construct a rainbow perfect matching by selecting a partial matching in-between the two dense bipartite graphs that balances the remainder, followed by extending it into a rainbow perfect matching using similar arguments to the ones displayed previously. In Section 4.6 we combine these two results, concluding that any Dirac bipartite graph with a μn-bounded edge colouring contains a rainbow perfect matching.

Our result can be extended to a more general setting by slightly strengthening the minimum degree condition. A *system of conflicts* for $E(G)$ is a set $\mathcal{F}$ of unordered pairs of edges of G. If $\{e, f\} \in \mathcal{F}$ we say that e and f conflict and call $\{e, f\}$ a *conflict*. A system of conflicts $\mathcal{F}$ for $E(G)$ is k-*bounded* if for each $e \in E(G)$, there are at most k conflicts that contain e.

Given a graph G and a system of conflicts $\mathcal{F}$ for $E(G)$, the subgraph $H \subseteq G$ is $\mathcal{F}$-*conflict-free* if for each distinct $e, f \in E(H)$, we have $\{e, f\} \notin \mathcal{F}$.

Rainbow subgraphs correspond to conflict-free subgraphs of transitive systems of conflicts. Given an edge colouring χ of G, we define the system of conflicts $\mathcal{F}_\chi$ for $E(G)$ as follows

$$\mathcal{F}_\chi = \{\{e, f\} : e, f \in E(G) \text{ and } \chi(e) = \chi(f)\}$$

Note that χ is k-bounded if and only $\mathcal{F}_\chi$ is $(k-1)$-bounded.

We obtain an asymptotic version of Theorem 4.3 for bounded systems of conflicts.

Theorem 4.4. *For all $\varepsilon > 0$ there exist $\mu > 0$ and $n_0 \in \mathbb{N}$ such that if $n \geq n_0$ and G is a balanced bipartite graph on $2n$ vertices with minimum degree $\delta(G) \geq (1/2 + \varepsilon)n$, then any μn-bounded system of conflicts $\mathcal{F}$ for $E(G)$ contains a conflict-free perfect matching.*

Theorem 4.4 follows as a corollary of the proof of Theorem 4.3 for robust expanders (Section 4.6). Section 4.7 contains two applications of Theorem 4.3 and Theorem 4.4, providing the existence of rainbow Δ-factors in Dirac graphs and of rainbow spanning subgraphs with bounded maximum degree in graphs with large minimum degree. We conclude the chapter in Section 4.8 with further remarks and some open questions.

4.2 Switching over 6-cycles

Our main tool to find conflict-free matchings is the p-Lopsided form of the Lovász Local Lemma (Corollary 3.4.) We will use the following bound on the probability which we may deduce from the choice of the x_i in the proof of this corollary. If $\mathcal{E}$ is the family of events which satisfy the assumptions of Corollary 3.4, then

$$\mathbb{P}\Bigg(\bigcap_{E \in \mathcal{E}} E^c\Bigg) \geq (1 - 2p)^{|\mathcal{E}|}.$$

The following notion will play a central role in showing the existence of conflict-free perfect matchings.

Definition. Let $G = (A \cup B, E)$ be a balanced bipartite graph on $2n$ vertices with at least one perfect matching. Suppose M is a perfect matching of G and let $x = a_1b_1 \in M$. An edge $y = ab \in E(G)$ is (x, M)-*switchable* if $y \notin M$ and the 6-cycle $a_1b_1a_2bab_2$ is a subgraph of G, where $a_2b, ab_2 \in M$.

The existence of many switchable edges in every perfect matching suffices to find a conflict-free perfect matching.

Lemma 4.5. *Suppose that $1/n \ll \mu \ll \gamma \leq 1$ where $n \in \mathbb{N}$. Let $G = (A \cup B, E)$ be a balanced bipartite graph on $2n$ vertices with at least one perfect matching. Suppose that for every perfect matching M of G and for every $x = a_1b_1 \in M$ there are at least γn^2 edges of G that are (x, M)-switchable. Given a μn-bounded system of conflicts for $E(G)$, the probability that a uniformly random perfect matching of G is conflict-free is at least $e^{-\mu^{1/2}n}$.*

Proof. Let $\Omega = \Omega(G)$ be the set of perfect matchings of G equipped with the uniform distribution. By assumption, note that $\Omega \neq \emptyset$. Let $M \in \Omega$ be a perfect matching chosen uniformly at random. Let $\mathcal{F}$ be a μn-bounded system of conflicts for $E(G)$.

For each unordered pair of edges $x, y \in E(G)$ let $E(x, y) = \{x, y \in M\}$ be the event that both x and y are simultaneously in M. Define

$$Q = \{\{x, y\} \in \mathcal{F} :\ x, y \text{ non-incident}\}\ ,$$

and let $q = |Q|$. Consider the collection of events $\mathcal{E} = \{E(x, y) : \{x, y\} \in Q\}$.

Write $\mathcal{E} = \{E_i : i \in [q]\}$ and let $\mathcal{H}$ be the graph with vertex set $[q]$ where $i, j \in [q]$ are adjacent if and only if the subgraph of G that is spanned by the set of edges $\{x, y, w, z\}$ is not a matching, where $E_i = E(x, y)$ and $E_j = E(w, z)$.

Observe that given $\{x, y\} \in Q$, there are at most $4n$ ways to choose an edge $w \in E(G)$ that is incident either to x or to y, and at most μn ways to choose an edge $z \in E(G)$ with $\{w, z\} \in \mathcal{F}$. Hence, the maximum degree in $\mathcal{H}$ is at most $d := 4\mu n^2$.

Our goal is to show that $\mathcal{H}$ is a p-dependency graph for $\mathcal{E}$, for a suitably small $p > 0$. Given $i \in [q]$ and $S \subseteq [q] \setminus N_{\mathcal{H}}[i]$ with $\mathbb{P}(\cap_{j \in S} {E_j}^c) > 0$, it suffices to show that (3.7) holds.

Let $E_i = E(x, y)$. We say that a perfect matching is S-*good* if it belongs to $\cap_{j \in S} E_j^c$. Since $\mathbb{P}(\cap_{j \in S} {E_j}^c) > 0$, there is at least one S-good perfect matching. Let $\mathcal{M} = \mathcal{M}(S)$ be the set of S-good perfect matchings and let $\mathcal{M}_0 \subseteq \mathcal{M}$ be the set of S-good perfect matchings that contain both x and y.

Construct an auxiliary bipartite graph $\mathcal{G} = (\mathcal{M}_0, \mathcal{M} \setminus \mathcal{M}_0, E(\mathcal{G}))$, where $M_0 \in \mathcal{M}_0$

and $M \in \mathcal{M}$ are adjacent (i.e. $M_0M \in E(\mathcal{G})$) if there exist edges $x_1, x_2, y_1, y_2 \in M_0$ and $x_3, x_4, x_5, y_3, y_4, y_5 \in M$ such that $x, x_3, x_2, x_5, x_1, x_4$ and $y, y_3, y_2, y_5, y_1, y_4$ are vertex disjoint 6-cycles contained in G (see Figure 4.1).

By double-counting the edges of $\mathcal{G}$, we obtain

$$\delta(\mathcal{M}_0)|\mathcal{M}_0| \leq |E(\mathcal{G})| \leq \Delta(\mathcal{M} \setminus \mathcal{M}_0)|\mathcal{M} \setminus \mathcal{M}_0| \,,$$

from which we deduce,

$$\mathbb{P}(E_i| \cap_{j\in S} E_j^c) = \frac{|\mathcal{M}_0|}{|\mathcal{M}|} \leq \frac{|\mathcal{M}_0|}{|\mathcal{M} \setminus \mathcal{M}_0|} \leq \frac{\Delta(\mathcal{M} \setminus \mathcal{M}_0)}{\delta(\mathcal{M}_0)} \,. \tag{4.1}$$

So, in order to prove (3.8) we need to bound $\Delta(\mathcal{M} \setminus \mathcal{M}_0)$ from above and $\delta(\mathcal{M}_0)$ from below.

We first bound $\Delta(\mathcal{M} \setminus \mathcal{M}_0)$ from above. Fix $M \in \mathcal{M} \setminus \mathcal{M}_0$ and let us count the number of 6-cycles of the form $xx_3x_2x_5x_1x_4$, with $x_3, x_4, x_5 \in M$. Since $x_5 \in M$ is not incident with x, once we have chosen x_5 the 6-cycle is completely determined, as the edges x_3 and x_4 are the ones in M that are incident to both endpoints of x. There are at most $|M| = n$ choices for x_5, so there are at most n 6-cycles containing x. Similarly there are at most n 6-cycles containing y. It follows that $\Delta(\mathcal{M} \setminus \mathcal{M}_0) \leq n^2$.

In order to bound $\delta(\mathcal{M}_0)$ from below, fix $M_0 \in \mathcal{M}_0$. Note here that not all pairs of disjoint 6-cycles containing x and y, respectively, will generate an edge in $\mathcal{G}$ as it may be

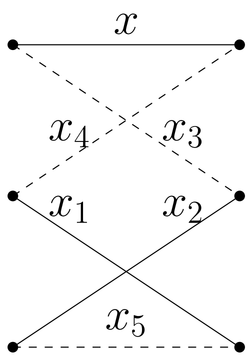


Figure 4.1: Switching for edge x.
x_5 is (x, M)-switchable.

that the perfect matching obtained by switching over the cycles is not S-good.

Define for $z \in \mathcal{M}_0$,

$$F_{\mathcal{M}_0}(z) = \{z' \in E(G) : z' \text{ is } (z, M_0)\text{-switchable and } \{w, z'\} \notin Q \text{ for all } w \in E(M_0)\} .$$

Let $F^*_{\mathcal{M}_0}(x) \subseteq F_{\mathcal{M}_0}(x)$ be the subset of edges that are not incident with x or y. By assumption, there are at least γn^2 edges that are (x, M_0)-switchable, from which at most μn^2 have conflicts with edges in M_0 and at most $2n$ are incident to y, implying $|F^*_{\mathcal{M}_0}(x)| \geq \gamma n^2/2$. Each edge $x_5 \in F^*_{\mathcal{M}_0}(x)$ uniquely determines a 6-cycle whose switching gives rise to a perfect matching. We claim that this matching is S-good. By adding x_3, x_4 and x_5, we can only create conflicts which use one of these edges. By definition of S, if $j \in S$, then the two edges defining E_j are not incident with x. Thus, x_3 and x_4 cannot create any conflict. Moreover, by the properties of $F_{\mathcal{M}_0}(x)$, x_5 does not conflict with the edges in M_0, so it cannot create any conflict. Given a choice of a 6-cycle of the form $xx_3x_2x_5x_1x_4$, let $F^{**}_{\mathcal{M}_0}(y) \subseteq F_{\mathcal{M}_0}(y)$ be the subset of edges that are not incident to the vertices of the fixed 6-cycle. Similarly as before, $|F^{**}_{\mathcal{M}_0}(y)| \geq \gamma n^2/2$ and each edge $y_5 \in F^{**}_{\mathcal{M}_0}(y)$ gives rise to a 6-cycle whose switching preserves the S-good condition. As for every choice of 6-cycle to switch out x and each choice of 6-cycle to switch out y we obtain an edge adjacent to M_0, we conclude that $\delta(\mathcal{M}_0) \geq \gamma^2 n^4/4$.

Substituting into (4.1), we obtain the desired bound

$$\mathbb{P}(E_i | \cap_{j \in S} E_j^c) \leq \frac{4}{\gamma^2 n^2} =: p .$$

Note that $|\mathcal{E}| \leq \mu n^3$. Since, $4pd = \frac{64\mu}{\gamma^2} \leq 1$, by the symmetric form of the p-Lopsided Lovász Local Lemma (Corollary 3.4), the probability that a uniformly random perfect matching is conflict-free is

$$\mathbb{P}(\cap_{E \in \mathcal{E}} E^c) \geq \left(1 - \frac{8}{\gamma^2 n^2}\right)^{\mu n^3} \geq e^{-\mu^{1/2} n} ,$$

and the lemma follows. $\square$

4.3 Dichotomy

In order to apply Lemma 4.5, we need to show that for every edge and every perfect matching containing it, there exist many switchable edges. However, this statement is not true for every Dirac bipartite graph. For instance, consider the graph G where $n = 2m+1$, $A = A_1 \cup A_2$ and $B = B_1 \cup B_2$ with $|A_1| = |B_2| = m$, and where $G[A_1, B_1]$ and $G[A_2, B_2]$ induce complete bipartite graphs and $G[A_2, B_1]$ induces a perfect matching. Clearly, G is a Dirac bipartite graph. However, for every edge in $x \in E(A_2, B_1)$, and independently of the choice of M containing x, there are at most m edges that are (x, M)-switchable as any such edge lies in $E(A_2, B_1)$. That any such edge lies in $E(A_2, B_1)$ follows from the fact that every perfect matching of G must contain precisely one edge from $E(A_2, B_1)$ and so if we remove this unique edge in a switching we must add back a different $A_2 - B_1$ edge such that the resulting subgraph is still a perfect matching of G.

Our proof proceeds by splitting the class of Dirac bipartite graphs into two cases: Robust Expanders, where we show the existence of many switchable edges, and Extremal Graphs, where we proceed carefully to handle the edges that produce a small number of switchings.

For $0 < \nu < 1$ and $X \subseteq V(G)$, the *ν-robust neighbourhood* of X in G is defined as

$$RN_\nu(X) := \{v \in V(G) : |N_G(v) \cap X| \geq \nu n\} \ .$$

Definition. Let $0 < \nu \leq \tau < 1$. A balanced bipartite graph $G = (A \cup B, E)$ on $2n$ vertices is a *bipartite robust (ν, τ)-expander* if for every set $X \subseteq V(G)$ with $\tau n \leq |X| \leq (1-\tau)n$ and either $X \subseteq A$ or $X \subseteq B$, we have

$$|RN_\nu(X)| \geq |X| + \nu n \ .$$

For a bipartite graph $G = (A \cup B, E)$ if $X \subseteq A$ and $Y \subseteq B$, we let $E(X,Y) = \{xy \in E(G) \colon x \in X, y \in Y\}$ and $e(X,Y) = |E(X,Y)|$.

Definition. Let $0 < \varepsilon < 1$. A balanced bipartite graph $G = (A \cup B, E)$ on $2n$ vertices is an *ε-extremal graph* if there exist partitions $A = A_1 \cup A_2$ and $B = B_1 \cup B_2$ such that the following properties are satisfied:

(P1) $||A_1| - |A_2|| \leq \varepsilon n$;

(P2) $||B_1| - |B_2|| \leq \varepsilon n$;

(P3) $e(A_1, B_2) \leq \varepsilon n^2$.

The following result establishes a dichotomy between these classes. Similar ideas have already appeared in previous work on Dirac graphs [32, 72, 74]. The proof follows the lines of the previous approaches but we include it here for the sake of completeness.

Theorem 4.6. *Suppose that $1/n \ll \nu \ll \varepsilon \ll \tau \ll 1$ where $n \in \mathbb{N}$. Let $G = (A \cup B, E)$ be a Dirac bipartite graph on $2n$ vertices. Then one of the following holds:*

i) G is a bipartite robust (ν, τ)-expander;

ii) G is an ε-extremal graph.

Proof. Let $0 < \delta < 1$ be such that $\nu \ll \delta \ll \varepsilon$. Suppose that G is not a bipartite robust (ν, τ)-expander. Thus, we may assume without loss of generality that there exists a set $X \subseteq A$ with $\tau n \leq |X| \leq (1-\tau)n$ and such that $|RN_\nu(X)| < |X| + \nu n$. We split the argument into three possible cases:

Case 1: $\tau n \leq |X| \leq \frac{n}{2} - \delta n$.

Since $e(X, N(X)) \geq |X|(n/2)$, we reach a contradiction:

$$e(X, N(X)) \leq |X||RN_\nu(X)| + \nu n^2 \leq |X| \left(\frac{n}{2} - (\delta - \nu)n\right) + \nu n^2 < |X|\frac{n}{2} \leq e(X, N(X)) \,,$$

where the second last inequality holds because $\nu \ll \delta, \tau$.

Case 2: $\frac{n}{2} - \delta n \leq |X| \leq \frac{n}{2} + \delta n$.

Define $A_1 = X$, $A_2 = A \backslash X$, $B_1 = RN_\nu(X)$, $B_2 = B \backslash RN_\nu(X)$. Note that $e(A_1, B_2) \leq \nu n|B_2| \leq \varepsilon n^2$; thus, (P3) holds. Now, (P1) and (P2) follow immediately by the conditions on $|X|$ and $|RN_\nu(X)|$. Hence G is an ε-extremal graph.

Case 3: $\frac{n}{2} + \delta n \leq |X| \leq (1-\tau)n$.

Each vertex in B has at least δn neighbours in X. So $RN_\nu(X) = B$. Hence, $|RN_\nu(X)| = n \geq |X| + \nu n$, a contradiction with the choice of X. □

4.4 Robust Expanders

As we will show below, the robust expansion property yields many (x, M)-switchable edges independently of the choice of x and M. Thus, Lemma 4.5 can be directly applied to obtain the existence of a conflict-free perfect matching in robust expanders.

Lemma 4.7. *Suppose $1/n \ll \gamma \ll \nu \ll \tau \ll 1$ where $n \in \mathbb{N}$. Let $G = (A \cup B, E)$ be a bipartite robust (ν, τ)-expander on $2n$ vertices with minimum degree at least $n/2$. Let M be a perfect matching of G and let $x \in M$. Then, there are at least γn^2 edges of G that are (x, M)-switchable.*

Proof. Let $f_M : A \to B$ be a bijective map defined as $f(a) = b$ if and only if $ab \in M$. Given $x = a_1b_1 \in M$ and an edge $y \notin M$ not incident to x, there is at most one 6-cycle in G that uses x, y and any two of the other edges in M. Also, note that y is (x, M)-switchable if and only if there is such a 6-cycle. Therefore, to count the number of (x, M)-switchable edges y, we count the number of 6-cycles containing x and any two of the other edges in M.

We will construct the 6-cycle by the sequence of vertices $a_1b_1a_2b_3a_3b_2$, where we have $a_2b_3, a_3b_2 \in M$. To bound from below the number of ways to choose the 6-cycle, we compute a lower bound on the number of choices for a_3 and b_3. Select,

$$a_3 \in RN_\nu(f_M(N(b_1))) \cap f_M^{-1}(N(a_1) \setminus \{b_1\}) ,$$

and then, $b_2 = f_M(a_3)$. Given the choice of a_3, select

$$b_3 \in f_M(N(b_1) \setminus \{a_1, a_3\}) \cap N(a_3) ,$$

and let $a_2 = f_M^{-1}(b_3)$. Recall that the minimum degree is at least $n/2$. As G is a bipartite robust (ν, τ)-expander, $|RN_\nu(f_M(N(b_1)))| \geq \frac{n}{2} + \nu n$, which implies that there are at least $\nu n - 1$ choices for a_3. Again, by the expansion properties of G, a_3 has at least νn neighbours in $f_M(N(b_1))$, so there are at least $\nu n - 2$ choices for b_3. In total, there are at least γn^2 choices of 6-cycles, $a_1b_1a_2b_3a_3b_2$, or equivalently, γn^2 edges $y = a_3b_3 \in E(G)$ that are (x, M)-switchable. $\square$

We can combine Lemma 4.5 and Lemma 4.7 together to conclude.

Corollary 4.8. *Suppose $1/n \ll \mu \ll \nu \ll \tau \ll 1$ where $n \in \mathbb{N}$. Let $G = (A \cup B, E)$ be a bipartite robust (ν, τ)-expander on $2n$ vertices with minimum degree at least $n/2$. Then, any μn-bounded system of conflicts for $E(G)$ contains a conflict-free perfect matching.*

4.5 Extremal Graphs

In this section we study the existence of rainbow perfect matchings for extremal graphs.

The example displayed at the beginning of Section 4.3 suggests that extremal graphs have special edges that are difficult to switch; namely, the ones between A_1 and B_2. Since the partitions $A = (A_1, A_2)$ and $B = (B_1, B_2)$ can be unbalanced, it may be unavoidable to select edges in $E(A_1, B_2)$ in a perfect matching of G. In fact, we may have to choose linearly many such edges.

A greedy approach for choosing the edges in $E(A_1, B_2)$ is likely to fail. By the properties of the edge colouring, the graph may contain vertices that only have a constant number of colours in the edges incident to them. If one selects a partial matching M^* in $E(A_1, B_2)$ and removes all the edges that have a colour in M^*, vertices that have few colours on their incident edges are likely to become isolated.

The way to handle this problem is given by Lemma 4.11, which shows that there is a way to select a rainbow partial matching M^* in $E(A_1, B_2)$ such that $|A_i \setminus V(M^*)| = |B_i \setminus V(M^*)|$, for $i \in \{1, 2\}$, and such that the degrees in the subgraph obtained after removing the colours in M^* are similar to the ones in the original graph.

4.5.1 A technical lemma

The core of the proof of Lemma 4.11 is a technical lemma that we present in this section.

We will be dealing both with multisets and with sets. We adopt the convention that a multiset will be defined with double brackets ($\{\{\ldots\}\}$) and a set with single brackets. For a multiset C and $k \in \mathbb{N}$, we denote by $m(k, C)$ the multiplicity of k in C. We define the operators $\cap^+$ and $\setminus^+$ both taking a multiset and a set and returning a multiset as follows: if A is a multiset and B is a set,

$$A \cap^+ B := \{\{x \in A : x \in B\}\} \qquad\qquad A \setminus^+ B := A \setminus (A \cap^+ B)$$

where $\setminus$ is the standard multiset difference. That is, $A \cap^+ B$ and $A \setminus^+ B$ are the multisets with multiplicity functions, $m(k, A \cap^+ B) = m(k, A) I(k \in B)$ and $m(k, A \setminus^+ B) = m(k, A) - m(k, A \cap^+ B)$.

Lemma 4.9. *Suppose that $1/N \ll \mu \ll \nu, 1/\alpha \ll \eta \leq 1$ where $N \in \mathbb{N}$. Let $C_1, \ldots, C_N$ be multisets of elements of $\mathbb{N}$ such that:*

(B1) $\nu N \leq |C_i| \leq N$, for every $i \in [N]$;

(B2) $\sum_{i=1}^{N} m(k, C_i) \leq \mu N$, for every $k \in \mathbb{N}$.

Let $\ell \in \mathbb{N}$ with $1 \leq \ell \ll \nu N$ and $\alpha\ell \in \mathbb{N}$. Let $U \subseteq \mathbb{N}$ be a set with $|U| = \alpha\ell$. Then, there exists $T \subseteq U$ such that:

(T1) $|T| \geq \ell$;

(T2) $|C_i \setminus^+ T| \geq (1-\eta)|C_i|$, *for every* $i \in [N]$.

Proof. If $\ell \leq 2\alpha$, then let T be an arbitrary subset of U of size ℓ. Since for every $i \in [N]$ and as $\mu \ll \eta, \nu$, we have $\mu N\ell \leq 2\mu\alpha N \leq \nu\eta N \leq \eta|C_i|$, $(T2)$ clearly holds. Throughout the proof we will assume that $\ell \geq 2\alpha$.

Let $0 < \varepsilon < 1$ be such that $\mu \ll \varepsilon \ll \nu, 1/\alpha$ and let $m_* := \frac{\varepsilon}{10\alpha^2} \cdot \frac{N}{\log N}$ and let $s := \log(\mu N/m_*)$. For every $i \in [N]$ and every $j \in [s]$, define the (multi)sets

$$P_i^j = \{\{k \in C_i : 2^{-j}\mu N \leq m(k, C_i) \leq 2^{-(j-1)}\mu N\}\},$$
$$S_i^j = \{k \in P_i^j\}.$$

Further, define

$$P_i = \cup_{j\in[s]} P_i^j,$$
$$S_i = \cup_{j\in[s]} S_i^j,$$
$$Q_i = C_i \setminus P_i.$$

Note that for every $k \in Q_i$, $m(k, C_i) \leq m_*$. Let $p_i^j = |P_i^j|$, $s_i^j = |S_i^j|$, $p_i = |P_i|$, $s_i = |S_i|$, $q_i = |Q_i|$, $c_i = |C_i|$. Then, these parameters satisfy

$$2^{-j}\mu N s_i^j \leq p_i^j \leq 2^{-(j-1)}\mu N s_i^j, \tag{4.2}$$
$$\sum_{j\in[s]} p_i^j = p_i, \tag{4.3}$$
$$p_i + q_i = c_i.$$

Let $T_0 \subseteq U$ be a random subset of U obtained by including each element of U independently at random with probability $\delta := 3\alpha^{-1}$. Note that $\mathbb{E}(|T_0|) = 3\ell$.

Claim. With probability $1 - o_N(1)$, for every $i \in [N]$, $|Q_i \setminus^+ T_0| \geq (1 - 4\alpha^{-1})q_i - \varepsilon N$.

Proof. Fix $i \in [N]$. If $q_i \leq \varepsilon N$ the statement is clearly true. So we may assume that $q_i \geq \varepsilon N$. For each $k \in Q_i$, define $m_k := m(k, C_i)$.

Then,

$$\sum_{k \in Q_i} m_k^2 \leq m_* \sum_{k \in Q_i} m_k = m_* q_i \leq \frac{\varepsilon N}{10\alpha^2 \log N} \cdot q_i. \tag{4.4}$$

Let $X_i = |Q_i \cap^+ T_0|$ and note that $\mathbb{E}(X_i) \leq \delta q_i$. By Azuma's Inequality (see e.g. [90]) with m_k satisfying (4.4) and the fact that $q_i \geq \varepsilon N$,

$$\mathbb{P}(X_i - \mathbb{E}(X_i) \geq \alpha^{-1} q_i) \leq 2\exp\left(\frac{-q_i^2}{2\alpha^2 \sum_{k \in Q_i} m_k^2}\right) \leq 2\exp\left(-5\frac{q_i \log N}{\varepsilon N}\right) \leq N^{-5}\ .$$

So, with probability $1 - o_N(1)$, for every $i \in [N]$, if $q_i \geq \varepsilon N$, then

$$|Q_i \setminus^+ T_0| \geq (1 - \alpha^{-1})q_i - \mathbb{E}(X_i) = (1 - 4\alpha^{-1})q_i \geq (1 - 4\alpha^{-1})q_i - \varepsilon N\ .$$

□

We now consider the sets P_i. For $\rho > 0$, a pair (i, j) is *ρ-dense* if $s_i^j \geq 2^{(j-1)/2}\rho$. Let R_i be the set of pairs (i, j) that are $\mu^{-1/2}$-dense. The contribution of non-dense pairs is negligible; using (4.2), we have

$$\sum_{j \notin R_i} p_i^j \leq \mu N \sum_{j \notin R_i} 2^{-(j-1)} s_i^j \leq \mu^{1/2} N \sum_{j \notin R_i} 2^{-(j-1)/2} \leq \mu^{1/3} N\ . \tag{4.5}$$

We say that $i \in [N]$ is *susceptible* if $|C_i \cap^+ U| \geq \eta |C_i|$. Let $D = \{i \in [N] : i \text{ is susceptible}\}$. Note that (T2) is satisfied for every $i \notin D$, as we have

$$|C_i \setminus^+ T| \geq |C_i \setminus^+ U| = |C_i| - |C_i \cap^+ U| \geq (1 - \eta)|C_i|\ .$$

Since $|C_i| \geq \nu N$, we can bound the size of D as follows

$$|D| \leq \alpha\ell \cdot \mu N/(\eta\nu N) \leq \ell\ . \tag{4.6}$$

Finally, for every $S \subseteq \mathbb{N}$ and $j \in [s]$ we say that $i \in [N]$ is *j-activated* by S if $|S_i^j \cap S| \geq 2\delta s_i^j$.

Consider the set $T \subseteq T_0$ defined as follows: we let T be a copy of T_0 where for each $i \in D$ and $j \in [s]$ we remove S_i^j if

i) i is j-activated by T_0, and

ii) $j \in R_i$ (i.e., (i,j) is $\mu^{-1/2}$-dense).

Observe that by removing elements from T_0 we only increase the size of $Q_i \setminus^+ T_0$. From the construction of T_0 and using (4.2) twice, it follows that for each $i \in D$, $j \in R_i$, we have

$$|P_i^j \cap^+ T| \leq \mu N 2^{-(j-1)} |S_i^j \cap T| \leq \mu N 2^{-(j-1)} \cdot 2\delta s_i^j \leq 4\delta p_i^j.$$

By combining this with (4.5), we obtain

$$\begin{aligned} |P_i \cap^+ T| &= \sum_{j\in[s]} |P_i^j \cap^+ T| = \sum_{j\in R_i} |P_i^j \cap^+ T| + \sum_{j\notin R_i} |P_i^j \cap^+ T| \\ &\leq 4\delta \sum_{j\in R_i} p_i^j + \sum_{j\notin R_i} p_i^j \leq 4\delta p_i + \mu^{1/3} N. \end{aligned}$$

Therefore, with probability $1-o_N(1)$, condition $(T2)$ is satisfied; that is, for every $i \in [N]$,

$$\begin{aligned} |C_i \setminus^+ T| &= |P_i \setminus^+ T| + |Q_i \setminus^+ T| \\ &\geq |P_i \setminus^+ T| + |Q_i \setminus^+ T_0| \\ &\geq (1-4\delta)p_i - \mu^{1/3}N + (1-4\alpha^{-1})q_i - \varepsilon N \geq (1-\eta)|C_i|. \end{aligned}$$

In order to conclude the proof of the lemma, it suffices to show that condition $(T1)$ holds with positive probability, from where we will deduce the existence of the desired set.

Claim. With probability at least $\frac{9}{10}$, we have $|T| \geq |T_0| - \ell$.

Proof. Since $|S_i^j \cap T_0|$ is stochastically dominated by a binomial random variable with parameters s_i^j and δ (there might be elements of S_i^j that are not in U), we can use Chernoff's inequality (see e.g. Corollary 2.3 in [58]) to show that

$$\mathbb{P}(i \text{ is } j\text{-activated}) \leq 2e^{-\frac{\delta s_i^j}{3}} .$$

If $j \in R_i$, then $s_i^j \geq 2^{(j-1)/2}\mu^{-1/2}$. Thus, $e^{-\frac{\delta s_i^j}{3}} \leq e^{-\frac{\delta 2^{(j-1)/2}}{3\mu^{1/2}}} \leq \mu^2 2^{-j}$. Hence, for $j \in R_i$

$$\mathbb{P}(i \text{ is } j\text{-activated}) \leq \mu^2 2^{-j} . \tag{4.7}$$

Recall the following inequality which follows from (4.2) and (4.3),

$$\sum_{j\in[s]} 2^{-j} s_i^j \leq \mu^{-1} . \tag{4.8}$$

Define the following random variable

$$Y := |T_0 \setminus T| \leq \sum_{i\in D} \sum_{j\in R_i} s_i^j \mathbb{1}(i \text{ is } j\text{-activated}) .$$

Note that the sets D and R_i are fully determined by $C_1, \ldots, C_N$. Then using (4.6), (4.7) and (4.8), it follows that

$$\begin{aligned}\mathbb{E}(Y) &\leq \sum_{i\in D} \sum_{j\in R_i} s_i^j \mathbb{P}(i \text{ is } j\text{-activated}) \leq \mu^2 \sum_{i\in D} \sum_{j\in R_i} 2^{-j} s_i^j \\ &\leq \mu^2 \sum_{i\in D} \sum_{j\in[s]} 2^{-j} s_i^j \leq \mu|D| \leq \frac{\ell}{10} .\end{aligned}$$

So, by Markov's inequality, $\mathbb{P}(Y \geq \ell) \leq 1/10$. □

Recall that $\ell \geq 2\alpha$. Since $|T_0|$ is distributed as a binomial random variable with parameters $\alpha\ell$ and δ, Chernoff's inequality implies that $\mathbb{P}(|T_0| \leq 2\ell) \leq 2e^{-\frac{\ell^2}{2\alpha\ell}} = 2e^{-\frac{\ell}{2\alpha}} \leq$

$\frac{2}{e}$. Thus, with positive probability, we have

$$|T| \geq 2\ell - \ell \geq \ell. \tag{4.9}$$

We conclude that there exists $T \subseteq U$ satisfying (T1) and (T2), concluding the proof of the lemma. $\square$

4.5.2 Superextremal graphs

We will use Lemma 4.9 to control the effect of colour deletions in the degrees of G. If degrees do not shrink significantly, the graphs $G_i = G[A_i, B_i]$, $i \in \{1, 2\}$, will still be fairly dense, and by applying Lemma 4.5 we will get the existence of a rainbow perfect matching.

However, the ε-extremal condition does not ensure that the graphs G_i have large minimum degree; that is, G_i is not necessarily Dirac. In this section we refine the notion of extremality and we obtain a partition where the degrees of each vertex within its part is controlled. Eventually, this will allow us to count the number of switchable edges.

Definition. Let $0 < \nu_1 \leq \nu_2 < 1$. A balanced bipartite graph $G = (A \cup B, E)$ on $2n$ vertices is a (ν_1, ν_2)-*superextremal graph* if there exist partitions $A = A_1 \cup A_2$ and $B = B_1 \cup B_2$ such that the following properties are satisfied for $i \in \{1, 2\}$:

(Q1) $e(v, B_i) \geq \frac{n}{2} - \nu_1 n$, for all but at most $\nu_1 n$ vertices $v \in A_i$;

(Q2) $e(v, B_i) \geq \nu_2 n$, for every $v \in A_i$;

(Q3) $e(v, A_i) \geq \frac{n}{2} - \nu_1 n$, for all but at most $\nu_1 n$ vertices $v \in B_i$;

(Q4) $e(v, A_i) \geq \nu_2 n$, for every $v \in B_i$;

(Q5) $||A_1| - |B_1||$, $||A_1| - |A_2|| \leq \nu_1 n$;

(Q6) $e(v, B_2) \leq \nu_2 n$, for every $v \in A_1$, unless $|A_1| = |B_1|$;

(Q7) $e(v, A_1) \le \nu_2 n$, for every $v \in B_2$, unless $|A_1| = |B_1|$;

(Q8) $|A_1| \ge |B_1|$;

(Q9) one of the following holds for $\ell := |A_1| - |B_1|$:

- $e(v, B_2) \ge \ell/2$, for every $v \in A_1$;
- $e(v, A_1) \ge \ell/2$, for every $v \in B_2$.

Lemma 4.10. *Suppose $1/n \ll \varepsilon \ll \nu_1 \ll \nu_2 \ll 1$ where $n \in \mathbb{N}$. Let $G = (A \cup B, E)$ be an ε-extremal Dirac bipartite graph on $2n$ vertices. Then, G is a (ν_1, ν_2)-superextremal graph.*

Proof. Since G is an ε-extremal graph, there exist partitions $A = A_1^1 \cup A_2^1$ and $B = B_1^1 \cup B_2^1$ satisfying $(P1)$, $(P2)$ and $(P3)$. Let $0 < \nu_3 \le \nu_4 < 1$ such that $\varepsilon \ll \nu_3 \ll \nu_1 \ll \nu_4 \ll \nu_2$ and define

$$\begin{aligned}
X_1^1 &= \left\{v \in A_1^1 : e(v, B_1^1) \le \frac{n}{2} - \nu_3 n\right\} & X_1^2 &= \left\{v \in A_1^1 : e(v, B_1^1) \le \frac{n}{4}\right\} \\
X_2^1 &= \left\{v \in A_2^1 : e(v, B_2^1) \le \frac{n}{2} - \nu_3 n\right\} & X_2^2 &= \left\{v \in A_2^1 : e(v, B_2^1) \le \frac{n}{4}\right\} \\
Y_1^1 &= \left\{v \in B_1^1 : e(v, A_1^1) \le \frac{n}{2} - \nu_3 n\right\} & Y_1^2 &= \left\{v \in B_1^1 : e(v, A_1^1) \le \frac{n}{4}\right\} \\
Y_2^1 &= \left\{v \in B_2^1 : e(v, A_2^1) \le \frac{n}{2} - \nu_3 n\right\} & Y_2^2 &= \left\{v \in B_2^1 : e(v, A_2^1) \le \frac{n}{4}\right\}
\end{aligned}$$

We double count edges to bound the size of these sets. Note that $e(A_1^1, B_1^1) \ge \frac{n}{2}|A_1^1| - \varepsilon n^2$ by counting from A_1^1. Alternately, we can also obtain that $e(A_1^1, B_1^1) \le |X_1^1|(\frac{n}{2} - \nu_3 n) + (|A_1^1| - |X_1^1|)|B_1^1|$. Combining these two inequalities yields

$$|X_1^1| \left(|B_1^1| - \frac{n}{2} + \nu_3 n\right) \le |A_1^1| \left(|B_1^1| - \frac{n}{2}\right) + \varepsilon n^2 \le 2\varepsilon n^2.$$

Observe that $|B_1^1| \ge \frac{n}{2} - \varepsilon n$ and so $|B_1^1| - \frac{n}{2} + \nu_3 n \ge \frac{\nu_3 n}{2}$. Therefore $|X_1^1| \le \nu_3 n$. Similarly, one can deduce that $|X_1^2| \le 9\varepsilon n$. Analogous computations lead to $|X_2^1|, |Y_1^1|, |Y_2^1| \le \nu_3 n$

and to $|X_2^2|, |Y_1^2|, |Y_2^2| \leq 9\varepsilon n$. Now, we define

$$A_1^2 = (A_1^1 \setminus X_1^2) \cup X_2^2 \qquad\qquad B_1^2 = (B_1^1 \setminus Y_1^2) \cup Y_2^2$$
$$A_2^2 = (A_2^1 \setminus X_2^2) \cup X_1^2 \qquad\qquad B_2^2 = (B_2^1 \setminus Y_2^2) \cup Y_1^2$$

Without loss of generality, $|A_1^2| \geq |B_1^2|$; otherwise we swap the labels of A_1^2 and A_2^2, and the labels of B_1^2 and B_2^2. By swapping the labels we lose control of $e(A_1^2, B_2^2)$. However, at this point, this condition is no longer needed, as we have a bound on the size of the sets X_i^j and Y_i^j, for $i, j \in \{1, 2\}$.

Let

$$X_1^3 = \{v \in A_1^2 : e(v, B_2^2) \geq \nu_4 n\} \qquad\qquad Y_2^3 = \{v \in B_2^2 : e(v, A_1^2) \geq \nu_4 n\}$$

If $|X_1^3| + |Y_2^3| \geq |A_1^2| - |B_1^2|$, choose $X_1^4 \subseteq X_1^3$ and $Y_2^4 \subseteq Y_2^3$ arbitrarily such that $|X_1^4| + |Y_2^4| = |A_1^2| - |B_1^2|$. Otherwise, let $X_1^4 = X_1^3$ and $Y_2^4 = Y_2^3$. Recall that, since G is ε-extremal, it satisfies $n/2 - \varepsilon n \leq |A_1^1|, |B_1^1| \leq n/2 + \varepsilon n$. Thus, we have

$$|X_1^4| \leq |A_1^2| - |B_1^2| \leq |A_1^1| - |B_1^1| + 9\varepsilon n + \nu_3 n \leq 10\varepsilon n + \nu_3 n \ .$$

and similarly for Y_2^4.

We define

$$A_1 = A_1^2 \setminus X_1^4 \qquad A_2 = A_2^2 \cup X_1^4 \qquad B_1 = B_1^2 \cup Y_2^4 \qquad B_2 = B_2^2 \setminus Y_2^4$$

We claim that the partitions $A = A_1 \cup A_2$ and $B = B_1 \cup B_2$ satisfy properties (Q1)-(Q9), and so, G is a (ν_1, ν_2)-superextremal graph.

Let us first check that property (Q1) is satisfied. Observe that all the vertices in A_1^1, excluding the ones in X_1^1, have degree at least $n/2 - \nu_3 n$. Since $|X_1^2| \leq 9\varepsilon n$, all the vertices in A_1^2 have degree at least $n/2 - \nu_3 n - 9\varepsilon n$ to B_1, excluding the ones in $X_1^1 \cup X_2^2$. Since

$|X_1^4| \leq 10\varepsilon n + \nu_3 n$, all the vertices in A_1 have degree at least $n/2 - 2\nu_3 n - 19\varepsilon n \geq n/2 - \nu_1 n$, excluding the ones in $X_1^1 \cup X_2^2$. Moreover, $|X_1^1 \cup X_2^2| \leq \nu_3 n + 9\varepsilon n \leq \nu_1 n$, so (Q1) follows. Similar arguments yield to properties (Q2)-(Q4) and (Q6)-(Q7).

Property (Q8) follows from the choice of X_1^4 and Y_2^4, since $|A_1| = |A_1^2| - |X_1^4| \geq |B_1^2| + |Y_2^4| = |B_1|$. Property (Q5) follows since $|A_1| - |B_1| \leq |A_1^2| - |B_1^2| \leq 20\varepsilon n \leq \nu_1 n$ (and since a similar computation bounds $||A_1| - |A_2||$).

Finally, Property (Q9) follows by noting that if $\ell = |A_1| - |B_1|$ (and $|A_2| = |B_2| - \ell$), then either $|B_1|$ or $|A_2|$ is at most $n/2 - \ell/2$ thus requiring minimum degree $\ell/2$ either from A_1 to B_2 or from B_2 to A_1, respectively.

□

4.5.3 Selecting a rainbow partial matching between parts

Given a superextremal graph G with partitions $A = A_1 \cup A_2$ and $B = B_1 \cup B_2$, in this section we will show the existence of a rainbow partial matching M^* in $G[A_1, B_2]$ of size $\ell = |A_1| - |B_1|$ such that the graph H resulting from removing all edges incident to M^* and all edges with colours that appear in M^*, has similar degrees as the graph G.

Lemma 4.11. *Suppose $1/n \ll \mu \ll \nu_1 \ll \nu_2 \ll \nu_3 \ll \eta_1 \ll 1$ where $n, \ell \in \mathbb{N}$. Let $G = (A \cup B, E)$ be a (ν_1, ν_3)-superextremal graph with partitions $A = A_1 \cup A_2$ and $B = B_1 \cup B_2$. Then, any μn-bounded edge colouring χ of G admits a rainbow matching M^* of size $\ell = |A_1| - |B_1|$ such that the following holds. Let $H = (A^H \cup B^H, E^H)$ be the graph where $A^H = A \setminus V(M^*)$, $B^H = B \setminus V(M^*)$ and*

$$E^H = \{x = ab \in E(G) : a, b \notin V(M^*), \chi(x) \notin \chi(E(M^*))\} .$$

Let $n_H := n - \ell$. Then, there exist partitions $A^H = A_1^H \cup A_2^H$ and $B^H = B_1^H \cup B_2^H$ that satisfy the following properties for $i \in \{1, 2\}$:

(R1) $e_H(v, B_i^H) \geq (1 - \eta_1)\frac{n_H}{2}$, for all but at most $\nu_1 n$ vertices $v \in A_i^H$;

(R2) $e_H(v, B_i^H) \geq \nu_2 n_H$, *for every* $v \in A_i^H$*;*

(R3) $e_H(v, A_i^H) \geq (1-\eta_1)\frac{n_H}{2}$, *for all but at most* $\nu_1 n$ *vertices* $v \in B_i^H$*;*

(R4) $e_H(v, A_i^H) \geq \nu_2 n_H$, *for every* $v \in B_i^H$*;*

(R5) $|A_1^H| = |B_1^H|$, $|A_2^H| = |B_2^H|$ *and* $|A_1^H| - |A_2^H| \leq \nu_1 n_H$.

Proof. We first greedily select a large rainbow matching in $G[A_1, B_2]$. Let $E_0 = E(A_1, B_2)$ and $M_0 = \emptyset$. By (Q5) and (Q8), note that $|E_0| \geq \frac{\ell}{2}(\frac{n}{2} - \nu_1 n)$. For every $i \geq 1$ and while $E_{i-1} \neq \emptyset$, we arbitrarily choose $x_i = a_i b_i \in E_{i-1}$ and define the graph M_i with $V(M_i) = V(M_{i-1}) \cup \{a_i, b_i\}$ and $E(M_i) = E(M_{i-1}) \cup \{x_i\}$. We let

$$E_i = \{x = ab \in E_{i-1} : \, a, b \notin V(M_i),\, \chi(x) \notin \chi(E(M_i))\} \, .$$

Since χ is μn-bounded, $|A_1| - |B_1| = \ell \geq 1$ and using (Q6)-(Q7), we have $|E_i| \geq |E_{i-1}| - (2\nu_3 + \mu)n$. Let $i^* = \lfloor \ell/(10\nu_3) \rfloor$. It follows that $E_i \neq \emptyset$, for every $0 \leq i \leq i^*$.

We now apply Lemma 4.9 with parameters $N = 2n$, $\alpha = i^*/\ell$, $\nu = \nu_3/2$, $\eta = \eta_1/2$ and $U = \{\chi(x) : x \in M_{i^*}\}$. For every $v \in A \cup B$, we choose $C_v = \{\{\chi(x) : v \in x\}\}$ to be the multiset of colours on edges incident with vertex v. By (Q2) and (Q4), we have $\nu N \leq |C_v| \leq N$, for each $v \in A \cup B$. As each edge has two endpoints and χ is μn-bounded, then $\sum_{v \in A \cup B} m(k, C_v) \leq 2\mu n = \mu N$. Hence, (B1) and (B2) hold.
Lemma 4.9 implies the existence of a set of colours $T \subseteq U$ of size ℓ satisfying (T1) and (T2). Let M^* be the subgraph of M_{i^*} induced by the colours in T. Then, M^* is a rainbow matching of size ℓ. It suffices to prove that H, as defined in the statement, satisfies (R1)-(R5).

For each $Z \in \{A, B\}$ and $i \in \{1, 2\}$, let $Z_i^H = Z_i \cap V(H)$ and $\bar{Z}_i^H$ be the other choice. Property (R5) follows since $|B_1^H| = |B_1| = |A_1| - \ell = |A_1^H|$ and using (Q5). Then, for every $v \in Z_i^H$, we have

$$e_H(v, \bar{Z}_i^H) \geq |C_v \setminus^+ T| - \ell \geq (1-\eta)|C_v| - \nu_1 n \geq \begin{cases} (1-\eta)(\frac{n}{2} - \nu_1 n) - \nu_1 n \geq (1-\eta_1)\frac{n_H}{2} & \text{if } v \text{ satisfies (Q1) or (Q3)} \\ (1-\eta)\nu_3 n - \nu_1 n \geq \nu_2 n_H & \text{if } v \text{ satisfies (Q2) or (Q4)} \end{cases}$$

Thus H satisfies (R1)-(R4), completing the proof. □

4.5.4 Completing the rainbow perfect matching

Consider the rainbow partial matching M^* and the graph H provided by Lemma 4.11. Note that H is vertex disjoint from M^* and has no edge with colour in $\chi(E(M^*))$. Thus, the union of any rainbow perfect matching of H and M^* will provide a rainbow perfect matching of G.

We will show that H satisfies the conditions of Lemma 4.5, to conclude the existence of a rainbow perfect matching there.

Of course, in order to have a rainbow perfect matching in H we need to ensure the existence of at least one perfect matching. We will use the Moon-Moser condition for the existence of Hamiltonian cycles in bipartite graphs to guarantee we can find a perfect matching.

Theorem 4.12. *(Moon, Moser [93]) Let $F = (R \cup S, E)$ be a balanced bipartite graph on $2m$ vertices with $R = \{r_1, \dots, r_m\}$ and $S = \{s_1, \dots, s_m\}$ that satisfies $d(r_1) \leq \dots \leq d(r_m)$ and $d(s_1) \leq \dots \leq d(s_m)$. Suppose that for every $1 \leq k \leq m/2$, we have $d(r_k) > k$ and $d(s_k) > k$. Then F has a Hamiltonian cycle.*

Lemma 4.13. *Suppose $1/n_H \ll \mu \ll \varepsilon \ll \nu_1 \ll \gamma \ll \nu_2 \ll \eta \ll 1$ where $n_H \in \mathbb{N}$. Let $H = (A^H \cup B^H, E^H)$ be a bipartite graph with $|A^H| = |B^H| = n_H$ and satisfying properties (R1)-(R5). Consider the subgraph $H_* = (A^H \cup B^H, E_*^H)$ of H with $E_*^H = E_{*,1}^H \cup E_{*,2}^H$ and,* *for* $i \in \{1,2\}$,

$$E^H_{*,i} = \{x = ab \in E(H) : \, a \in A_i, \, b \in B_i \text{ and } \max\{e_H(a,B_i), e_H(b,A_i)\} \geq (1-\eta)n_H/2\}$$

Then, H_* *has at least one perfect matching.*

Moreover, if M_* *is a perfect matching of* H_* *and* $x = a_1b_1 \in E(M_*)$*, then there are at least* γn_H^2 *edges of* H_* *that are* (x, M_*)*-switchable.*

Proof. It is easy to check that H_* has only two connected components. We will show that it is true in $H_1 = H_*[A^H_1, B^H_1]$ and the same argument also applies to $H_*[A^H_2, B^H_2]$. Note that H_1 is a balanced bipartite graph on $2m$ vertices for some $m \in (n_H(1/2 - \nu_1), n_H(1/2 + \nu_1))$.

We will use the Moon-Moser condition (Lemma 4.12) to show the existence of a Hamiltonian cycle in H_1. Let $A^H_1 = \{r_1, \ldots, r_m\}$ and $B^H_1 = \{s_1, \ldots, s_m\}$ with $d(r_1) \leq \cdots \leq d(r_m)$ and $d(s_1) \leq \cdots \leq d(s_m)$.

If $1 \leq k \leq 2\nu_1 n_H < 5\nu_1 m - 1$, then, by (R2), $d(r_k) \geq \nu_2 n_H \geq 5\nu_1 m > k$, so there is nothing to prove. If $2\nu_1 n_H \leq k \leq m/2$, then, by (R1), $d(r_k) \geq (1-\eta)m > k$. An identical argument works for s_k using (R3) and (R4). Thus we satisfy the Moon-Moser condition. So, H_1 has a Hamiltonian cycle, which implies the existence of a perfect matching.

Let M_* be a perfect matching of H_*. Consider the bijective map $f_{M^*} : A^H \to B^H$ defined as $f(a) = b$ if and only if $ab \in M_*$. Let $x = a_1b_1 \in M_*$, and, without loss of generality, assume that $a_1 \in A^H_1$, so $b_1 \in B^H_1$. In order to prove the second part of the lemma, we need to show that there are many edges $y = ab$ that are (x, M_*)-switchable.

Let $0 < \delta < 1$ such that $\gamma \ll \delta \ll \nu_2$. Observe that the minimum degree in H_* is at least $(\nu_2 - \nu_1)n_H \geq \delta m$. By construction, there is no pair of vertices both of degree less than $(1-\eta)m$ that are connected by an edge in H_*. Thus, without loss of generality, we may assume that $e_{H_*}(a_1, B^H_1) \geq \delta m$ and that $e_{H_*}(b_1, A^H_1) \geq (1-\eta)m$.

Since $|f^{-1}_{M_*}(N_{H_*}(a_1))| \geq \delta m$ and since there are at most $2\nu_1 m$ vertices of degree less than $(1-\eta)m$, there are at least $\delta m/2$ choices for $a \in f^{-1}_{M_*}(N_{H_*}(a_1) \setminus \{b_1\})$ that satisfy

$e_{H_*}(a, B_1^H) \geq (1-\eta)m$.

Fix such a vertex a and note that

$$e_{H_*}\left(a, B_1^H \setminus f_{M_*}(N_{H_*}(b_1) \setminus \{a_1, a\})\right) \leq |B_1^H| - (1-\eta)m + 2 \leq \eta m + 2 \ .$$

Therefore,

$$\begin{aligned} e_{H_*}\left(a, f_{M_*}(N_{H_*}(b_1) \setminus \{a_1, a\})\right) &= e_{H_*}(a, B_1^H) - e_{H_*}\left(a, B_1^H \setminus f_{M_*}(N_{H_*}(b_1) \setminus \{a_1, a\})\right) \\ &\geq (1-\eta)m - (\eta m + 2) \\ &\geq (1-3\eta)m \ . \end{aligned}$$

Thus, there are at least $(1-3\eta)m$ choices for $b \in f_{M_*}(N_{H_*}(b_1) \setminus \{a_1, a\})$ with $ab \in E(H_*)$. It follows that there are at least $(\delta m/2)(1-3\eta)m \geq \gamma n_H^2$ choices of an edge $y = ab \in E(H_*)$ such that there exists a 6-cycle that contains x, y and two other edges of M_*. We conclude that there are at least γn_H^2 edges of H_* that are (x, M_*)-switchable.

□

An application of Lemma 4.5, Lemma 4.11 and Lemma 4.13 yields the following immediate corollary.

Corollary 4.14. *Suppose $1/n \ll \mu \ll \varepsilon \ll 1$ where $n \in \mathbb{N}$. Let $G = (A \cup B, E)$ be an ε-extremal Dirac bipartite graph on $2n$ vertices. Then, any μn-bounded edge colouring of G contains a rainbow perfect matching.*

4.6 Proofs of Theorem 4.3 and Theorem 4.4

We finally prove our main theorems.

Proof of Theorem 4.3. Let G be a Dirac bipartite graph on $2n$ vertices and suppose $1/n \ll \mu \ll \varepsilon \ll \nu \ll \tau \ll 1$. Consider a μn-bounded edge colouring χ of G. By

Lemma 4.6, the graph G is either ε-extremal or a bipartite robust (ν, τ)-expander. If G is a bipartite robust (ν, τ)-expander, then G has a rainbow perfect matching by Corollary 4.8 with $\mathcal{F} = \mathcal{F}_\chi$. If G is an ε-extremal graph, then G has a rainbow perfect matching by Corollary 4.14. $\square$

Proof of Theorem 4.4. Let $G = (A \cup B, E)$ be a balanced bipartite graph on $2n$ vertices with minimum degree at least $(1/2 + \varepsilon)n$. Suppose that $1/n \ll \mu \ll \varepsilon \ll 1$. We will show that G is a bipartite robust $(\varepsilon/8, 1/4)$-expander. Let X be a subset of either A or B with $n/4 \leq |X| \leq 3n/4$; without loss of generality, we may assume that $X \subseteq A$. By the minimum degree condition we have $e(X, B) \geq (1/2 + \varepsilon)n|X|$ and, by the definition of robust neighbourhood, we have $e(X, B) \leq |X||RN_{\varepsilon/8}(X)| + \varepsilon n(n - |RN_{\varepsilon/8}(X)|)/8$. Combining these inequalities yields $|X||RN_{\varepsilon/8}(X)| + \varepsilon n^2/8 \geq (1/2+\varepsilon)n|X|$ and, as $|X| \geq n/4$, upon rearrangement, we have that $|RN_{\varepsilon/8}(X)| \geq (1/2+\varepsilon/2)n$. If $n/4 \leq |X| \leq n/2$, then $|RN_{\varepsilon/8}(X)| \geq |X| + \varepsilon n/8$ and we are done. If $n/2 \leq |X| \leq 3n/4$, by the minimum degree condition, each $v \in B$ has at least εn neighbours in X. Thus $RN_{\varepsilon/8}(X) = B$ and $|RN_{\varepsilon/8}(X)| = n \geq |X| + \varepsilon n/8$. So G is a bipartite robust $(\varepsilon/8, 1/4)$-expander. Corollary 4.8 completes the proof. $\square$

The following proposition shows that $\mu \leq 1/4$ (see Section 4.8 for a discussion).

Proposition 4.15. *For every $t \in \mathbb{N}$, there exists a Dirac bipartite graph G on $n = 4t(t+1)$ vertices and a $\left(\frac{t+1}{4t}\, n\right)$-bounded edge colouring of G such that G does not contain a rainbow perfect matching.*

Proof. Let $m = 2t$. Consider the bipartite graph $G = (A \cup B, E)$ constructed as follows. The vertex set is partitioned into $A = A_1 \cup A_2$ and $B = B_1 \cup B_2$, with

$$A_1 = \{A_1^1, \ldots, A_1^{m-1}\}$$
$$A_2 = \{A_2^1, \ldots, A_2^{m+1}\}$$
$$B_1 = \{B_1^1, \ldots, B_1^{m+1}\}$$
$$B_2 = \{B_2^1, \ldots, B_2^{m-1}\} \ ,$$

where $|A_k^i| = |B_k^i| = t+1$. The edge set of G is consists of two complete bipartite graphs induced by $G[A_1, B_1]$ and $G[A_2, B_2]$, and of $m+1$ smaller complete bipartite graphs induced by $G[A_2^i, B_1^i]$, for $i \in [m+1]$. Note that G is a Dirac bipartite graph.

Consider the edge colouring that assigns colour $c_{k,\ell}^{i,j}$ to the edges in $G[A_k^i, B_\ell^j]$. Since each set has size $t+1$, the colouring is $(t+1)^2 = \left(\frac{t+1}{4t}\, n\right)$-bounded.

Suppose that G admits a rainbow perfect matching M. Note that M contains at most $m+1$ edges in $G[A_2, B_1]$. Otherwise there exists $i \in [m+1]$ such that M contains two edges in $E[A_2^i, B_1^i]$, contradicting the fact that it is rainbow, since both edges have colour $c_{2,1}^{i,i}$. Since all the edges incident to A_1 are also incident to B_1, we must have $|A_1| \geq |B_1| - (m+1)$. However

$$|A_1| = (m-1)(t+1) = (m+1)(t+1) - 2(t+1) = |B_1| - (m+2),$$

a contradiction. We conclude that, G has no rainbow perfect matching.

□

4.7 Applications

In the following section we provide some applications of our main theorems on the existence of rainbow spanning subgraphs of graphs with large minimum degree that are not necessarily bipartite.

We first discuss the existence of rainbow Δ-factors in Dirac graphs for a wide range of Δ. Recall that a *Dirac graph* on n vertices is a graph with minimum degree at least $n/2$. The existence of $(n/2)$-factors in Dirac graphs was proved by Katerinis [61]. Our next result extends Theorem 4.3 to Δ-factors of Dirac graphs.

Theorem 4.16. *There exist $\mu > 0$ and $n_0 \in \mathbb{N}$ such that if $n \geq n_0$ then for every even $1 \leq \Delta \leq \mu n$ the following holds. Let G be a Dirac graph on n vertices, then any $(\mu n/\Delta)$-bounded colouring of G contains a rainbow Δ-factor.*

Note that this theorem is tight in its dependence on n and Δ as a Δ factor contains $n\Delta/2$ edges.

Proof. We construct an auxiliary bipartite graph $Q = (V(Q), E(Q))$ as follows. The vertex set is $V(Q) = A \cup B$, where $A = \{u_{v,i} : v \in V(G), 1 \leq i \leq \Delta/2\}$ and $B = \{u_{v,i} : v \in V(G), \Delta/2 < i \leq \Delta\}$. The edge set is defined as

$$E(Q) = \{u_{v,i}u_{w,j} : u_{v,i} \in A,\ u_{w,j} \in B \text{ and } vw \in E(G)\} .$$

Note that Q is a bipartite Dirac graph on $2N = \Delta n$ vertices. Let $\chi : E(G) \to \mathbb{N}$ be a μn-bounded edge colouring of G. Construct the edge colouring $\chi_Q : E(Q) \to \mathbb{N}$ defined by $\chi_Q(u_{v,i}u_{w,j}) = \chi(vw)$, for every $u_{v,i}u_{w,j} \in E(Q)$. Since $2 \cdot (\Delta/2)^2 \cdot \mu n/\Delta = \mu N$, the colouring is μN-bounded. Thus, by Theorem 4.3, Q has a rainbow perfect matching M.

Consider the subgraph $H = (V(H), E(H))$ of G with $V(H) = V(G)$ and edge set

$$E(H) = \{vw \in E(G) : \text{there exist } 1 \leq i \leq \Delta/2 < j \leq \Delta \text{ such that } u_{v,i}u_{w,j} \in E(M)\} .$$

We claim that H is a rainbow Δ-factor of G. Since $u_{v,i} \in V(Q)$ for every $i \in [\Delta]$ and M is a perfect matching of Q, we have $d_H(v) = \Delta$. Since $u_{v,i}u_{v,j} \notin E(Q)$ for every $i, j \in [\Delta]$, H has no self loops. Finally, since M is a rainbow perfect matching of Q, and by definition of the colouring χ, H has no multiple edges and each colour in χ appears at most once in M. Thus, H is a simple rainbow Δ-regular spanning subgraph of G. $\square$

Our second corollary concerns bipartite subgraphs of graphs with large minimum degree. Consider two graphs G and H on n vertices with $\Delta(H) \leq \Delta$. The Bollobás-Eldridge-Catlin conjecture [11, 19], states that if $\delta(G) \geq (1 - 1/(\Delta + 1))\, n - 1/(\Delta + 1)$, then G contains a copy of H. Sauer and Spencer [104] showed that the conjecture holds if $\delta(G) \geq (1 - 1/2\Delta)\, n - 1$. This result has been improved for large values of Δ [62]. The existence of rainbow copies of H in k-bounded edge colourings of K_n was studied in [14], provided that $k = O(n/\Delta^2)$. In [111], it was observed that similar techniques allow to replace K_n by a graph G with $\delta(G) \geq (1 - c/\Delta)\, n$, for a sufficiently small constant $c > 0$.

Our last result partially extends the result in [14] at the Sauer-Spencer minimum degree threshold.

Theorem 4.17. *For every $\gamma > 0$ there exists $\mu > 0$ such that for every $\Delta \in \mathbb{N}$ there exists $n_0 \in \mathbb{N}$ such that for every even $n \geq n_0$ the following holds. Let G be a graph on n vertices with $\delta(G) \geq (1 - 1/2\Delta + \gamma)\, n$ and let H be a balanced bipartite graph on n vertices with $\Delta(H) \leq \Delta$, then any proper $(\mu n/\Delta^2)$-bounded edge colouring of G contains a rainbow copy of H.*

Sudakov and Volec [111] showed that there exist a graph H with maximum degree at most Δ and a $3.9n/\Delta^2$-bounded edge colouring of K_n which does not contain a rainbow copy of H. Therefore this theorem is also tight up to constant factors.

Proof. By Lemma 2.3 in [5] there is a balanced bipartite spanning subgraph $G' = (A \cup B, E)$ of G with minimum degree $\delta(G') \geq (1 - 1/2\Delta + \gamma/2)m$, where $2m = n$. By Theorem 3.5 in [42], the minimum degree condition ensures the existence of a subgraph J of G' that is isomorphic to H. For each $a \in A$, let $N_a = \{b \in B : ab \in E(J)\}$ denote the neighbourhood of a in J. Construct an auxiliary bipartite graph $Q = (V(Q), E(Q))$. The vertex set is the multiset $V(Q) = A \cup \Gamma$, where $\Gamma = \{\{N_a : a \in A\}\}$ as a multiset. The edge set is defined as

$$E(Q) = \{a_1 N_{a_2} : \, N_{a_2} \subseteq N_G(a_1)\} \, .$$

Note that Q is a balanced bipartite graph on $2m$ vertices. We first show that $\delta(Q) \geq (1/2 + \varepsilon)m$, where $\varepsilon = \gamma/2$. For each $a \in A$, there at most $(1/2\Delta - \varepsilon)m$ vertices $b \in B$ such that $ab \notin E(G)$. Since $\Delta(J) \leq \Delta$, for each $b \in B$ there exist at most Δ vertices $a' \in A$ such that $b \in N_{a'}$. Thus, there are at most $(1/2 - \varepsilon\Delta)m$ vertices $a' \in A$ such that $N_{a'}$ is not included in $N_G(a)$. In particular, we have $d_Q(a) \geq (1/2 + \varepsilon)m$. For each

$a \in A$, we have $N_Q(N_a) = \cap_{b\in N_a} N_G(b)$. So, by inclusion-exclusion,

$$d_Q(N_a) = |\cap_{b\in N_a} N_G(b)| \geq m - \sum_{b\in N_a}(m - |N_G(b)|) \geq (1/2+\varepsilon)m \ .$$

Hence, $\delta(Q) \geq (1/2+\varepsilon)n$.

Let χ be a proper μn-bounded edge colouring of G. Consider the following system of conflicts,

$$\begin{aligned}\mathcal{F}_Q = \{\{a_1N_{a_1'}, a_2N_{a_2'}\} :& \exists\, x, y \in E(G) \text{ with } \chi(x) = \chi(y) \\ & \text{and } \{x,y\} \subseteq E_G(a_1, N_{a_1'}) \cup E_G(a_2, N_{a_2'})\} \ .\end{aligned}$$

Fix an edge $a_1N_{a_1'} \in E(Q)$. For each $b_1 \in N_{a_1'}$, there are at most $\mu n/\Delta^2$ edges a_2b_2 with $\chi(a_2b_2) = \chi(a_1b_1)$. Again, since $\Delta(J) \leq J$, b_2 is in at most Δ neighbourhoods $\Gamma_{a_2'}$. So, for each $b \in N_{a_1'}$, there are most $\mu n/\Delta$ edges $a_2N_{a_2'}$ conflicting with $a_1N_{a_1'}$. Since $|N_{a_1'}| \leq \Delta$, the total number of conflicts involving edge $a_1N_{a_1'}$ is at most $\mu n = 2\mu m$. So $\mathcal{F}$ is $2\mu m$-bounded.

We can apply Theorem 4.4 to the balanced bipartite graph Q and the system of conflicts $\mathcal{F}_Q$ to deduce the existence of a $\mathcal{F}_Q$-conflict-free perfect matching M in Q. Define the subgraph $R = (V(R), E(R))$ of G as follows. The vertex set is $V(R) = V(G)$ and edge set is

$$E(R) = \{ab \in E(G) : \text{ there exists } a' \in A \text{ such that } aN_{a'} \in E(M) \text{ and } b \in N_{a'}\} \ .$$

We claim that R is a rainbow subgraph of G isomorphic to H. Consider a bijective map $f : V(G) \to V(G)$, such that $f(u) = v$ if and only if $uN_v \in M$ for $u \in A$ and f is the identity map on B. We claim that f is an isomorphism from R to J. To see this, first observe that f is an automorphism of $V(G)$. Now, consider an edge $ab \in E(R)$ and note that $f(a)f(b) = f(a)b$ where $f(a)$ is such that $aN_{f(a)} \in M$. As M is a matching, there is only one choice $N_{a'}$ such that $aN_{a'} \in E(M)$, implying that $a' = f(a)$. By definition of $E(R)$, we have that $b \in N_{f(a)}$, so $f(a)b = f(a)f(b) \in E(J)$. Similarly, one can show that for all edges $ab \in E(J)$, $f^{-1}(a)f^{-1}(b) = f^{-1}(a)b \in E(R)$. Thus f is an isomorphism between R and J, and since J is isomorphic to H, so is R.

Finally, suppose for contradiction that there exist $x, y \in E(R)$ with $\chi(x) = \chi(y)$. If $x = a_1b_1$ and $y = a_2b_2$, let $a_1', a_2' \in A$ be such that $a_1N_{a_1'}, a_2'N_{a_2'} \in E(M)$. Then, as $x, y \in E(R)$, we have $b_1 \in N_{a_1'}$ and $b_2 \in N_{a_2'}$, implying that $a_1N_{a_1'}$ and $a_2'N_{a_2'}$ conflict under $\mathcal{F}_Q$. This is a contradiction as M is a $\mathcal{F}_Q$-conflict-free perfect matching. So R is rainbow. $\square$

4.8 Further remarks

We conclude this chapter with a number of remarks and open questions.

1) The condition on the minimum degree in Theorem 4.3 is best possible. However, the value of μ that follows from our proof is far from being optimal. In Section 4.6, we showed that the statement is not true if $\mu > 1/4$. Obtaining the best possible value for μ is a difficult problem, since it would imply a minimum degree version of the Ryser-Brualdi-Stein conjecture, which is wide open.

2) We believe that the statement of Theorem 4.3 should also hold for system of conflicts. The only obstacle in our proof is Lemma 4.9, which, in its current form, cannot be adapted to deal with conflicts instead of colours.

3) As shown in Section 4.7, the methods presented in this chapter are of potential interest to embed other conflict-free spanning structures in graphs with large minimum degree, beyond perfect matchings. We study H-factors and Hamilton cycles in the following two chapters. Krivelevich et al. [72] proved the existence of $\mathcal{F}$-conflict-free Hamiltonian cycles in Dirac graphs, provided that the conflicts in $\mathcal{F}$ are local. Their proof is substantially different from ours and relies on Pósa rotations.

4) Lu and Székely [82] generalised the idea of system of conflicts to include, not only unordered pairs of edges, but sets of edges of any size. Under some sparsity conditions on the set of conflicts, they proved the existence of conflict-free perfect matchings in $K_{n,n}$. Our results can be seen as a first step towards extending the Lu-Székely framework to Dirac graphs.

5) Csaba [30] proved the Bollobás-Eldridge-Catlin conjecture for embedding bipartite graphs of maximum degree Δ into any graph G of minimum degree at least $(1-\beta)(1-1/(\Delta+1))n$ for some $\beta > 0$. It would be of interest to determine whether a form of Theorem 4.17 holds in this setting, since it does not follow as a direct consequence of Theorem 4.3 nor indeed is it immediate from the rainbow blow-up lemma of Glock and Joos [44].

CHAPTER 5

RAINBOW H-FACTORS

5.1 Introduction

A fundamental question in Extremal Combinatorics is to determine conditions on a hypergraph G that guarantee an embedded copy of some other hypergraph H. The Turán problem for an r-graph H asks for the maximum number of edges in an H-free r-graph G on n vertices; we usually think of H as fixed and n as large. For $r = 2$ (ordinary graphs) this problem is fairly well understood (except when H is bipartite), but for general r and general H we do not even have an asymptotic understanding of the Turán problem (see the survey [63]). For example, a classical theorem of Mantel determines the maximum number of edges in a triangle-free graph on n vertices (it is $\lfloor n^2/4 \rfloor$), but we do not know even asymptotically the maximum number of edges in a tetrahedron-free 3-graph on n vertices. On the other hand, if we seek to embed a spanning hypergraph then it is most natural to consider minimum degree conditions. Such questions are known as Dirac-type problems, after the classical theorem of Dirac that any graph on $n \geq 3$ vertices with minimum degree at least $n/2$ contains a Hamilton cycle. There is a large literature on such problems for graphs and hypergraphs, surveyed in [78, 76, 102, 120].

One of these problems, finding hypergraph factors, will be our topic for the remainder of this chapter. To describe it we introduce some notation and terminology. Let G be an r-graph on $[n] = \{1, \dots, n\}$. For any $L \subseteq V(G)$ the degree of L in G is the number of edges of G containing L. The minimum ℓ-degree $\delta_\ell(G)$ is the minimum degree in G over all $L \subseteq V(G)$ of size ℓ. Let H be an r-graph with $|V(H)| = h \mid n$. A partial H-factor F in G of size m is a set of m vertex-disjoint copies of H in G. If $m = n/h$ we call F an H-factor. We let $\delta_\ell(H, n)$ be the minimum δ such that $\delta_\ell(G) > \delta n^{r-\ell}$ guarantees an H-factor in G. Then we define the asymptotic ℓ-degree threshold for H-factors as

$$\delta_\ell^*(H) := \lim_{m\to\infty} \delta_\ell(H, mh)\ . \tag{5.1}$$

We refer to Section 2.1 in [120] for a summary of the known bounds on $\delta_\ell^*(H)$ (using different notation). As for the Turán problem, $\delta_1^*(H)$ is well-understood for graphs [70, 77], but there are few results for hypergraphs. Even for perfect matchings (the case when H is a single edge) there are many cases for which the problem remains open (this is closely connected to the Erdős Matching Conjecture [36]).

Note that the limit in the definition of $\delta_\ell^*(H)$ does indeed exist. The proof of this is a relatively straightforward adaptation of the proof of our main theorem which we sketch in Appendix A.

Let us now introduce colours on the edges of G and ask for conditions under which we can embed a copy of H that is *rainbow*, meaning that its edges have distinct colours. Besides being a natural problem in its own right, this general framework also encodes many other combinatorial problems. Perhaps the most well-known of these is the Ryser-Brualdi-Stein Conjecture (Conjecture 4.1) on transversals in Latin squares, which is equivalent to saying that any proper edge-colouring of the complete bipartite graph $K_{n,n}$ has a rainbow matching of size $n-1$. There are several other well-known open problems that can be encoded as finding certain rainbow subgraphs in graphs with an edge-colouring that is locally k-bounded for some k, meaning that each vertex is in at most k edges of any given colour (so $k = 1$ is proper colouring). For example, a recent result of Montgomery, Pokrovskiy and Sudakov [91] shows that any locally k-bounded edge-colouring of K_n contains a rainbow copy of any tree of size at most $n/k - o(n)$, and this implies asymptotic solutions to the conjectures of Ringel [101] on decompositions by trees and of Graham and Sloane [47] on harmonious labellings of trees.

We now consider rainbow versions of the extremal problems discussed above. The rainbow Turán problem for an r-graph H is to determine the maximum number of edges in a properly edge-coloured r-graph G on n vertices with no rainbow H. This problem was introduced by Keevash, Mubayi, Sudakov and Verstraëte [64], who were mainly concerned with degenerate Turán problems (the case of even cycles encodes a problem from Number Theory), but also observed that a simple supersaturation argument shows that the threshold for non-degenerate rainbow Turán problems is asymptotically the same as that for the usual Turán problem, even if we consider locally $o(n)$-bounded edge-colourings.

For Dirac-type problems, it seems reasonable to make stronger assumptions on our colourings, as we have already noted that even locally bounded colourings of complete graphs encode many problems that are still open. For example, Erdős and Spencer [39] showed the existence of a rainbow perfect matching in any edge-colouring of $K_{n,n}$ that is $(n-1)/16$-bounded, meaning that are at most $(n-1)/16$ edges of any given colour. In the previos chapter we obtained a Dirac-type version of this result, showing that any μn-bounded edge-colouring of a subgraph of $K_{n,n}$ with minimum degree at least $n/2$ has a rainbow perfect matching. One could consider this a 'local resilience' version (as in [112]) of the Erdős-Spencer theorem. This is suggestive of a more general phenomenon, namely that for any Dirac-type problem, the rainbow problem for bounded colourings should have asymptotically the same degree threshold as the problem with no colours. A result of Yuster [119] on H-factors in graphs adds further evidence (but only for the weaker property of finding an H-factor in which each copy of H is rainbow). For graph problems, the general phenomenon was recently confirmed in considerable generality by Glock and Joos [44], who proved a rainbow version of the blow-up lemma of Komlós, Sárközy and Szemerédi [68] and the Bandwidth Theorem of Böttcher, Schacht and Taraz [15].

Our main result establishes the same phenomenon for hypergraph factors. We will use the following boundedness assumption for our colourings, in which we include the natural r-graph generalisations of both the local boundedness and boundedness assumptions from above (for $r = 2$ boundedness implies local boundedness, but in general they are incomparable assumptions).

Definition 5.1.1. An edge-colouring of an r-graph on n vertices is *μ-bounded* if for every colour c:

i) there are at most μn^{r-1} edges of colour c,

ii) for any set I of $r-1$ vertices, there are at most μn edges of colour c containing I.

Note that we cannot expect any result without some "global" condition as in Definition 5.1.1.i, since any H-factor contains linearly many edges. Similarly, some "local" condition along the lines of Definition 5.1.1.ii is also needed. Indeed, consider the edge-colouring of the complete r-graph K_n^r by $\binom{n}{r-1}$ colours identified with $(r-1)$-subsets of $[n]$, where each edge is coloured by its $r-1$ smallest elements. Suppose H has the property that every $(r-1)$-subset of $V(H)$ is contained in at least 2 edges of H (e.g. suppose H is also complete). Then there are fewer than n edges of any given colour, but there is no rainbow copy of H (let alone an H-factor), as in any embedding of H all edges containing the $r-1$ smallest elements have the same colour.

Theorem 5.1.2. *Let $1/n \ll \mu \ll \varepsilon \ll 1/h \leq 1/r < 1/\ell \leq 1$ with $h|n$. Let H be an r-graph on h vertices and G be an r-graph on n vertices with $\delta_\ell(G) \geq (\delta_\ell^*(H) + \varepsilon)n^{r-\ell}$. Then any μ-bounded edge-colouring of G admits a rainbow H-factor.*

Furthermore, we shall show that there are graphs H such that one cannot completely remove the requirement for some ε in the above theorem.

Theorem 5.1.3. *Let $t \geq 3$ and H be K_t the clique on t vertices. Then for any fixed $k \geq 2$ and n sufficiently large there exists a graph G of minimum degree at least*

$$\delta(G) \geq \frac{t-1}{t}n + k - 2$$

and a $\max\{2, \binom{k}{2}\}$*-bounded colouring of its edges which has no rainbow* H*-factor.*

In particular, a rainbow Hajnal-Szemerédi theorem does not exist even for 2*-bounded colourings.*

Throughout the remainder of this chapter we fix ℓ, r, h, ε, μ, n, H and G as in the statement of Theorem 5.1.2. We also fix an integer m with $\mu \ll 1/m \ll \varepsilon$ and define $\gamma = (mh)^{-m}$.

5.2 Proof modulo lemmas

The outline of the proof of Theorem 5.1.2 is the same as that given by Erdős and Spencer [39] for the existence of Latin transversals: we consider a uniformly random H-factor $\mathcal{H}$ in G (there is at least one by definition of $\delta^*_\ell(H)$) and apply the Lopsided Lovász Local Lemma to show that $\mathcal{H}$ is rainbow with positive probability. We will show that the local lemma hypotheses hold if there are many feasible switchings, defined as follows.

Definition 5.2.1. Let F_0 be an H-factor in G and $H_0 \in F_0$. An (H_0, F_0)-switching is a partial H-factor Y in G with $V(H_0) \subseteq V(Y)$ such that

1. for each $H' \in F_0$ we have $V(H') \subseteq V(Y)$ or $V(H') \cap V(Y) = \emptyset$, and
2. each $H' \in Y$ shares at most one vertex with H_0.

Let Y' be obtained from Y by deleting all vertices in $V(H_0)$ and their incident edges. We call Y *feasible* if Y' is rainbow and does not share any colour with any $H' \in F_0 \setminus V(Y)$.

The idea of the switching defined above is that we may replace a small number of copies of H in the H-factor F_0 with different copies in order to remove the "bad copy" H_0 which prevents F_0 from being rainbow. See Figure 5.1 for an example of a switching.

The following lemma, proved in Section 5.4, reduces the proof of Theorem 5.1.2 to showing the existence of many feasible switchings. For this lemma and the rest of the chapter we shall define the size of a partial H-factor to be the number of copies of H which it contains.

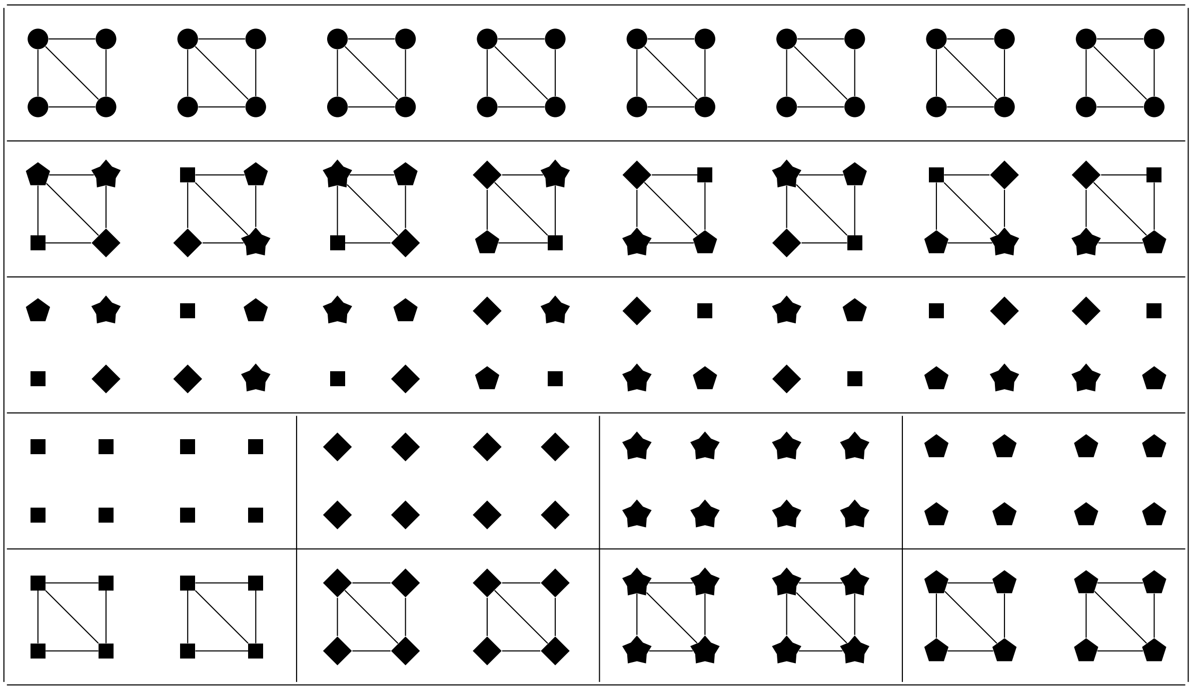


Figure 5.1: The process of an (H_0, F_0)-switching of size 8. We start with a partial H-factor of size 8 in the first line, produce a transverse partition as seen in the fourth line, and pick a new partial H-factor within each part of the transverse partition in the fifth line.

Lemma 5.1. *Suppose that for every H-factor F_0 of G and $H_0 \in F_0$ there are at least γn^{m-1} feasible (H_0, F_0)-switchings of size m. Then G has a rainbow H-factor.*

Note the exponent $m-1$ in Lemma 5.1 comes from the fact that we use $m-1$ additional copies of H in addition to H_0 in order to perform an (H_0, F_0)-switching of size m. That is Lemma 5.1 states that a constant fraction of candidate switchings being feasible suffices to find a rainbow H-factor.

We will construct switchings by randomly choosing some copies of H from F_0 and considering a random transverse partition in the sense of the following definition.

Definition 5.2.2. Let F_0 be an H-factor in G and $H_0 \in F_0$. Let $X \subseteq F_0$ be a partial H-factor in G with $H_0 \in X$. We call $S \subseteq V(X)$ transverse if $|V(H') \cap S| \leq 1$ for all $H' \in X$. We call a partition of $V(X)$ transverse if each part is transverse. For any edges e and f let $X(e, f) = \{H' \in X : |V(H') \cap (e \cup f)| \geq 2\}$. We call X suitable if

1. for any transverse $I \subseteq V(X) \setminus V(H_0)$ with $|I| = r - 1$ there are at most $\varepsilon|X|/4$

vertices $v \in V(X)$ such that $I \cup \{v\} \in E(G)$ shares a colour with some $H' \in F_0$, and

2. for any transverse edges e and f disjoint from $V(H_0)$ of the same colour we have $X(e,f) \neq \emptyset$, and, furthermore, if $e \cap f = \emptyset$, then $|X(e,f)| \geq 2$.

The following lemma, proved in Section 5.5, shows that a suitable partial H-factor has an associated feasible switching if it has a transverse partition whose parts each satisfy the minimum degree condition for an H-factor.

Lemma 5.2. *Let F_0, H_0 and X be as in Definition 5.2.2, suppose X is suitable and $|X| = m$. Let $\mathcal{P} = (V_1, \ldots, V_h)$ be a transverse partition of $V(X)$ (so $|V_i| = m$ for each i) and suppose $\delta_\ell(G[V_i]) \geq (\delta_\ell^*(H) + \varepsilon/2)m^{r-\ell}$ for all $i \in [h]$. Then there is a partial H-factor Y in G with $V(Y) = V(X)$ such that Y is a feasible (H_0, F_0)-switching.*

The following lemma, proved in Section 5.6, gives a lower bound on the number of partial H-factors X with some transverse partition $\mathcal{P}$ satisfying the conditions of the previous lemma.

Lemma 5.3. *Let F_0 be an H-factor in G and $H_0 \in F_0$. Let $X \subseteq F_0$ be a random partial H-factor where $H_0 \in X$ and each $H' \in F_0 \setminus \{H_0\}$ is included independently with probability $p = \frac{m}{n/h-1}$. Let $\mathcal{P} = (V_1, \ldots, V_h)$ be a uniformly random transverse partition of $V(X)$. Then with probability at least $1/m$ we have X suitable, $|X| = m$ and all $\delta_\ell(G[V_i]) \geq (\delta_\ell^*(H) + \varepsilon/2)m^{r-\ell}$.*

We conclude this section by showing how Theorem 5.1.2 follows from the above lemmas.

Proof of Theorem 5.1.2. By Lemma 5.1, it suffices to show that for every H-factor F_0 of G and $H_0 \in F_0$ there are at least γn^{m-1} feasible (H_0, F_0)-switchings of size m. There are $\binom{n/h-1}{m-1} \geq (n/mh - 1)^{m-1}$ partial H-factors X of size m with $H_0 \in X \subseteq F_0$. By Lemma 5.3, at least $m^{-1}(n/mh-1)^{m-1} > \gamma n^{m-1}$ of these are suitable and have a transverse partition $\mathcal{P} = (V_1, \ldots, V_h)$ with all $\delta_\ell(G[V_i]) \geq (\delta_\ell^*(H) + \varepsilon/2)m^{r-\ell}$. By Lemma 5.2, each such X has an associated feasible (H_0, F_0)-switching. $\square$

5.3 Probabilistic methods

In this section we collect various probabilistic tools that will be used in the proofs of the lemmas stated in the previous section. We will also use the bounded version of the **p**-lopsided Lovász local lemma (Corollary 3.5.)

We start with Talagrand's Inequality, see e.g. [90, page 81].

Theorem 5.4. *Let $c, r > 0$ and let $X \geq 0$ be a random variable determined by n independent trials which is:*

c-Lipschitz. *Changing the outcome of any one trial can affect X by at most c.*

r-certifiable. *For each $s \geq 1$, if $X \geq s$ then there is a set of at most rs trials whose outcomes certify $X \geq s$.*

Then for any $0 \leq t \leq \mathbb{E}[X]$,

$$\mathbb{P}[|X - \mathbb{E}[X]| > t + 60c\sqrt{r\mathbb{E}[X]}] \leq 4e^{-t^2/(8c^2 r\mathbb{E}[X])}.$$

Next we state an inequality of Janson [56].

Definition 5.3.1. Let $\{I_i\}_{i\in\mathcal{I}}$ be a finite family of indicator random variables. We call a graph Γ on $\mathcal{I}$ a *strong dependency graph* if the families $\{I_i\}_{i\in A}$ and $\{I_i\}_{i\in B}$ are independent whenever A and B are disjoint subsets of $\mathcal{I}$ with no edge of Γ between A and B.

Theorem 5.3.2. *In the setting of Definition 5.3.1, let $p_j = \mathbb{E}(I_j)$, $S = \sum_{i\in\mathcal{I}} I_i$, $\mu = \mathbb{E}[S]$, $\delta = \max_{i\in\mathcal{I}} \sum_{j:ij\in E(\Gamma)} p_j$ and $\Delta = \sum_{ij\in E(\Gamma)} \mathbb{E}[I_i I_j]$. Then for any $0 < \eta < 1$,*

$$\mathbb{P}[S < (1-\eta)\mu] \leq \exp(-\min\{(\eta\mu)^2/(8\Delta + 2\mu), \eta\mu/(6\delta)\}).$$

We conclude with a standard bound on the probability that a binomial is equal to its mean.

Lemma 5.5. *Let X be a binomial random variable with parameters n and p such that $np = m \in \mathbb{N}$ and $m^2 = o(n)$. Then $\mathbb{P}[X = m] \geq 1/(3\sqrt{m})$.*

Proof. The stated bound follows from $\mathbb{P}[X = m] = \binom{n}{m} p^m (1-p)^{n-m} \geq m!^{-1}(n - m)^m p^m (1-p)^{n-m} = m!^{-1} m^m (1-p)^n$, $(1-p)^n = e^{-np+O(np^2)}$ and $m! \leq e^{1-m} m^{m+1/2}$. $\square$

5.4 Applying the local lemma

In this section we prove Lemma 5.1, which applies the local lemma to reduce the proof of Theorem 5.1.2 to finding many feasible switchings.

Proof of Lemma 5.1. Suppose that for every H-factor F_0 of G and $H_0 \in F_0$ there are at least γn^{m-1} feasible (H_0, F_0)-switchings of size m. We need to show that G has a rainbow H-factor.

We will apply Corollary 3.5 to a uniformly random H-factor $\mathcal{H}$ in G, where $\mathcal{E} = \mathcal{A} \cup \mathcal{B}$ consists of all events of the following two types. For every copy H_0 of H in G and any two edges e and f in H_0 of the same colour we let $A(e, f, H_0)$ be the event that $H_0 \in \mathcal{H}$; we let $\mathcal{A} = \{A(e, f, H_0) : \mathbb{P}[A(e, f, H_0)] > 0\}$. Note that $A(e, f, H_0)$ does not actually depend on e, f, however their inclusion assists with counting how many of these events we have (the same is also true of $B(e, f, H_1, H_2)$ defined subsequently). For every pair H_1, H_2 of vertex-disjoint copies of H in G and edges e_1 of H_1 and e_2 of H_2 of the same colour we let $B(e_1, e_2, H_1, H_2)$ be the event that $H_1 \in \mathcal{H}$ and $H_2 \in \mathcal{H}$; we let $\mathcal{B} = \{B(e_1, e_2, H_1, H_2) : \mathbb{P}[B(e_1, e_2, H_1, H_2)] > 0\}$. Then $\mathcal{H}$ is rainbow iff none of the events in $\mathcal{E}$ occur.

We define the supports of $A = A(e, f, H_0)$ to be $\mathrm{supp}(A) = V(H_0)$ and also of $B = B(e_1, e_2, H_1, H_2)$ as $\mathrm{supp}(B) = V(H_1) \cup V(H_2)$. Let Γ be the graph on $\mathcal{A} \cup \mathcal{B}$ where $E_1, E_2 \in V(\Gamma)$ are adjacent if and only if $\mathrm{supp}(E_1) \cap \mathrm{supp}(E_2) \neq \emptyset$. Our goal is to show that there exist suitably small $p_{\mathcal{A}}, p_{\mathcal{B}}$ such that Γ is a p-dependency graph for $\mathcal{A} \cup \mathcal{B}$, where $p_A = p_{\mathcal{A}}$ for all $A \in \mathcal{A}$ and $p_B = p_{\mathcal{B}}$ for all $B \in \mathcal{B}$. For $\mathcal{X} \in \{\mathcal{A}, \mathcal{B}\}$, let $d_{\mathcal{X}}$ be the maximum over $E \in V(\Gamma)$ of the number of neighbours of E in $\mathcal{X}$. To apply Corollary 3.5, it suffices to show $p_{\mathcal{A}} d_{\mathcal{A}} + p_{\mathcal{B}} d_{\mathcal{B}} \leq 1/4$.

To bound the degrees, we will first estimate the number of events in $\mathcal{A}$ and $\mathcal{B}$ whose support contains any fixed vertex $v \in V(G)$.

Claim. There are at most $2^{r+1}h!\mu n^{h-1}$ events $A(e, f, H_0) \in \mathcal{A}$ with $v \in V(H_0)$.

Proof. To see this, first consider those events with $v \notin e \cup f$. For any $s < r$, as the colouring is μ-bounded, the number of choices of e and f of the same colour with $|e \cap f| = s$ is at most $n^r \cdot \binom{r}{s}\mu n^{r-s}$. For any such e and f with $v \notin e \cup f$, there are at most $h!n^{h-(2r-s+1)}$ copies of H containing $e \cup f \cup \{v\}$, so summing over s we obtain at most $2^r h!\mu n^{h-1}$ such events. Now we consider events $A(e, f, H_0)$ with $v \in e \cup f$. The number of choices of e and f of the same colour with $|e \cap f| = s$ and $v \in e \cup f$ is at most $n^{r-1} \cdot \binom{r}{s}\mu n^{r-s}$. For any such e and f there are at most $h!n^{h-(2r-s)}$ copies of H containing $e \cup f \cup \{v\}$, so summing over s we obtain at most $2^{r+1}h!\mu n^{h-1}$ such events. The claim follows. $\square$

Claim. There are at most $2(h!)^2\mu n^{2h-2}$ events $B(e_1, e_2, H_1, H_2) \in \mathcal{B}$ with $v \in V(H_1) \cup V(H_2)$.

Proof. To see this, first consider those events with $v \in e_1 \cup e_2$. By definition of $\mathcal{B}$, we may consider only disjoint edges e_1, e_2. There are at most $h!n^{h-r}$ choices for each of H_1 and H_2 given e_1 and e_2. Also, the number of choices for e_1 and e_2 is at most $n^{r-1} \cdot \mu n^{r-1} = \mu n^{2r-2}$. Thus, we obtain at most $(h!)^2\mu n^{2h-2}$ such events. A similar argument applies to events $B(e_1, e_2, H_1, H_2)$ with $v \notin e_1 \cup e_2$, and the claim follows. $\square$

In particular, these two claims allow us to deduce that there is some constant $C = C(r, h)$ so that

$$d_{\mathcal{A}} < C\mu n^{h-1} \quad \text{and} \quad d_{\mathcal{B}} < C\mu n^{2h-2}. \tag{5.2}$$

Now we will bound $p_{\mathcal{A}}$ and $p_{\mathcal{B}}$ using switchings. For $p_{\mathcal{A}}$ we need to bound $\mathbb{P}[A \mid \cap_{E \in \mathcal{E}'}\overline{E}]$ for any $A = A(e, f, H_0) \in \mathcal{A}$ and $\mathcal{E}' \subseteq \mathcal{E}$ such that $AE \notin E(\Gamma)$ for all $E \in \mathcal{E}'$ and $\mathbb{P}[\cap_{E \in \mathcal{E}'}\overline{E}] > 0$. Let $\mathcal{F}$ be the set of H-factors of G that satisfy $\cap_{E \in \mathcal{E}'}\overline{E}$; then $\mathcal{F} \neq \emptyset$ as $\mathbb{P}[\cap_{E \in \mathcal{E}'}\overline{E}] > 0$. Let $\mathcal{F}_0 = \{F_0 \in \mathcal{F} : H_0 \in F_0\}$. We consider the auxiliary bipartite multigraph $\mathcal{G}_A$ with parts $(\mathcal{F}_0, \mathcal{F} \setminus \mathcal{F}_0)$, where for each $F_0 \in \mathcal{F}_0$ and feasible (H_0, F_0)-switching Y of size m we add an edge from F_0 to F obtained by replacing $F_0[V(Y)]$ with Y; we note that $F \in \mathcal{F} \setminus \mathcal{F}_0$ as Y is rainbow and shares no colours with $H' \in F_0 \setminus V(Y)$ by Definition 5.2.1 hence F still satisfies $\cap_{E\in\mathcal{E}'}\overline{E}$. Let δ_A be the minimum degree in $\mathcal{G}_A$ of vertices in $\mathcal{F}_0$ and Δ_A be the maximum degree in $\mathcal{G}_A$ of vertices in $\mathcal{F} \setminus \mathcal{F}_0$. By double-counting the edges of $\mathcal{G}_A$ we obtain $\mathbb{P}[A \mid \cap_{E\in\mathcal{E}'}\overline{E}] = |\mathcal{F}_0|/|\mathcal{F}| \leq \Delta_A/\delta_A$.

We therefore need an upper bound for Δ_A and a lower bound for δ_A. By the hypotheses of the lemma, we have $\delta_A \geq \gamma n^{m-1}$. To bound Δ_A, we fix any $F \in \mathcal{F} \setminus \mathcal{F}_0$ and bound the number of pairs (F_0, Y) where $F_0 \in \mathcal{F}_0$ and Y is a feasible (H_0, F_0)-switching of size m that produces F. Each vertex of $V(H_0)$ must belong to a different copy of H in F, as otherwise there are no (H_0, F_0)-switchings that could produce F. Thus we identify h copies of H in F whose vertex set must be included in $V(Y)$. There at most n^{m-h} choices for the other copies of H to include in $V(Y)$ and then at most $(hm)!$ choices for Y, so $\Delta_A \leq (hm)!n^{m-h}$. We deduce

$$\mathbb{P}[A| \cap_{E\in\mathcal{E}'} \overline{E}] \leq (hm)!\gamma^{-1}n^{1-h} =: p_\mathcal{A} \ . \tag{5.3}$$

The argument to bound $p_\mathcal{B}$ is very similar. This time we do a double switching: one to remove each of the two copies of H which share a colour. We need to bound $\mathbb{P}[B \mid \cap_{E\in\mathcal{E}'}\overline{E}]$ for any $B = B(e_1, e_2, H_1, H_2) \in \mathcal{B}$ and $\mathcal{E}' \subseteq \mathcal{E}$ such that $BE \notin E(\Gamma)$ for all $E \in \mathcal{E}'$ and $\mathbb{P}[\cap_{E\in\mathcal{E}'}\overline{E}] > 0$. Let $\mathcal{F}$ be the set of H-factors of G that satisfy $\cap_{E\in\mathcal{E}'}\overline{E}$; then $\mathcal{F} \neq \emptyset$. Let $\mathcal{F}' = \{F' \in \mathcal{F} : \{H_1, H_2\} \subseteq F'\}$. We consider the auxiliary bipartite multigraph $\mathcal{G}_B$ with parts $(\mathcal{F}', \mathcal{F} \setminus \mathcal{F}')$, where there is an edge from $F' \in \mathcal{F}'$ to F for each pair (Y, Z), where Y is a feasible (H_1, F')-switching of size m producing some H-factor F'' containing H_2 but not H_1, and Z is a feasible (H_2, F'')-switching of size m with $V(Z) \cap V(H_1) = \emptyset$ producing F; note that then $F \in \mathcal{F} \setminus \mathcal{F}'$.

We have $\mathbb{P}[B \mid \cap_{E\in\mathcal{E}'}\overline{E}] \leq \Delta_B/\delta_B$, where Δ_B and δ_B are defined analogously to Δ_A and δ_A. The condition $V(Z) \cap V(H_1) = \emptyset$ rules out at most hn^{m-2} choices of Z given

H_1, and similarly the condition that F'' contains H_2 and not H_1 rules out at most hn^{m-2} choices of Y given H_2. So $\delta_B \geq (\gamma n^{m-1} - hn^{m-2})^2 > \frac{1}{2}\gamma^2 n^{2m-2}$. Similarly to before we have $\Delta_B \leq ((hm)! n^{m-h})^2$, so

$$\mathbb{P}[B | \cap_{E \in \mathcal{E}'} \overline{E}] \leq 2(hm)!^2 \gamma^{-2} n^{2-2h} =: p_{\mathcal{B}} \; . \tag{5.4}$$

Combining (5.2), (5.3) and (5.4) we have $p_{\mathcal{A}} d_{\mathcal{A}} + p_{\mathcal{B}} d_{\mathcal{B}} \leq 1/4$, so the lemma follows from Corollary 3.5. $\square$

5.5 Switchings

In this section we prove Lemma 5.2, which shows how to obtain a feasible switching from a suitable partial H-factor and transverse partition whose parts have high minimum degree.

Proof of Lemma 5.2. Let F_0 be an H-factor in G and $H_0 \in F_0$. Let $X \subseteq F_0$ be a suitable partial H-factor in G of size m with $H_0 \in X$. Let $\mathcal{P} = (V_1, \dots, V_h)$ be a transverse partition of $V(X)$ such that all $\delta_\ell(G[V_i]) \geq (\delta^*_\ell(H) + \varepsilon/2) m^{r-\ell}$. We need to find a partial H-factor Y in G with $V(Y) = V(X)$ such that Y is a feasible (H_0, F_0)-switching.

We construct Y by successively choosing H-factors Y_i of $G[V_i]$ for $1 \leq i \leq h$. For each i we let $V'_i = V_i \setminus V(H_0)$ and we will show that $G[V'_i]$ is rainbow by Definition 5.2.2.ii. This is because every subset of V'_i is transverse by definition. However, if edges e and f are both transverse and have the same colour then by Definition 5.2.2.ii their union is not transverse. At step i, we let G_i be the r-graph obtained from $G[V_i]$ by deleting all edges disjoint from $V(H_0)$ that share a colour with any H' in F_0 or $\cup_{j<i} Y_j$. It suffices to show that G_i has an H-factor Y_i, as then $Y = \cup_{i=1}^h Y_i$ will be feasible.

By definition of $\delta^*_\ell(H)$, it suffices to show that for each $L \subseteq V_i$ with $|L| = \ell$ that we delete at most $\frac{\varepsilon}{2} m^{r-\ell}$ edges containing L. We can assume L is disjoint from $V(H_0)$, as otherwise we do not delete any edges containing L. There are $\binom{m-\ell}{r-1-\ell}$ choices of I of size $r-1$ with $L \subseteq I \subseteq V_i$. For each such I, by Definition 5.2.2.i, the number of edges containing I deleted due to sharing a colour with any $H' \in F_0$ is at most $\varepsilon m/4$. Thus we delete at most $\frac{\varepsilon}{4}m^{r-\ell}$ such edges containing L.

It remains to consider edges containing L that are deleted due to sharing a colour with any H' in $\cup_{j<i}Y_j$. As $G[V_i']$ is rainbow, any colour in $\cup_{j<i}Y_j$ accounts for at most one deleted edge. In the case $\ell \le r-2$ we can crudely bound the number of deleted edges by the total number of edges in $\cup_{j<i}Y_j$, which is at most $ie(H)m < mh^{r+1} < \frac{\varepsilon}{4}m^{r-\ell}$.

Now we may suppose $\ell = r-1$. Consider any edge e containing L that is deleted due to having the same colour as some edge f in some Y_j with $j < i$. By Definition 5.2.2.ii and $|e \setminus L| = 1$ there is a copy H' of H in X that intersects both L and f. To bound the number of choices for e, note that there are $|L| = r-1$ choices for H' and $i-1$ choices for j. These choices determine a vertex in V_j, and so a copy of H in Y_j, which contains at most h^{r-1} choices for f. Then the colour of f determines at most one deleted edge in e. Thus the number of such deleted edges e containing L is at most $(r-1)(i-1)h^{r-1} < \frac{\varepsilon}{4}m$, as required. $\square$

5.6 Transverse partitions

To complete the proof of Theorem 5.1.2, it remains to prove Lemma 5.3, which bounds the probability that a random partial H-factor and transverse partition satisfy the hypotheses of Lemma 5.2.

Proof of Lemma 5.3. Let F_0 be an H-factor in G and $H_0 \in F_0$. Let $X \subseteq F_0$ be a random partial H-factor where $H_0 \in X$ and each $H' \in F_0 \setminus \{H_0\}$ is included independently with probability $p = \frac{m-1}{n/h-1} \le \frac{hm}{n}$. Let $\mathcal{P} = (V_1, \dots, V_h)$ be a uniformly random transverse partition of $V(X)$. Note that each copy H' of H in X has one vertex in each V_i, according to a uniformly random bijection between $V(H')$ and $[h]$, and that these bijections are independent for different choices of H'. Consider the events

$$\mathcal{E}_1 = \{|X| = m\}, \qquad \mathcal{E}_2 = \{X \text{ satisfies Definition 5.2.2.ii}\},$$
$$\mathcal{E}_3 = \{X \text{ satisfies Definition 5.2.2.i}\}, \qquad \mathcal{E}_4 = \cap_{i=1}^{h}\{\delta_\ell(G[V_i]) \geq (\delta_\ell^*(H) + \varepsilon/2)m^{r-\ell}\}.$$

We need to show that $\mathbb{P}[\cap_{i=1}^{4}\mathcal{E}_i] > 1/m$. To do so, we first recall from Lemma 5.5 that $\mathbb{P}[\mathcal{E}_1] \geq 1/(4\sqrt{m})$. To complete the proof, we will show that $\mathbb{P}[\mathcal{E}_i] \geq 1-1/m$ for $i = 2, 3, 4$. Throughout, for $I \subseteq V(G)$ we let $F_I \subseteq F_0$ be the partial H-factor consisting of all copies of H in F_0 that intersect I.

Bounding $\mathbb{P}[\mathcal{E}_2]$.

For $s \in [r-1]$ let $\mathcal{Z}_s$ be the set of pairs (e, f) of transverse edges disjoint from $V(H_0)$ of the same colour with $|e \cap f| = s$ and $X(e, f) = \emptyset$. As the colouring is μ-bounded, we have $|\mathcal{Z}_s| \leq n^r \cdot \binom{r}{s}\mu n^{r-s}$. For any $(e, f) \in \mathcal{Z}_s$ we have $|F_{e\cup f}| = 2r - s$, so $\mathbb{P}[e \cup f \subseteq V(X)] = p^{2r-s}$. By a union bound, the probability that any such event holds is at most $\sum_{s=1}^{r-1}\binom{r}{s}\mu n^{2r-s}p^{2r-s} < (hm)^r(hm+1)^r\mu < 1/2m$.

Similarly, let $\mathcal{Z}_0$ be the set of pairs (e, f) of transverse edges disjoint from $V(H_0)$ of the same colour with $e \cap f = \emptyset$ and $|X(e, f)| \leq 1$. As the colouring is μ-bounded, we have $|\mathcal{Z}_0| \leq n^r \cdot \mu n^{r-1}$. For any $(e, f) \in \mathcal{Z}_0$, $|F_{e\cup f}| \geq 2r - 1$ and $\mathbb{P}[e \cup f \subseteq V(X)] \leq p^{2r-1}$. Thus the probability that any such event holds is at most $\mu(hm)^{2r-1} < 1/2m$.

Bounding $\mathbb{P}[\mathcal{E}_3]$.

For any transverse $I \subseteq V(X) \setminus V(H_0)$ with $|I| = r - 1$ we let B_I be the set of $v \in V(G) \setminus (V(F_I) \cup V(H_0))$ such that $I \cup \{v\}$ is an edge sharing a colour with some $H' \in F_0$. Write $Y_I = |V(X) \cap B_I|$. It suffices to bound the probability that there is any $I \subseteq V(X)$ with $Y_I > \varepsilon m/5$. Indeed, the number of $v \in V(F_I) \cup V(H_0)$ such that $I \cup \{v\}$ is an edge is at most $rh < \varepsilon m/20$.

First we show that X is unlikely to contain any I in $\mathcal{B} := \{I : |B_I| > \varepsilon n/10h\}$. Indeed, as the colouring is μ-bounded, there are at most $e(F_0)\mu n^{r-1} = \mu e(H)n^r/h$ edges with colours in F_0, so $|\mathcal{B}| < \mu\varepsilon^{-2}n^{r-1}$. For each transverse I we have $\mathbb{P}[I \subseteq V(X)] = p^{r-1}$, so by a union bound, the probability that X contains any I in $\mathcal{B}$ is at most $\mu\varepsilon^{-2}(hm)^{r-1} < 1/2m$.

Now for each $I \notin \mathcal{B}$ we bound Y_I by Talagrand's inequality, where the independent trials are the decisions for each $H' \in F_0 \setminus \{H_0\}$ of whether to include H' in X. As $I \notin \mathcal{B}$ we have $\mathbb{E}[Y_I] = p|B_I| \leq \varepsilon m/10$. Also, Y_I is clearly h-Lipschitz as $|H| = h$ and 1-certifiable as we can simply list the successful trials containing the vertices of $V(X) \cap B_I$. We apply Theorem 5.4 to $Y_I' = Y_I + \varepsilon m/30$, with $t = \varepsilon m/30 \leq \mathbb{E}[Y_I']$, $c = h$ and $r = 1$ to deduce $\mathbb{P}[Y_I > \varepsilon m/5] \leq 4e^{-10^{-4}h^{-2}\varepsilon^2 m} < m^{-2r}$.

As we excluded $V(F_I)$ from B_I, the events $\{I \subseteq V(X)\}$ and $Y_I > \varepsilon m/5$ are independent, so both occur with probability at most $p^{r-1}m^{-2r}$. Taking a union bound over at most n^{r-1} choices of I, we obtain $\mathbb{P}[\overline{\mathcal{E}_3}] < 1/m$.

Bounding $\mathbb{P}[\mathcal{E}_4]$.

For $L \subseteq V(G)$ with $|L| = \ell$ and $i \in [h]$ we define

$$\mathcal{J}_L = \{J \subseteq V(G) \setminus V(H_0) : \ F_L \cap F_J = \emptyset \ \text{ and } \ L \cup J \in E(G) \text{ is transverse}\}.$$

We say L is i-bad if $L \subseteq V_i$ and $d_i'(L) := |\{J \in \mathcal{J}_L : J \subseteq V_i\}| < (\delta_\ell^*(H) + \varepsilon/2)m^{r-\ell}$. We will give an upper bound on the probability that there is any i-bad L.

First we note that the events $\{L \subseteq V_i\}$ and $\{J \subseteq V_i\}$ are independent for any $J \in \mathcal{J}_L$. There are at most n^ℓ choices of L with $L \cap V(H_0) = \emptyset$, each of which has $\mathbb{P}[L \subseteq V_i] = (p/h)^\ell$, and at most $hn^{\ell-1}$ choices of L with $|L \cap V(H_0)| = 1$, each of which has $\mathbb{P}[L \subseteq V_i] \leq (p/h)^{\ell-1}$. By a union bound, it suffices to show for every transverse L and $i \in [h]$ that $\mathbb{P}[d_i'(L) < (\delta_\ell^*(H) + \varepsilon/2)m^{r-\ell}] < m^{-2r}$.

We also note that $|\mathcal{J}_L| \geq (\delta_\ell^*(H) + 0.9\varepsilon)n^{r-\ell}$, as there are at least $(\delta_\ell^*(H) + \varepsilon)n^{r-\ell}$ choices of J with $L \cup J \in E(G)$, of which the number excluded due to $J \cap V(H_0) \neq \emptyset$, $F_L \cap F_J \neq \emptyset$ or $L \cup J$ not being transverse is at most

$$hn^{r-\ell-1} + \ell hn^{r-\ell-1} + \tfrac{n}{h}\tbinom{h}{2}n^{r-\ell-2} < 0.1\varepsilon n^{r-\ell}.$$

We will apply Janson's inequality to $d'_i(L) = \sum_{J \in \mathcal{J}_L} I_J$, where each I_J is the indicator of $\{J \subseteq V_i\}$. As $\mathbb{P}[J \subseteq V_i] = (p/h)^{r-\ell}$ for each $J \in \mathcal{J}_L$, we have $\mu = \mathbb{E}[d'_i(L)] > (\delta^*_\ell(H) + 0.9\varepsilon)m^{r-\ell}$. We use the dependency graph Γ where JJ' is an edge iff $F_J \cap F_{J'} \neq \emptyset$. Note that for any $J \in \mathcal{J}_L$ and $s \in [r-\ell]$ the number of choices of J' with $|F_J \cap F_{J'}| = s$ is at most $\binom{r-\ell}{s} h^s n^{r-\ell-s}$, and for each we have $\mathbb{P}[J \cup J' \subseteq V_i] = (p/h)^{2(r-\ell)-s}$. Thus we can bound the parameter Δ in Theorem 5.3.2 as

$$\Delta \leq |\mathcal{J}_L| \sum_{s=1}^{r-\ell} \binom{r-\ell}{s} h^s n^{r-\ell-s} \left(\frac{p}{h}\right)^{2(r-\ell)-s} \leq m^{r-\ell} \sum_{s=1}^{r-\ell} \binom{r-\ell}{s} h^s m^{r-\ell-s} < 2h(r-\ell) m^{2(r-\ell)-1}.$$

We also have

$$\delta \leq \sum_{s=1}^{r-\ell} \binom{r-\ell}{s} h^s n^{r-\ell-s} (p/h)^{r-\ell-s} \leq \sum_{s=1}^{r-\ell} \binom{r-\ell}{s} h^s m^{r-\ell-s} < 2h(r-l) m^{r-\ell-1}.$$

By Theorem 5.3.2, there is some constant $c = c(r, \varepsilon, h)$ independent of m so that

$$\mathbb{P}[d'_i(L) < (\delta^*_\ell(H) + \varepsilon/2)m^{r-\ell}] < e^{-cm} < m^{-2r},$$

as required. $\square$

5.7 Non Coincidence of Thresholds

In this section we prove Theorem 5.1.3. That is we construct coloured graphs with K_t-factors and no rainbow K_t factor using as few as two copies of each colour.

Proof of Theorem 5.1.3. Choose $m \in \mathbb{N}$ such that $k | m+1$ and let $n = tm$. We construct the graph G on n vertices as follows.

Let A and B be disjoint sets of $(t-1)m - 1$ and $m+1$ vertices respectively. Let $V(G) = A \cup B$. We shall make use of the fact that any K_t-factor in G must have at least one copy of K_t which uses two or more vertices from B.

Let F be an arbitrary partition of B into k-sets (one exists as $k | m+1$).

Then, the following shall be the edges of G.

$$E(G) = \{xy | x \in A, y \in A \cup B, x \neq y\} \cup E(F)$$

It is easy to check that G has the claimed minimum degree.

We shall now colour the edges of G. First give each edge with both ends in A a unique colour from $\mathbb{N}$, and also give each k-clique in B an unique colour from $\mathbb{N}$ i.e. give all its edges the same colour. For each edge ab such that $a \in A$ and $b \in B$, we will give it the colour $\{a, c(b)\}$ where $c(b)$ is the colour of the clique containing b. Then, the colours of edges inside A, between A and B and inside B are disjoint. The colouring is 1-bounded inside A, k-bounded between A and B and $\binom{k}{2}$-bounded inside B. Thus the colouring is $\max\{2, \binom{k}{2}\}$-bounded.

To prove that there is no rainbow K_t-factor, of G note that any K_t-factor must contain at least one copy, say Q, of K_t with at least one edge inside B. If $|V(Q) \cap B| \geq 3$, then $Q[V(Q) \cap B]$ is a subgraph of some element of F. Thus $Q[V(Q) \cap B]$ is monochromatic with at least 3 edges so Q is not rainbow. Thus, Q contains at least one vertex $a \in A$ (as $t \geq 3$). Suppose $V(Q) \cap B = \{b_1, b_2\}$. Then, ab_1 and ab_2 both have colour $\{a, c(b_1)\}$ (as b_1 and b_2 are in the same clique so $c(b_1) = c(b_2)$. Hence Q is not monochromatic. So G has no rainbow K_t-factor. □

5.8 Concluding remarks

Our result and those of [44] suggest that for any Dirac-type problem, the rainbow problem for bounded colourings should have asymptotically the same degree threshold as the problem with no colours. In particular, it may be interesting to establish this for Hamilton cycles in hypergraphs (i.e. a Dirac-type generalisation of [34]) in the next chapter we study this question in graphs suggesting that a hypergraph analogue is also likely to hold. The local resilience perspective emphasises analogies with the recent literature on Dirac-type problems in the random setting (see the surveys [13, 110]), perhaps suggests looking for common generalisations, e.g. a rainbow version of [80]: in the random graph $G(n,p)$ with $p > C(\log n)/n$, must any $o(pn)$-bounded edge-colouring of any subgraph H with minimum degree $(1/2 + o(1))pn$ have a rainbow Hamilton cycle?

Furthermore, in light of the counterexample in section 5.7, it would be interesting to understand which subgraphs have their "rainbow threshold" in the same place as their extremal threshold. That is, consider a sequence $\mathcal{H} = H_n$ of graphs (where $v(H_n) = n$.) Suppose that any graph G of minimum degree at least $\delta_{\mathcal{H}}(n)$ on n vertices contains H_n and there exists a graph on n vertices of minimum degree $\delta_{\mathcal{H}}(n) - 1$ that does not contain H_n. When is it the case that there exists $\mu > 0$ such that any μn-bounded edge colouring of any graph G with n vertices and minimum degree at least $\delta_{\mathcal{H}}(n)$ contains a rainbow copy of H_n?

CHAPTER 6

RAINBOW HAMILTON CYCLES

6.1 Introduction

Recall from Chapter 4 the Ryser-Brualdi-Stein Conjecture.

Conjecture 6.1 (Ryser, Brualdi, Stein [17, 103, 109]). *Every Latin square of order n contains a partial transversal of size at least $n-1$.*

Recall further that Latin squares are in bijection with proper n-colourings of the edges of the complete bipartite graph $K_{n,n}$. If G is an edge-coloured graph and $H \subseteq G$, we say that H is *rainbow* if no two edges of H have the same colour. In the setting of edge-coloured graphs, the Ryser-Brualdi-Stein conjecture states that any proper edge-colouring of $K_{n,n}$ using n colours has a rainbow matching of size at least $n-1$. Looking instead at symmetric Latin squares which can be seen to be in bijection with proper colourings of K_n, the conjecture implies that any proper edge-colouring of K_n using n colours has a rainbow subgraph with at least $n-2$ edges and maximum degree 2. It is natural to ask whether similar phenomena occur under weaker conditions on the colourings. Recall that an edge-colouring of G such that no colour appears more than k times on its edges is a *k-bounded colouring* of $E(G)$. In this framework, Hahn gave the following conjecture:

Conjecture 6.2 (Hahn [49]). *Any $(n/2)$-bounded colouring of $E(K_n)$ contains a rainbow Hamilton path.*

Hahn's conjecture was disproved by Maamoun and Meyniel [87] who showed it was not even true for even proper colourings of K_{2^t} for integers $t \geq 2$ which disproves Hahn's conjecture as proper colourings are $n/2$-bounded.

Motivated by Hahn's conjecture, one could ask for which k any k-bounded colouring of K_n contains a rainbow Hamilton path (or cycle). Hahn and Thomassen [50] showed that $k = o(n^{1/3})$ is sufficient. This was subsequently improved by Albert, Frieze and Reed [2] who used the local lemma to prove that one can take $k = n/64$. This question has also been studied for Hamilton cycles in complete hypergraphs [34, 35] and generalised to embedding rainbow copies of other spanning subgraphs H in complete structures [14, 59, 111]. In addition, there has been recent progress on approximate rainbow decompositions [65, 92].

Here we will be interested in embedding rainbow subgraphs into sparser graphs. Due to the nature of the proofs, most of the previous results can be adapted to host graphs G with minimum degree $\delta(G) = (1 - O(1/\Delta))n$, where Δ is the maximum degree of H. However, the bound obtained for the minimum degree seems far from being tight. Recent work has shown that for certain spanning subgraphs H (including Hamilton cycles), the minimum degree threshold for rainbowly embedding H is asymptotically the same as for embedding H [18, 27, 44].

In this chapter we determine the exact minimum degree threshold at which rainbow Hamilton cycles appear. In his famous theorem [33], Dirac showed that any graph G on n vertices with minimum degree at least $n/2$ has a Hamilton cycle. We call such graphs, *Dirac graphs*. Krivelevich, Lee and Sudakov [72] proved the existence of properly coloured Hamilton cycles in edge-coloured Dirac graphs where each colour appears at most $k = o(n)$ times in the edges incident to each vertex. In fact, their result applies to the more general setting of incompatibility systems, solving a conjecture of Häggkvist.

The main result of this chapter is a Dirac theorem for rainbow Hamilton cycles that holds for $o(n)$-bounded colourings.

Theorem 6.3. *There exist $\mu > 0$ and $n_0 \in \mathbb{N}$ such that if $n \geq n_0$ and G is a Dirac graph* *on n vertices, then any μn-bounded colouring of $E(G)$ contains a rainbow Hamilton cycle.*

Our theorem can also be seen as a rainbow analogue of the result of Krivelevich, Lee and Sudakov.

Note that a linear bound on the number of occurrences of each colour is necessary as otherwise we could have less than n colours in total and no rainbow Hamilton cycle would exist. Our next result shows that we need $\mu \leq 1/8$.

Theorem 6.4. *For every sufficiently large $n \in \mathbb{N}$ and every $\mu > 1/8$, there exists a Dirac graph G on n vertices and a μn-bounded colouring of $E(G)$ such that G does not contain a rainbow Hamilton cycle.*

The proof of Theorem 6.3 extends the ideas introduced in Chapter 4 to deal with perfect matchings in bipartite graphs. Firstly, we use a classification for Dirac graphs observed by Kühn, Lapinskas and Osthus in [73]: either the graph has good expansion properties (*robust expander*, see e.g. [79]) or the graph is extremal in some sense: it either resembles a disjoint pair of cliques or a complete balanced bipartite graph. Similar classifications for Dirac graphs have been used in the literature (see e.g. [69, 71]). For extremal graphs, we fix a partial rainbow matching only using atypical edges and we extend it to a rainbow Hamilton cycle with an application of the lopsided version of the Lovász Local Lemma [39]. For robust expanders, we apply the recent Rainbow Blow-up Lemma of Glock and Joos [44] to embed a rainbow Hamilton cycle. Here, we only require the graph to have linear minimum degree. In both cases we use a key lemma that allows us to fix a partial embedding of a cycle that has a negligible effect to the rest of the graph. Finally, we combine these two results, to conclude that any Dirac graph with a $o(n)$-bounded edge colouring contains a rainbow Hamilton cycle.

As an application of Theorem 6.3, we obtain the following corollary on the vertex-degree threshold for the existence of Berge Hamilton cycles in hypergraphs. A *Berge cycle* in a hypergraph H is a sequence $v_1, e_1, v_2, e_2, v_3, \ldots, v_\ell, e_\ell$ where $v_i \in V(H)$ and $e_i \in E(H)$ are pair-wise distinct, and $\{v_i, v_{i+1}\} \subset e_i$ (addition modulo ℓ).

Corollary 6.5. *Let H be an r-uniform hypergraph on n vertices and suppose that $r = o(\sqrt{n})$. If H has minimum vertex-degree $\delta_1(H) > \binom{\lceil n/2 \rceil - 1}{r-1}$, then H contains a Berge Hamilton cycle.*

This result is best possible as for even n, the union of two complete r-uniform hypergraphs of size $n/2$ has minimum degree $\binom{n/2-1}{r-1}$ and no Berge Hamilton cycle. It also improves the bound observed in [24].

For a graph $G = (V, E)$ and $A, B \subseteq V$, we denote by $G[A]$ the subgraph induced by A in G and by $G[A, B]$ the subgraph induced by the edges between A and B in G. We use $E(A)$ and $E(A, B)$ to denote the set of edges of $G[A]$ and $G[A, B]$, respectively. We denote by $e(A) = |E(A)|$ and $e(A, B) = |E(A, B)|$. For $v \in V$, we use $N_G(x)$ to denote the set of vertices in V adjacent to x, and $d_G(x) = |N_G(x)|$. We also use $d_G(x, A)$ for the number of vertices in A that are adjacent to x. If the graph G is clear from the context, we use $N(x), d(x)$ and $d(x, A)$ instead. Finally, we will use $\delta(G)$ and $\Delta(G)$ to denote the minimum and maximum degree of G, respectively.

6.2 A trichotomy for Dirac graphs

Our proof proceeds by splitting the class of Dirac graphs into three families: robust expanders, graphs that resemble a complete bipartite graph $K_{n/2,n/2}$ and graphs that resemble the disjoint union of two complete graphs $K_{n/2}$, denoted by $2K_{n/2}$. This trichotomy was originally introduced by Kühn Lapinskas and Osthus [73]. We will state the version of this lemma from [31]. Note that this definition is very similar to the definintion of bipartite robust expanders in Chapter 4.

For $0 < \nu < 1$ and $X \subseteq V(G)$, the *ν-robust neighbourhood* of X in G is defined as

$$RN_\nu(X) := \{v \in V(G) : |N_G(v) \cap X| \geq \nu n\} .$$

Let $0 < \nu \leq \tau < 1$. A graph $G = (V, E)$ on n vertices is a *robust (ν, τ)-expander* if for every set $X \subseteq V(G)$ with $\tau n \leq |X| \leq (1-\tau)n$, we have

$$|RN_\nu(X)| \geq |X| + \nu n \ .$$

Let $0 < \gamma < 1$. A graph G on n vertices is

- γ-*close to* $K_{n/2,n/2}$ if there exists $A \subseteq V(G)$ with $|A| = \lfloor \frac{n}{2} \rfloor$ such that $e(A) \leq \gamma n^2$.
- γ-*close to* $2K_{n/2}$ if there exists $A \subseteq V(G)$ with $|A| = \lfloor \frac{n}{2} \rfloor$ such that $e(A, V(G)\backslash A) \leq \gamma n^2$.

We will use the following classification of Dirac graphs.

Lemma 6.6 (Lemma 1.3.2 in [31] for Dirac graphs)**.** *Suppose that* $0 < 1/n \ll \nu \ll \tau, \gamma < 1$ *where* $n \in \mathbb{N}$*. Let* G *be a graph on* n *vertices with* $\delta(G) \geq n/2$*. Then* G *satisfies one of the following properties:*

i) G *is* γ*-close to* $K_{n/2,n/2}$*;*

ii) G *is* γ*-close to* $2K_{n/2}$*;*

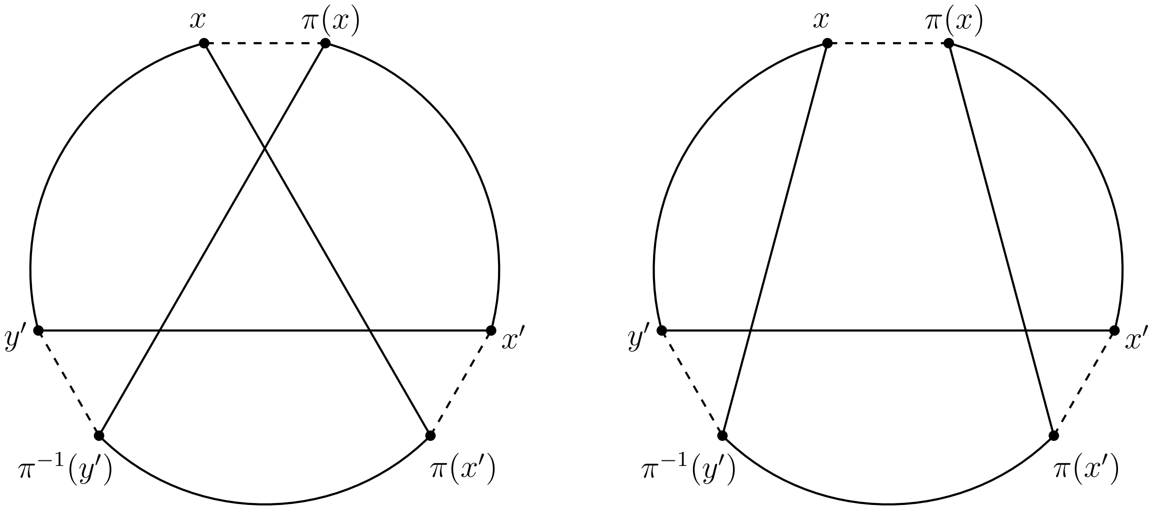
iii) G *is a robust* (ν, τ)*-expander.*

6.3 A Switching Lemma

In the previous two chapters, we introduced the connection between the existence of many local operations (switchings) for a given perfect matching, and the existence of a rainbow perfect matching. In this section, we adapt this idea to the Hamilton cycle case.

For the sake of convenience, we will define the switching operation on directed cycles. A directed cycle $\vec{H}$ on a finite set V is a spanning cycle with an orientation of the edges so every vertex has out-degree one. We denote by H the undirected cycle obtained by

Figure 6.1



Switching $s_1(\vec{H}, e, e')$ Switching $s_2(\vec{H}, e, e')$

removing the orientation of the edges in $\vec{H}$. A directed cycle defines a *successor function* $\pi : V \to V$ so $(x, \pi(x))$ is a directed edge of $\vec{H}$ for every $x \in V$. In this chapter, a *switching* is a map s that given a directed cycle $\vec{H}$ on V and edges $e \in E(H)$, $e' \notin E(H)$, assigns a directed cycle $\vec{H_0} := s(\vec{H}; e, e')$ of V such that $e' \in E(H_0)$ and $e \notin E(H_0)$.

We now define the switchings that we will use in the proofs.

Definition 6.3.1. Given a directed cycle $\vec{H}$ on V with successor function π, $e = x\pi(x) \in E(H)$ and $e' = x'y' \notin E(H)$ with x in the directed path from y' to x' induced by $\vec{H}$, we define $\vec{H_1} = s_1(\vec{H}; e, e')$ and $\vec{H_2} = s_2(\vec{H}; e, e')$ as the directed cycles that contain the directed edge (x', y') and whose undirected cycles are, respectively,

$$H_1 = (H - \{e, x'\pi(x'), \pi^{-1}(y')y'\}) + \{e', x\pi(x'), \pi^{-1}(y')\pi(x)\}$$
$$H_2 = (H - \{e, x'\pi(x'), \pi^{-1}(y')y'\}) + \{e', x\pi^{-1}(y'), \pi(x)\pi(x')\} .$$

(See Fig. 6.1 for a diagram.)

Note that $s_i(\vec{H}; e, e')$ always produces one single cycle and that there is a unique way to orient its edges to obtain a directed cycle that contains (x', y'). So s_i is a well-defined switching. Moreover, both switchings are involutions, that is to say:

$$\vec{H} = s_1(s_1(\vec{H}; e, e'); e', e)$$
$$\vec{H} = s_2(s_2(\vec{H}; e, e'); e', e) \ .$$

6.3.1 Using switchings to find rainbow Hamilton cycles

Given a graph G, a directed cycle $\vec{H}$ on $V(G)$, $e \in E(H)$ and $e' \in E(G) \setminus E(H)$, we say that $\vec{H}_0 = s_i(\vec{H}; e, e')$ is *admissible* if H_0 is a subgraph of G. Under the assumption that we have many admissible switchings for each directed Hamilton cycle of G and each edge in the cycle, we can prove that G has a rainbow Hamilton cycle using the local lemma. Here we will prove a stronger result: given a small set of edges, one can find a rainbow Hamilton cycle that contains it.

Theorem 6.7. *Suppose $1/n \ll \mu \ll \alpha \ll \beta \leq 1/2$ where $n \in \mathbb{N}$. Let G be a graph on n vertices and χ a μn-bounded colouring of $E(G)$. Let $Z \subseteq E(G)$ with $|Z| \leq \alpha n$ such that each colour of an edge in Z only appears once as the colour of an edge $e \in E(G)$. Suppose that G has at least one Hamilton cycle that contains Z. Suppose that for every directed Hamilton cycle $\vec{H}$ of G with $Z \subseteq E(H)$ and every edge $e \in E(H) \setminus Z$, there are at least βn^2 admissible switchings $s_i(\vec{H}; e, e')$ for some $e' \in E(G) \setminus E(H)$ and $i \in \{1, 2\}$. Then G has a rainbow Hamilton cycle that contains Z.*

Proof. Let $\Omega = \Omega(G, Z)$ be the set of undirected Hamilton cycles of G that contain Z, equipped with the uniform distribution. By assumption, note that $\Omega \neq \emptyset$. Let H be a Hamilton cycle chosen uniformly at random from Ω.

For each unordered pair of edges $e, f \in E(G)$ let $E(e, f) = \{e, f \in H\}$ be the event that both e and f are simultaneously in H. Let $\mathrm{supp}(E(e, f))$ be set of vertices that are incident to either e or f. Let $Q \subseteq \binom{E(G)}{2}$ be the set of unordered pairs of edges e, f with $\chi(e) = \chi(f)$, and let $q = |Q|$. Furthermore, define $Q(e) = \{f \in E(G) : \{e, f\} \in Q\}$. Consider the collection of events $\mathcal{E} = \{E(e, f) : \{e, f\} \in Q\}$.

Write $\mathcal{E} = \{E_i : i \in [q]\}$ and let $\mathcal{D}$ be the graph with vertex set $[q]$ where $i, j \in [q]$ are adjacent if and only if $\mathrm{supp}(E_i) \cap \mathrm{supp}(E_j) \neq \emptyset$.

Given $\{e, f\} \in Q$ there are at most $4n$ ways to choose an edge $e' \in E(G)$ that is incident either to e or to f, and at most μn ways to choose an edge $f' \in E(G)$ with $\chi(f') = \chi(e')$. Hence, the maximum degree of $\mathcal{D}$ is at most $d := 4\mu n^2$.

Our goal is to show that $\mathcal{D}$ is a $\mathbf{p}$-dependency graph for $\mathcal{E}$ where $\mathbf{p} = (p, p, \dots, p)$ for some suitably small $p > 0$. Given $i \in [q]$ and $S \subseteq [q] \setminus (N_{\mathcal{D}}(i) \cup \{i\})$ with $\mathbb{P}(\cap_{j \in S} {E_j}^c) > 0$, it suffices to show that (3.8) holds.

Fix $E_i = E(e_i, f_i)$ and $S \subseteq [q] \setminus (N_{\mathcal{D}}(i) \cup \{i\})$. A Hamilton cycle is S-*good* if it belongs to $\cap_{j \in S} E_j^c$. Since $\mathbb{P}(\cap_{j \in S} {E_j}^c) > 0$, there is at least one S-good Hamilton cycle that contains Z. Let $\mathcal{H} \subseteq \Omega$ be the set of S-good Hamilton cycles that contain Z and let $\mathcal{H}_0 \subseteq \mathcal{H}$ be the ones that also contain e_i and f_i.

Construct an auxiliary bipartite multigraph $\mathcal{G} = (\mathcal{H}_0, \mathcal{H} \setminus \mathcal{H}_0, E(\mathcal{G}))$, where we add an edge between $H_0 \in \mathcal{H}_0$ and $H \in \mathcal{H} \setminus \mathcal{H}_0$ for every orientation $\vec{H}_0$ of H_0 and $\vec{H}$ of H, every $k, \ell \in \{1, 2\}$ and e_i', f_i' such that

$$\vec{H} = s_k(s_\ell(\vec{H}_0; e_i, e_i'); f_i, f_i') \ .$$

By double-counting the edges of $\mathcal{G}$, we obtain

$$\delta(\mathcal{H}_0)|\mathcal{H}_0| \leq e(\mathcal{G}) \leq \Delta(\mathcal{H} \setminus \mathcal{H}_0)|\mathcal{H} \setminus \mathcal{H}_0| \ ,$$

from which we may deduce,

$$\mathbb{P}(E_i | \cap_{j \in S} E_j^c) = \frac{|\mathcal{H}_0|}{|\mathcal{H}|} \leq \frac{|\mathcal{H}_0|}{|\mathcal{H} \setminus \mathcal{H}_0|} \leq \frac{\Delta(\mathcal{H} \setminus \mathcal{H}_0)}{\delta(\mathcal{H}_0)} \ . \tag{6.1}$$

So, in order to prove (3.8) we need to bound $\Delta(\mathcal{H} \setminus \mathcal{H}_0)$ from above and $\delta(\mathcal{H}_0)$ from below.

We first bound $\Delta(\mathcal{H} \setminus \mathcal{H}_0)$ from above. Fix $H \in \mathcal{H} \setminus \mathcal{H}_0$. There are two choices for

$\vec{H}$, at most n choices for $e_i' \in E(H)$ and at most 2 choices for ℓ that yield an admissible switching and create an edge in $\mathcal{G}$. The same argument applies to f_i. It follows that $\Delta(\mathcal{H} \setminus \mathcal{H}_0) \leq 16n^2$.

In order to bound $\delta(\mathcal{H}_0)$ from below, fix $H_0 \in \mathcal{H}_0$ and choose one of the two orientations $\vec{H}_0$. Note here that not all pairs of disjoint admissible switchings for e_i and f_i, respectively, will generate an edge in $\mathcal{G}$ as it may be that the Hamilton cycle resulting from the switchings is not S-good or does not contain Z.

For $e \in \{e_i, f_i\}$, define

$$F_Z(e) = \{e' \in E(G) \setminus E(H_0) : \exists \ell \in \{1,2\} \text{ with } s_\ell(\vec{H}_0; e, e') \text{ admissible containing } Z\};$$
$$F(e) = \{e' \in E(G) \setminus \cup_{f \in E(H_0)} Q(f) : \operatorname{supp}(E_i) \cap e' = \emptyset\} \cap F_Z(e) \, .$$

Every edge $e_i' \in F_Z(e_i)$ determines at least one choice of $\ell \in \{1,2\}$ such that $s_\ell(\vec{H}_0; e_i, e_i')$ is admissible and contains Z. Moreover, if $e_i' \in F(e_i)$, then $s_\ell(\vec{H}_0; e_i, e_i')$ is S-good. The key point is that S is the intersection of events that have support disjoint from $\operatorname{supp}(E_i)$, so we only need to make sure that the colour of e_i' is not in H_0, as the other two new edges in $s_\ell(\vec{H}_0; e_i, e_i')$ are incident to $\operatorname{supp}(E_i)$.

Let us compute the size of $F(e_i)$. As the colours on edges of Z are unique amongst colours of $E(G)$, we have $e_i \notin Z$ and there are at least βn^2 choices of e_i' and $\ell \in \{1,2\}$ such that $s_\ell(\vec{H}_0; e_i, e_i')$ is admissible. From these, there are at most $8|Z|n \leq 8\alpha n^2$ switchings that do not preserve Z, so $|F_Z(e_i)| \geq (\beta/2 - 8\alpha)n^2$. There are at most μn^2 edges in $\cup_{f \in E(H_0)} Q(f)$ and at most $4n$ edges e' with $\operatorname{supp}(E_i) \cap e' \neq \emptyset$, so $|F(e_i)| \geq \beta n^2/4$.

Fix $e_i' \in F(e_i)$, let $\vec{H}_* = s_\ell(\vec{H}_0; e_i, e_i')$ and let π be the successor function in $\vec{H}_*$. If $e_i = uv$, let

$$F' = \{e \in E(G) : e \cap \{u, \pi(u), \pi^{-1}(u), v, \pi(v), \pi^{-1}(v)\} \neq \emptyset\} \cup \{e \in E(G) : e \in Q(e_i')\} \, .$$

Consider $F^*(f_i) = F(f_i) \setminus F'$ and note that for every $f_i' \in F^*(f_i)$ there exists $k \in \{1,2\}$ with $\vec{H} = s_k(\vec{H}_*; f_i, f_i')$ admissible, containing Z, S-good and not containing e_i and f_i, so $H \in \mathcal{H} \setminus \mathcal{H}_0$. Arguing as before and noting that $|F'| \leq 8n$, we have $|F^*(f_i)| \geq \beta n^2/4$. As there are two possible orientations for H_0, we conclude that $\delta(\mathcal{H}_0) \geq \beta^2 n^4/8$.

Substituting into (6.1), we obtain the desired bound

$$\mathbb{P}(E_i | \cap_{j \in S} E_j^c) \leq \frac{128}{\beta^2 n^2} =: p \ .$$

As $\mu \ll \beta$, $4pd \leq 1$ and by the lopsided version of the local lemma (Corollary 3.4) implies that the probability that a uniformly random Hamilton cycle containing Z is rainbow is positive, so there exists at least one. $\square$

6.4 A technical lemma

In this section we prove a technical lemma that we will use in the proof of our main theorem to fix a set of edges Z of the rainbow Hamilton cycle such that the graph obtained after removing edges with the same colour as Z still has a large minimum degree.

For a multiset C of $\mathbb{N}$ and $t \in \mathbb{N}$, we denote by $\mathrm{mult}(t, C)$ the multiplicity of t in C. Given a set T, we use $C \setminus^+ T$ to denote the multiset obtained by removing all elements in T from C and $C \cap^+ T$ to denote the multiset obtained by removing all elements not in T from C.

The following result is an extension of Lemma 4.9, although the proof is different.

Lemma 6.8. *Let $b, m \in \mathbb{N}$ and suppose that $1/n \ll \mu \ll \nu \ll 1/a \ll \eta, 1/b \leq 1$ where $a, n \in \mathbb{N}$. Let $C_1, \dots, C_m$ be multisets of $\mathbb{N}$ such that:*

(S1) $\nu n \leq |C_i| \leq n$, for every $i \in [m]$;

(S2) $\sum_{i=1}^m \mathrm{mult}(t, C_i) \leq \mu n$, for every $t \in \mathbb{N}$.

Let $\ell \in \mathbb{N}$ and let $U_k \subseteq \mathbb{N}$ for $k \in [\ell]$ be disjoint sets with $|U_k| = a$ and $U = \biguplus_{k=1}^{\ell} U_k$. Then, there exists $T \subseteq U$ such that:

(T1) $|T \cap U_k| \geq b$, for every $k \in [\ell]$;

(T2) $|C_i \setminus^+ T| \geq (1-\eta)|C_i|$*, for every* $i \in [m]$*.*

Proof. Let $s := \lceil \log(\mu n) \rceil$. For every $i \in [m]$ and every $j \in [s]$, define the (multi)sets

$$C_i^j = \{\{t \in C_i : 2^{-j}\mu n \leq \mathrm{mult}(t, C_i) \leq 2^{-(j-1)}\mu n\}\},$$
$$S_i^j = \{t \in C_i^j\},$$
$$S_i = \cup_{j\in[s]} S_i^j \ .$$

Let $c_i^j = |C_i^j|$, $s_i^j = |S_i^j|$, $c_i = |C_i|$ and $s_i = |S_i|$. Then, these parameters satisfy

$$2^{-j}\mu n s_i^j \leq c_i^j \leq 2^{-(j-1)}\mu n s_i^j, \tag{6.2}$$
$$\sum_{j\in[s]} c_i^j = c_i \ .$$

For every $j \in [s]$ and $u \in U$, define $n_j(u) = |\{i : u \in S_i^j\}|$. Note that

$$\sum_{j\in[s]} n_j(u) 2^{-j} \leq 1. \tag{6.3}$$

Choose δ with $1/a \ll \delta \ll \eta, 1/b$. Let T be a random subset of U obtained by including each element of U independently at random with probability δ.

A pair (i,j) is *dense* if $s_i^j \geq 2^{(j-1)/2}\mu^{-1/2}$. Let R_i be the set of $j \in [s]$ such that (i,j) is dense. The contribution of non-dense pairs is negligible; using (6.2), we have

$$\sum_{j\notin R_i} c_i^j \leq \mu n \sum_{j\notin R_i} 2^{-(j-1)} s_i^j \leq \mu^{1/2} n \sum_{j\notin R_i} 2^{-(j-1)/2} \leq \mu^{1/3} n \ . \tag{6.4}$$

For every $S \subseteq \mathbb{N}$ and $j \in [s]$ we say that $i \in [m]$ is j*-activated* by S if $|S_i^j \cap S| \geq 2\delta s_i^j$.

We define two event types that we would like T to avoid:

- Type A: for every $k \in [\ell]$, A_k is the event that $|T \cap U_k| < b$, with support, $\mathrm{supp}(A_k) = U_k$.

- Type B: for every $i \in [m]$ and $j \in [s]$ such that $j \in R_i$, B_i^j is the event that i is

j-activated by T, with support, $\text{supp}(B_i^j) = S_i^j$.

Denote by $\mathcal{E} = \{E_1, \ldots, E_q\}$ the collection of events of type A and B defined above. Let $\mathcal{D}$ be the *dependency graph of* $\mathcal{E}$, the graph with vertex set $[q]$ constructed by adding an edge between $i, j \in [q]$ if and only if $\text{supp}(E_i) \cap \text{supp}(E_j) \neq \emptyset$. We will apply the weighted version of the local lemma (Corollary 3.6) to show that there exists a choice of T that avoids all events in $\mathcal{E}$.

Let $p = e^{-2} \leq 1/4$. We first bound the probabilities of the events in $\mathcal{E}$. Let $w(A_k) := \delta a/8$ and $w(B_i^j) := \delta s_i^j/8$. Let $X_k = |T \cap U_k|$. Note that X_k is binomially distributed with mean δa. By Chernoff inequality (see e.g. Corollary 2.3 in [58]) with $t = 3/4$, we have

$$\mathbb{P}(A_k) \leq \mathbb{P}(X_k \leq \delta a/4) \leq e^{-(9/32)\delta a} \leq e^{-\delta a/4} = p^{w(A_k)} \, . \tag{6.5}$$

Let $Y_i^j = |S_i^j \cap T|$, which is stochastically dominated by a binomial random variable with mean δs_i^j. Recall that $B_i^j = \{Y_i^j \geq 2\delta s_i^j\}$. Chernoff's inequality with $t = 1$ implies

$$\mathbb{P}(B_i^j) \leq \mathbb{P}(Y_i^j \geq 2\delta s_i^j) \leq e^{-\delta s_i^j/4} = p^{w(B_i^j)} \, . \tag{6.6}$$

To apply the local lemma, it suffices to check that for every $E \in \mathcal{E}$, we have

$$\sum_{A_k \sim E} (2p)^{w(A_k)} + \sum_{B_i^j \sim E} (2p)^{w(B_i^j)} \leq \frac{w(E)}{2} \, .$$

Since two events are adjacent only if their supports intersect, for each $u \in U$ we will compute the contribution of the events whose support contains u.

As the sets U_k are disjoint, there is only one event of Type A whose support intersects u. Using $2p \leq e^{-1}$ and (6.5), we have

$$\sum_{\text{supp}(A_k) \ni u} (2p)^{w(A_k)} \leq e^{-\delta a/8} \, .$$

For events of Type B, $j \in R_i$ implies $s_i^j \geq 2^{(j-1)/2}\mu^{-1/2}$, and since $\mu \ll \delta \ll 1$, we obtain

$$(2p)^{w(B_i^j)} \leq e^{-\delta s_i^j/8} \leq e^{-\delta 2^{(j-7)/2}\mu^{-1/2}} \leq \mu 2^{-j} .$$

Recall that, for every $j \in [s]$, u appears in $n_j(u)$ sets S_i^j. It follows from (6.3) and (6.6) that

$$\sum_{\mathrm{supp}(B_i^j)\ni u} (2p)^{w(B_i^j)} \leq \sum_{j\in[s]} n_j(u)(2p)^{\min\{w(B_i^j):\, j\in R_i\}} \leq \mu \sum_{j\in[s]} n_j(u)2^{-j} \leq \mu .$$

Observe that for any type of event $E \in \mathcal{E}$, we have $|\mathrm{supp}(E)| = 8\delta^{-1}w(E)$. Thus,

$$\sum_{A_k\sim E} (2p)^{w(A_k)} + \sum_{B_i^j\sim E} (2p)^{w(B_i^j)} = 8\delta^{-1}w(E)(e^{-\delta a/8} + \mu) \leq \frac{w(E)}{2} .$$

By the weighted form of the local lemma, we obtain the existence of a set T that avoids all the events in $\mathcal{E}$. The set T satisfies (T1) as it avoids A_k for $k \in [\ell]$. Let us show that (T2) follows from the events of type B.

Using (6.2) twice, it follows that for each $i \in [m]$, $j \in R_i$, we have

$$|C_i^j \cap^+ T| \leq \mu n 2^{-(j-1)}|S_i^j \cap T| \leq \mu n 2^{-(j-1)} \cdot 2\delta s_i^j \leq 4\delta c_i^j$$

By combining this with (6.4), for $i \in [m]$ we obtain

$$|C_i \cap^+ T| = \sum_{j\in R_i} |C_i^j \cap^+ T| + \sum_{j\notin R_i} |C_i^j \cap^+ T| \leq 4\delta \sum_{j\in R_i} c_i^j + \sum_{j\notin R_i} c_i^j \leq 4\delta c_i + \mu^{1/3}n \leq \eta|C_i| ,$$

where we used that $|C_i| \geq \nu n$ and $\mu \ll \nu \ll \delta \ll \eta \ll 1$. Thus, (T2) is satisfied. $\square$

6.5 Graphs which are close to $2K_{n/2}$

In this section, we prove Theorem 6.3 for graphs that resemble the disjoint union of two complete graphs.

Theorem 6.9. *Suppose $1/n \ll \mu \ll \gamma \ll 1$ where $n \in \mathbb{N}$. Let G be a graph on n vertices with $\delta(G) \geq n/2$ that is γ-close to $2K_{n/2}$. Let χ be a μn-bounded colouring of $E(G)$. Then G has a rainbow Hamilton cycle.*

6.5.1 ε-superextremal two-cliques

Note that in a graph which is γ-close to $2K_{n/2}$ we have no real control of the minimum degree within the partition. We can however make some small adjustments to the partition of G to get large minimum degree.

Definition 6.5.1. A graph G on n vertices is an ε-superextremal two-clique if there exists a partition $V(G) = A \uplus B$ with the following properties:

(A1) $||A| - |B|| \leq \varepsilon n$;

(A2) $d_G(a, A) \geq (1/2 - \varepsilon)n$ for all but at most εn vertices $a \in A$;

(A3) $d_G(a, A) \geq (1/4 - \varepsilon)n$ for all vertices $a \in A$;

(A4) $d_G(b, B) \geq (1/2 - \varepsilon)n$ for all but at most εn vertices $b \in B$;

(A5) $d_G(b, B) \geq (1/4 - \varepsilon)n$ for all vertices $b \in B$.

Lemma 6.10. *Suppose $1/n \ll \gamma \ll \varepsilon \ll 1$ where $n \in \mathbb{N}$. Let G be a graph on n vertices with $\delta(G) \geq n/2$ that is γ-close to $2K_{n/2}$. Then, G is an ε-superextremal two-clique with partition $V(G) = A \uplus B$. Moreover, $G[A, B]$ either has minimum degree at least 1 or the minimum degree from either A or B is at least 2.*

Proof. As G is γ-close to $2K_{n/2}$ there is a partition of $V(G)$ into parts A_0, B_0 of size $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$, respectively, such that $e(A_0, B_0) \leq \gamma n^2$. Define the sets

$$X_A = \{v \in A_0 : d_G(v, A_0) \leq n/4\} \qquad X_B = \{v \in B_0 : d_G(v, B_0) \leq n/4\}$$

Choose $\gamma \ll \delta \ll \varepsilon$. Note that as G has minimum degree at least $n/2$, $2e(A_0) \geq n|A_0| - \gamma n^2 \geq n^2/2 - \gamma n^2$, from which we deduce $|X_A|, |X_B| \leq \delta n$. Define $A = (A_0 \setminus X_A) \cup X_B$, $B = (B_0 \setminus X_B) \cup X_A$ and (A1)-(A5) follow immediately.

If $|A| = |B|$, then $G[A, B]$ has minimum degree at least 1, and otherwise, assuming $|A| < |B|$, A has minimum degree to B at least 2. $\square$

As G is γ-close to $2K_{n/2}$, it is an ε-superextremal two-clique with partition $V(G) = A \uplus B$. Consider a μn-bounded colouring χ of $E(G)$ with $1/n \ll \mu \ll \varepsilon$. We now choose a rainbow set of edges Z. By the second part of the previous lemma, we can find two vertex-disjoint edges f and f' between A and B with distinct colours. Henceforth, we set $Z = \{f, f'\}$.

In order to find a rainbow Hamilton cycle containing Z using Theorem 6.7, it will be more convenient to work with a spanning subgraph of G. Let $\hat{G}$ be the graph obtained from G by deleting all the edges in $E(A, B) \setminus Z$ and all the edges with the same colour as an edge in Z. It is easy to see that $\hat{G}$ is a 2ε-superextremal two-clique and that

(C1) $E_{\hat{G}}(A, B) = Z$;

(C2) each edge in Z has a unique colour in $E(\hat{G})$.

6.5.2 Finding the switchings

The next step is to show that Theorem 6.7 applies to the case of ε-superextremal two-cliques with Z given in the previous section. First we show that there is at least one Hamilton cycle. We will use the following sufficient condition for the existence of Hamilton cycles.

Theorem 6.11 (Chvátal [23]). *Let G be a graph on m vertices with degree sequence $d_1 \leq d_2 \leq \ldots \leq d_m$. Suppose that for $k \in \{1, \ldots, m/2\}$, if $d_k \leq k$ then $d_{m-k} \geq m - k$. Then G has a Hamilton cycle.*

The following result shows that there is at least one Hamilton cycle containing Z.

Lemma 6.12. *Let G be an ε-superextremal two-clique with partition $V(G) = A \uplus B$ and let f, f' be two vertex-disjoint edges between A and B. Then G has a Hamilton cycle which includes f and f'.*

Proof. Suppose that $f = ab$ and $f' = a'b'$ where $a, a' \in A$. It suffices to show that there is a spanning path in A from a to a' and similarly in B.

To prove this consider the graph G_A obtained from G by removing all vertices in B and adding an auxiliary vertex x which we connect only to a and a'. Vertex x has degree two, up to at most εn vertices have degree at least $n/4 - \varepsilon n$ and the remainder have degree at least $n/2 - \varepsilon n > |A|/2$. So the degree sequence of G_A satisfies $d_k > k$ for each $k \leq v(G_A)/2$. Thus, we can use Theorem 6.11 on G_A to obtain a cycle H_A that spans $A \cup \{x\}$. Since x has degree two, H_A contains a path P_A spanning A with endpoints a and a'. The same argument yields a spanning path P_B for B. Hence, G has a Hamilton cycle obtained by concatenating P_A and P_B using edges f, f'. $\square$

Let us show that there are many switchings in ε-superextremal two-cliques, for every edge not in Z.

Lemma 6.13. *Suppose $1/n \ll \mu \ll \varepsilon \ll 1$ where $n \in \mathbb{N}$. Let $\hat{G}$ be an ε-superextremal two-clique with partition $V(G) = A \uplus B$ satisfying (C1), where $Z = \{f, f'\}$ is composed by two vertex-disjoint edges between A and B. Let $\vec{H}$ be a directed Hamilton cycle of G. For every $e \in E(H) \setminus Z$, there are at least $n^2/300$ admissible switchings $s_i(\vec{H}; e, e')$ for some $e' \in E(G) \setminus E(H)$ and $i \in \{1, 2\}$.*

Proof. Suppose that $f = ab$ and $f' = a'b'$ where $a, a' \in A$. As G satisfies (C1) and $e \notin Z$, without loss of generality, we may assume that $e \in E(A)$ and that $\vec{H}[A]$ induces a directed path P_A from a to a'. Let π be the successor function of $\vec{H}$ and consider the total order $<_{\vec{H}}$ on A that satisfies $u <_{\vec{H}} \pi(u)$ for all $u \in A \setminus \{a'\}$. Write $e = u\pi(u)$ for $u \in A$. Define $X = N(u) \setminus \{a, b, a', b'\}$ and $Y = N(\pi(u)) \setminus \{a, b, a', b'\}$. Let X^- be the first $\lfloor |X|/2 \rfloor$ vertices in X with respect to $<_{\vec{H}}$ and $X^+ = X \setminus X^-$. Define Y^- and Y^+ analogously. We split the proof in two cases:

Case 1: $x <_{\vec{H}} y$ for all $x \in X^-, y \in Y^+$.

Define

$$X^{--} = \{x \in X^- \colon x \leq_{\vec{H}} u\} \qquad X^{-+} = \{x \in X^- \colon u <_{\vec{H}} x\}$$

$$Y^{+-} = \{y \in Y^+ \colon y \leq_{\vec{H}} u\} \qquad Y^{++} = \{y \in Y^+ \colon u <_{\vec{H}} y\}$$

Clearly, either $|X^{--}| \geq \lfloor |X|/4 \rfloor$ or $|X^{-+}| \geq \lfloor |X|/4 \rfloor$ and let X^* be the largest of the two sets. Similarly, define Y^*. By the hypothesis of the case and depending on the position of u in P_A, either $X^{-+} = \emptyset$ or $Y^{+-} = \emptyset$, so $(X^*, Y^*) \neq (X^{-+}, Y^{+-})$. This leaves the following cases for (X^*, Y^*):

- Case 1.1: If $(X^*, Y^*) = (X^{--}, Y^{++})$, then we set $X_0 = \pi(X^*)$ and $Y_0 = \pi^{-1}(Y^*)$. For a directed edge e' from Y_0 to X_0, $s_2(\vec{H}; e, e')$ is admissible.
- Case 1.2: If $(X^*, Y^*) \neq (X^{--}, Y^{++})$, then we set $X_0 = \pi^{-1}(X^*)$ and $Y_0 = \pi(Y^*)$. For a directed edge e' from X_0 to Y_0, $s_1(\vec{H}; e, e')$ is admissible.

It suffices to count the edges between X_0 and Y_0. Let $X_1 = \{x \in X_0 \colon d_G(x, A) \geq (1/2 - \varepsilon)n\}$ and define Y_1 analogously. By (A2) and (A3), $|X_1|, |Y_1| \geq (1/16 - 2\varepsilon)n$. Using (A1) and (A2) again, we may also deduce that each vertex in X_1 is adjacent to all but at most $2\varepsilon n$ of the vertices in Y_1. Hence, $e(X_0, Y_0) \geq e(X_1, Y_1) \geq (1/16 - 2\varepsilon)(1/16 - 4\varepsilon)n^2 \geq n^2/300$.

Case 2: $y <_{\vec{H}} x$ for all $y \in Y^-, x \in X^+$.

The proof is almost identical to the one for Case 1, up to defining the sets X_0 and Y_0 properly in terms of most common ordering of $x \in X^+$, $y \in Y^-$ and u, and choosing the correct switching type in each case.

Hence, we obtain at least $n^2/300$ admissible switchings $s_i(\vec{H}; e, e')$. $\square$

We finally prove the main theorem of this section.

Proof of Theorem 6.9. Let $\gamma \ll \varepsilon \ll 1$. By Lemma 6.10 and the discussion after it, G has a subgraph $\hat{G}$ which is a 2ε-superextremal two-clique with partition $V(\hat{G}) = A \uplus B$ that satisfies (C1)-(C2) for $Z = \{f, f'\}$, where f, f' are two vertex-disjoint edges between A and B. By Lemma 6.12, there exists at least one Hamilton cycle in $\hat{G}$ that contains Z. Finally, Lemma 6.13 implies that for every directed Hamilton cycle H of $\hat{G}$ and every $e \in E(H) \setminus Z$ there are at least $n^2/300$ admissible switchings. Thus we may apply Theorem 6.7 to the graph $\hat{G}$ to obtain a rainbow Hamilton cycle (that contains Z). As $\hat{G}$ is a spanning subgraph of G, the desired result follows. $\square$

6.6 Graphs which are close to $K_{n/2,n/2}$

In this section, we prove Theorem 6.3 for graphs that resemble the complete bipartite graph.

Theorem 6.14. *Suppose $1/n \ll \mu \ll \gamma \ll 1$ where $n \in \mathbb{N}$. Let G be graph on n vertices with $\delta(G) \geq n/2$ that is γ-close to $K_{n/2,n/2}$. Let χ be a μn-bounded colouring of $E(G)$. Then G has a rainbow Hamilton cycle.*

6.6.1 $(\alpha, \varepsilon, \nu)$-superextremal bicliques

Let G be a graph that is γ-close to $K_{n/2,n/2}$ with partition $V(G) = A \uplus B$. As in the previous section, we could have vertices in A with no neighbours in B. We can make small adjustments to the partition in order to guarantee a minimum degree condition.

Definition 6.6.1. A graph G on n vertices is an $(\alpha, \varepsilon, \nu)$-*superextremal biclique* if there exists a partition $V(G) = A \uplus B$ with the following properties:

(B1) $0 \leq |B| - |A| \leq \alpha n$;

(B2) $d(a, B) \geq (1/2 - \varepsilon)n$ for all but at most αn vertices $a \in A$;

(B3) $d(a, B) \geq \nu n$ for all vertices $a \in A$;

(B4) $d(b, A) \geq (1/2 - \varepsilon)n$ for all but at most αn vertices $b \in B$;

(B5) $d(b, A) \geq (1/4 - \varepsilon)n$ for all vertices $b \in B$;

(B6) $d(b, B) \leq 2\nu n$ for all vertices $b \in B$, unless $|A| = \lfloor n/2 \rfloor$.

Lemma 6.15. *Suppose $1/n \ll \gamma \ll \alpha, \varepsilon \ll \nu \ll 1$ where $n \in \mathbb{N}$. Let G be a graph on n vertices with $\delta(G) \geq n/2$ that is γ-close to $K_{n/2,n/2}$. Then G is an $(\alpha, \varepsilon, \nu)$-superextremal biclique.*

Proof. Let $V(G) = A_0 \uplus B_0$ be the partition given by the fact that G is γ-close to $K_{n/2,n/2}$. Define

$$X_A = \{a \in A_0 : d(a, B_0) \leq (1/4 - \gamma)n\} \qquad X_B = \{b \in B_0 : d(b, A_0) \leq (1/4 - \gamma)n\}.$$

Choose $\gamma \ll \delta \ll \alpha, \varepsilon$. If $a \in X_A$, $d(a, A_0) \geq (1/4 + \gamma)n$ and, as $e(A_0) \leq \gamma n^2$, $|X_A| \leq \delta n$. As there are at least $|A_0|n/2 - \gamma n^2$ edges from A_0 to B_0, we may similarly deduce that $|X_B| \leq \delta n$. Now, let $A_1 = (A_0 \setminus X_A) \cup X_B$ and $B_1 = (B_0 \setminus X_B) \cup X_A$. Assume that $|B_1| \geq |A_1|$. If not, we shall swap their labels. Let $Y_B = \{b \in B_1 : d(b, B_1) \geq 2\nu n\}$. Note that it is entirely possible for Y_B to be very large (it could even be all of B_1 in some cases), so in the case that $|Y_B| \geq (|B_1| - |A_1|)/2$ select an arbitrary set $Y'_B \subseteq Y_B$ of size $\lfloor (|B_1| - |A_1|)/2 \rfloor$ and otherwise let $Y'_B = Y_B$. Define $A = A_1 \cup Y'_B$, $B = B_1 \setminus Y'_B$.

We claim that this partition satisfies all the properties of a superextremal biclique partition. Property (B1) follows from the fact that we swap sets of size at most δn

between A_0 and B_0 to obtain A_1 and B_1, that we assume $|B_1| \geq |A_1|$ and that $|Y'_B| \leq \lfloor (|B_1| - |A_1|)/2 \rfloor$. Properties (B2) and (B4) follow similarly to the bounds on the sizes of X_A and X_B. Properties (B3), (B5) and (B6) can all be deduced similarly from the definitions of X_A, X_B and Y'_B.

□

6.6.2 Finding the protected set Z

The main difference between this extremal case and the previous one, is that here we will need to protect a set of edges Z of up to linear size in order to balance both parts of the partition. If we choose Z greedily as before, when removing edges with the same colour as edges in Z, we will be deleting up to a quadratic number of edges, and thus it will be possible to isolate a vertex. We will use the technical lemma from Section 6.4 to ensure that we can choose Z, so deleting edges with the same colour will not have a significant effect on the degree of each vertex.

Lemma 6.16. *Suppose* $1/n \ll \mu \ll \alpha, \varepsilon \ll \nu \ll 1$ *where* $n \in \mathbb{N}$. *Let* G *be an* $(\alpha, \varepsilon, \nu)$*-superextremal biclique with partition* $V(G) = A \uplus B$ *and denote* $m = |B| - |A|$. *Let* χ *be a* μn*-bounded colouring of* $E(G)$. *Then* $G[B]$ *has a rainbow matching of size at least* $m/20\nu$.

Proof. We choose a matching M greedily. At each step, add an arbitrary edge of $E(G[B])$ to M which is not incident to M and has a colour which is not the same as the colour of any edge in M. By (B3) and as $\alpha \ll \nu$, observe that $d(b, B) \geq m/2$ for every $b \in B$, so $e(B) \geq m|B|/2$. If $m = 1$, then any edge in $E(G[B])$ forms the desired matching. Otherwise $|A| < \lfloor n/2 \rfloor$ and by (B6), for each edge we add to M there are at most $4\nu n$ edges incident to it in $G[B]$ and at most μn edges with the same colour, including the edge itself. Thus the choice of this edge removes at most $4\nu n + \mu n$ edges which we could have added to M in subsequent steps. As $d(b, B) \geq m/2$ for each $b \in B$, there are at least $m|B|/4$ edges in $G[B]$. Hence, when we can no longer add any more edges M has

size at least
$$|M| \geq \frac{m|B|/4}{(4\nu+\mu)n} \geq \frac{m}{20\nu}.$$

□

We will use Lemma 6.8 to select a partial matching of size $|B|-|A|$ from the matching obtained in the previous lemma. The edges of the matching will form the protected set Z.

Lemma 6.17. *Suppose $1/n \ll \mu \ll \alpha, \varepsilon \ll \nu \ll \eta \ll 1$ where $n \in \mathbb{N}$. Let G be a $(\alpha, \varepsilon, \nu)$-superextremal biclique on n vertices with $\delta(G) \geq n/2$ and partition $V(G) = A \uplus B$. Let χ be a μn-bounded colouring of $E(G)$. Then, there exist a matching M in B of size $|B|-|A|$ and a spanning subgraph $\hat{G}$ of G which is an $(\alpha, \eta, \nu/2)$-superextremal biclique with the same partition as G satisfying*

(D1) $E_{\hat{G}}(A) = \emptyset$ and $E_{\hat{G}}(B) = E(M)$;

(D2) $\max\{d_{\hat{G}}(a, B), d_{\hat{G}}(b, A)\} \geq (1/2-\eta)n$ for all $a \in A$, $b \in B$ with $ab \in E(\hat{G})$;

(D3) each edge in M has a unique colour in $E(\hat{G})$.

Proof. Let M_0 be the rainbow matching obtained from Lemma 6.16 and set $U = \{\chi(e) : e \in M_0\}$. Let $\nu \ll 1/a \ll \eta$. Assume that a divides $|U|$ (otherwise we can delete some elements from U so it holds) and let $\ell = |U|/a$. Choose an arbitrary partition $U = U_1 \uplus \cdots \uplus U_\ell$ with $|U_k| = a$ for $k \in [\ell]$. For a vertex $v \in V(G)$, let C_v be the multiset of colours on the edges in $E(A, B)$ incident to v. Properties (B2)-(B5) imply that $\nu n \leq |C_v| \leq n$ and the properties of the colouring imply that $\sum_{v \in V(G)} \mathrm{mult}(t, C_v) \leq 2\mu n$ for $t \in \mathbb{N}$. We apply Lemma 6.8 to this setup with the parameters as in the following table.

Use	2μ	$\eta/2$	$(\lvert B\rvert-\lvert A\rvert)/\ell$
In place of	μ	η	b

Let T_0 be the set of colours in U given by the lemma and note that $|T_0| \geq |B|-|A|$. Select an arbitrary subset T of T_0 of size $|B|-|A|$. Define M as the matching with edge set

$\{e \in E(M_0) : \chi(e) \in T\}$ and note that M is rainbow as M_0 was. Let $\hat{G}$ be the subgraph obtained from G by deleting all the edges $e \notin E(M)$ with either $e \in E(A) \cup E(B)$ or $\chi(e) \in T$, so it satisfies (D1) and (D3), and after that, deleting all edges between vertices of degree at most $(1/2 - \eta)n$. As $\varepsilon \ll \nu \ll \eta \ll 1$, Properties (B1)-(B6), (D1) and (T2), imply that $\hat{G}$ is an $(\alpha, \eta, \nu/2)$-superextremal biclique. As we deleted edges between low degree vertices, $\hat{G}$ also satisfies (D2).

□

6.6.3 Finding the switchings

In this section we will show that the graph $\hat{G}$ satisfies the hypothesis of Theorem 6.7 with $Z = M$. First, we show that there exists at least one Hamilton cycle that contains Z. We will use the following sufficient condition for the existence of Hamilton cycles in bipartite graphs:

Theorem 6.18. *(Moon and Moser [93]) Let $G = (R \cup S, E)$ be a balanced bipartite graph on $2m$ vertices with $R = \{r_1, \ldots, r_m\}$ and $S = \{s_1, \ldots, s_m\}$ that satisfies $d(r_1) \leq \ldots \leq d(r_m)$ and $d(s_1) \leq \ldots \leq d(s_m)$. Suppose that for every $k \in \{1, \ldots, m/2\}$, we have $d(r_k) > k$ and $d(s_k) > k$. Then G has a Hamilton cycle.*

Lemma 6.19. *Suppose $1/n \ll \alpha \ll \nu \ll \eta \ll 1$ where $n \in \mathbb{N}$. Let G be an (α, η, ν)-superextremal biclique on n vertices with partition $V(G) = A \uplus B$ and M a matching in $G[B]$ of size $|B| - |A|$. Let G be an (α, η, ν)-superextremal biclique on n vertices with partition $V(G) = A \uplus B$ and M a matching in $G[B]$ of size $|B| - |A|$. Then G has a Hamilton cycle that contains M.*

Proof. First note that any pair of vertices in B can be connected in G by many paths of length at most 4. As $|E(M)| \leq \alpha n$, we can connect the vertices of M with disjoint paths of length at most 4, obtaining a path P of length at most $5|E(M)| \leq 5\alpha n$ which contains $E(M)$ and has endpoints $b, b' \in B$. Note that P uses $|E(M)| + 1$ more vertices in B than in A. Let $\tilde{G}$ be the balanced bipartite graph obtained by deleting all the edges in $E(A) \cup E(B)$ and all the internal vertices of P, and adding an auxiliary vertex x to A only adjacent to b and b'. Every vertex in $\tilde{G}$ different from x satisfies the properties (B2)-(B5) of an $(\alpha, \eta+5\alpha, \nu-5\alpha)$-superextremal biclique, so we have control on the minimum degrees. In particular, the hypothesis of Theorem 6.18 are satisfied and we deduce that $\tilde{G}$ has a Hamilton cycle $\tilde{H}$. As w has degree two, $\tilde{H}$ contains the edges xb and xb'. The subgraph H of G obtained by replacing the path bxb' by P in $\tilde{H}$ is a Hamilton cycle of G that contains M. $\square$

Next lemma shows that in any Hamilton cycle H containing M, that there are a large number of admissible switchings for any edge of H which is not in M.

Lemma 6.20. *Suppose that $1/n \ll \mu \ll \alpha \ll \beta \ll \nu \ll \eta \ll 1$ where $n \in \mathbb{N}$. Let $\hat{G}$ be an (α, η, ν)-superextremal biclique on n vertices with partition $V(G) = A \uplus B$. Let M be a matching in $\hat{G}[B]$ with $|E(M)| \leq \alpha n$ and set $Z = E(M)$. Suppose G and M satisfy (D1)-(D2). Then for every directed Hamilton cycle $\vec{H}$ of $\hat{G}$ and every edge $e \in E(H) \setminus Z$, there are at least βn^2 admissible switchings $s_i(\vec{H}; e, e')$ for some $e' \in E(G) \setminus E(H)$ and $i \in \{1, 2\}$.*

The proof of this lemma is very similar to the one of Lemma 6.13 and we will omit some arguments that are analogous.

Proof. By (D1) and since $e \notin Z$, we may assume that $e = ab$ for some $a \in A$ and $b \in B$. As $ab \in E(\hat{G})$, by (B3) and (D2) we will assume that $d_{\hat{G}}(a, B) \geq \nu n$ and $d_{\hat{G}}(b, A) \geq (1/2 - \eta)n$, the symmetric case can be proved analogously.

Define $X = N(a) \setminus V(Z)$ and $Y = N(b) \setminus B$. As in the proof of Lemma 6.13, we can find $X_0 \subseteq X$, $Y_0 \subseteq Y$ with $|X_0| \geq \lfloor |X|/4 \rfloor \geq (\nu-\alpha)n/4$ and $|Y_0| \geq \lfloor |Y|/4 \rfloor \geq (1/8-\eta)n$ such that for every directed e' from X_0 to Y_0 (or from Y_0 to X_0), $s_i(\vec{H}; e, e')$ is admissible for some $i \in \{1, 2\}$. Letting $X_1 \subseteq X_0$ and $Y_1 \subseteq Y_0$ be the vertices of degree at least $(1/2-\eta)n$, by (B3) and (B5) and since $\alpha \ll \nu$, we get $|X_1| \geq (\nu/8)n$ and $|Y_1| \geq (1/8 - 2\eta)n$. As $|A| \leq n/2$ by (B1), it follows that $e(X_1, Y_1) \geq (1/8 - 3\eta)n|X_1| \geq \beta n^2$, as desired.

$\square$

We now have all the ingredients to prove the existence of a rainbow Hamilton cycle.

Proof of Theorem 6.14. Let $\mu \ll \alpha, \varepsilon \ll \gamma \ll \beta \ll \nu \ll \eta \ll 1$. By Lemma 6.15, G is an $(\alpha, \varepsilon, \nu)$-superextremal biclique with partition $V = A \uplus B$. By Lemma 6.17, we can choose a rainbow matching M in $G[B]$ of size $|B| - |A|$, denote $Z = E(M)$, and an $(\alpha, \eta, \nu/2)$-superextremal subgraph $\hat{G}$ of G satisfying (D1)-(D3). Lemma 6.19 ensures that $\hat{G}$ has at least one Hamilton cycle containing Z. Applying Theorem 6.20 to $\hat{G}$, we obtain that the hypothesis of Theorem 6.7 are satisfied. Thus $\hat{G}$ has a rainbow Hamiltonian cycle and so does G. □

6.7 Robust expanders

In this section we prove our main theorem for robust expanders.

Theorem 6.21. *Suppose $1/n \ll \mu \ll \nu \ll \tau \ll \gamma < 1$ with $n \in \mathbb{N}$. Let G be graph on n vertices with $\delta(G) \geq \gamma n$ that is a robust (ν, τ)-expander. Let χ be a μn-bounded colouring of $E(G)$. Then G has a rainbow Hamilton cycle.*

6.7.1 Regularity Lemma and rainbow blow-up lemma

We first introduce the regularity concepts and tools we will use in the proof. For $r \in \mathbb{N}$, let $[r]_0 = [r] \cup \{0\}$. We will use the shorthand $x = (a \pm b)c$ to mean $x \in (ac - bc, ac + bc)$. For X, Y disjoint sets of vertices, we define their *density* as $d(X, Y) = \frac{e(X,Y)}{|X||Y|}$. For X, Y disjoint sets of vertices, we define their *density* as $d(X, Y) = \frac{e(X,Y)}{|X||Y|}$. A bipartite graph on $A \cup B$ with all edges between A and B is called a *pair* and we denote it by (A, B). A pair (A, B) is *ε-regular* if for each $X \subseteq A$, $Y \subseteq B$ such that $|X| > \varepsilon|A|$ and $|Y| > \varepsilon|B|$, we have $|d(X, Y) - d(A, B)| < \varepsilon$. A pair (A, B) is *(ε, d)-super-regular* if it is ε-regular, $d(a) = (d \pm \varepsilon)|B|$ for each $a \in A$ and $d(b) = (d \pm \varepsilon)|A|$ for each $b \in B$ where here $d(a)$ means the degree of a. We will use the following version of the regularity lemma.

Theorem 6.22 (Szemerédi's Regularity Lemma [113]). *Let $M, M', n \in \mathbb{N}$ and suppose $1/n \ll 1/M \ll \varepsilon, 1/M' \le 1$ and $d > 0$. For any graph G on n vertices, there exists a partition $(V_i)_{i\in[r]_0}$ of $V(G)$ with $r \in (M', M)$ and a spanning subgraph G' of G such that:*

- *$|V_0| \le \varepsilon n$;*
- *$|V_i| = |V_j|$ for all $i, j \in [r]$;*
- *$d_{G'}(v) \ge d_G(v) - (\varepsilon + d)n$ for all $v \in V(G)$;*
- *$e(G'[V_i]) = 0$ for all $i \in [r]$;*
- *For all $i \ne j \in [r]$, the pair (V_i, V_j) in G' is either empty or ε-regular with density at least d.*

We call $(V_i)_{i\in[r]_0}$ an *(ε, d)-regular partition of G*. The sets $V_1, \dots, V_r$ are the *clusters* and V_0 is the *exceptional set*. The *reduced graph R associated to* $(V_i)_{i\in[r]_0}$ is the graph with vertices $V_1, \dots, V_r$ in which V_iV_j is an edge if and only if the pair (V_i, V_j) is (ε, d)-super-regular in G'.

A standard tool to embed bounded degree spanning subgraphs in G is the Blow-Up Lemma of Komlós, Sárközy and Szemerédi [68]. This lemma has been recently extended by Glock and Joos [44] to embed rainbow spanning subgraphs with bounded degrees in bounded colourings. We first introduce some notation.

Definition 6.7.1. A tuple $(H, G, R, (X_i)_{i\in[r]_0}, (V_i)_{i\in[r]_0})$ is a *blow-up instance* if the following hold:

- H and G are graphs, $(X_i)_{i\in[r]_0}$ is a partition of $V(H)$ into independent sets, $(V_i)_{i\in[r]_0}$ is a partition of $V(G)$ and $|X_i| = |V_i|$ for all $i \in [r]_0$;
- R is a graph with $V(R) = \{V_1, \dots, V_r\}$ and for $i \ne j \in [r]$ the graph $H[X_i, X_j]$ is empty if $V_iV_j \notin E(R)$.

Definition 6.7.2. The pair (A, B) is *lower (ε, d)-super-regular* if the following hold:

- $d(S,T) \geq d - \varepsilon$, for all $S \subseteq A$, $T \subseteq B$ with $|S| \geq \varepsilon|A|$, $|T| \geq \varepsilon|B|$;

- $d(a) \geq (d-\varepsilon)|B|$, for each $a \in A$;

- $d(b) \geq (d-\varepsilon)|A|$, for each $b \in B$.

A blow-up instance $(H, G, R, (X_i)_{i\in[r]_0}, (V_i)_{i\in[r]_0})$ is *lower* (ε, d)*-super-regular* if for all $ij \in E(R)$, $G[V_i, V_j]$ is lower (ε, d)-super-regular.

The blow-up lemma embeds H into G such that each X_i is embedded in V_i. In applications, one may want to restrict the candidates in V_i for each vertex in X_i. We will encode these restrictions using candidacy graphs.

Definition 6.7.3. For each $i \in [r]$, a *candidacy graph* A^i is a pair (X_i, V_i). A blow-up instance $(H, G, R, (X_i)_{i\in[r]_0}, (V_i)_{i\in[r]_0})$ with candidacy graphs $(A^i)_{r\in[r]}$ is *lower* (ε, d)*-super-regular* if $(H, G, R, (X_i)_{i\in[r]_0}, (V_i)_{i\in[r]_0})$ is lower (ε, d)-super-regular and A^i is lower (ε, d)-super-regular for each $i \in [r]$.

The main idea of the rainbow blow-up lemma is that, given a pre-embedding of X_0 into V_0 satisfying certain conditions, one can extend it to $V(H)$ to find a rainbow copy of H in G.

Definition 6.7.4. Given a blow-up instance $(H, G, R, (X_i)_{i\in[r]_0}, (V_i)_{i\in[r]_0})$ with candidacy graphs $(A^i)_{r\in[r]}$ and a colouring χ of $E(G)$, a bijection $\phi_0 : X_0 \to V_0$ is *feasible* if the following conditions hold:

(F1) for all $x_0 \in X_0$, $j \in [r]$ and $x \in N_H(x_0) \cap X_j$, we have $N_{A^j}(x) \subseteq N_G(\phi_0(x_0))$;

(F2) for all $j \in [r]$, $x \in X_j$, $v \in N_{A^j}(x)$ and distinct $x_0, x_0' \in N_H(x) \cap X_0$, we have $\chi(\phi_0(x_0)v) \neq \chi(\phi_0(x_0')v)$.

Informally speaking, (F1) ensures that every candidate image for x is a neighbour of $\phi_0(x_0)$ in G and (F2) ensures that the set of edges in the copy of H in G between a candidate image for x and V_0 is rainbow.

We are now able to state the rainbow blow-up lemma for bounded colourings:

Theorem 6.23 (Rainbow Blow-Up Lemma (Lemma 5.1 in [44])). *Let $n, \Delta, r \in \mathbb{N}$ and suppose $1/n \ll \mu, \varepsilon \ll d, 1/\Delta$ and $\mu \ll 1/r$. Suppose that $(H, G, R, (X_i)_{i\in[r]_0}, (V_i)_{i\in[r]_0})$ with candidacy graphs $(A^i)_{i\in[r]}$ is a lower (ε, d)-super-regular blow-up instance and assume further that*

(RB1) $\Delta(R), \Delta(H) \leq \Delta$;

(RB2) $|V_i| = (1 \pm \varepsilon)n/r$ *for all* $i \in [r]$

(RB3) for all $i \in [r]$, at most $(2\Delta)^{-4}|X_i|$ vertices in X_i have a neighbour in X_0.

Let χ be a μn-bounded colouring of $E(G)$. Suppose that there exists a feasible bijection $\phi_0 : X_0 \to V_0$. Then there exists a rainbow embedding ϕ of H into G which extends ϕ_0 such that $\phi(x) \in N_{A^i}(x)$ for all $i \in [r]$ and $x \in X_i$.

6.7.2 Collection of short paths

In order to apply the rainbow blow-up lemma, first we need to find a blow-up instance for robust expanders. The following result states that the reduced graph of a robust expander, is a also robust expander.

Lemma 6.24 (Lemma 14 in [79]). *Suppose $1/n \ll \varepsilon \ll d \ll \nu, \tau, \eta \leq 1$ where $n \in \mathbb{N}$. Let G be a robust (ν, τ)-expander graph on n vertices with $\delta(G) \geq \eta n$. Let R be the reduced graph of G associated to an (ε, d)-super-regular partition of it. Then R is a robust $(\nu/2, 2\tau)$-expander with $\delta(R) \geq (\eta - d - 2\varepsilon)|R|$.*

We also use the following result on the existence of Hamilton cycles in robust expanders.

Lemma 6.25 (Lemma 16 in [79]). *Suppose $1/n \ll \nu \ll \tau \ll \eta \leq 1$ where $n \in \mathbb{N}$. Let G be a robust (ν, τ)-expander with $\delta(G) \geq \eta n$. Then G has a Hamilton cycle.*

Lemmas 6.24 and 6.25 are stated for directed graphs, but they can also be applied to undirected graphs G by considering the digraph obtained from G by replacing each edge by arcs in both directions.

Henceforth, consider the hierarchy of parameters

$$1/n \ll \varepsilon_1, 1/M' \ll \varepsilon_2 \ll \varepsilon_3 \ll d_2 \ll d_1 \ll \nu \ll \tau \ll \eta < 1$$

and let G be a robust (ν, τ)-expander with $\delta(G) \geq \eta n$. Let $(V_i)_{i\in[r]_0}$ be an $(\varepsilon_1/4, d_1+2\varepsilon_1)$-regular partition of G and R be its associated reduced graph, where $r = |R| \geq M'$. If $r = |R|$ is odd, we can add all vertices of V_r to the exceptional set V_0, and the reduced graph will still have the same properties with slightly different parameters. Thus, without loss of generality we may assume that r is even. By Lemmas 6.24 and 6.25, R has a Hamilton cycle. We may add at most $(\varepsilon_1/2)n$ vertices from each vertex class to V_0 such that the pairs defining edges in R are (ε_1, d_1)-super-regular. Relabel the clusters of the super-regular partition so they follow the cyclic order. Let M be the matching of R formed by the pairs $V_{2i-1}V_{2i}$ for $i \in [r/2]$. Abusing notation, we also allow M to denote the involution on $V(R)$ defined by $M(V_{2i-1}) = V_{2i}$ for $i \in [r/2]$.

We will connect the vertices of V_0 to the rest of the graph by short rainbow paths, constructing a feasible pre-embedding $\phi_0\colon X_0 \to V_0$ so we can apply the rainbow blow-up lemma. We select the paths in such a way that we maintain the balance between pairs of clusters from M, so that upon removal of these paths, these pairs form balanced bipartite graphs.

Definition 6.7.5. Let G be a graph and $(V_i)_{i\in[r]_0}$ a partition of $V(G)$. Let M be the matching formed by the pairs $V_{2i-1}V_{2i}$ for $i \in [r/2]$. A *balanced path* for $v \in V_0$ of length $2k$ is a path $P = u_{-1}u_0u_1 \dots u_{2k-1}$ such that,

- $u_0 = v$ and $u_j \notin V_0$ for all $j \in \{-1, 1, 2, \dots, 2k-1\}$;
- $u_{-1} \in V_i$ and $u_{2k-1} \in M(V_i)$, for some $i \in [r]$;
- $|V(P) \cap V_{2i}| = |V(P) \cap V_{2i-1}|$, for every $i \in [r/2]$.

The next lemma shows that we can find a large number of balanced paths of length $2k$ that only intersect in V_0 and that use different colours. This will allow us to obtain a partial embedding of a rainbow Hamilton cycle of G.

Lemma 6.26. *Let $n, M', t \in \mathbb{N}$ and suppose*

$$1/n \ll \mu \ll \varepsilon_1, 1/M' \ll \varepsilon_2 \ll d_2 \ll d_1 \ll \nu \ll \tau, 1/t \ll \eta \leq 1.$$

Let G, $(V_i)_{i \in [r]_0}$, R, M be as above with $\delta(G) \geq \eta n$. Let χ be a μn-bounded colouring of $E(G)$. Then, there exists $\mathcal{P} = \cup_{v \in V_0} \mathcal{P}(v)$, where $\mathcal{P}(v) = \{P_1(v), \ldots, P_t(v)\}$ is a collection of t balanced paths of length $2k := 2\lceil 2/\nu \rceil$ for v satisfying

(P1) $|V(\mathcal{P}) \cap V_i| \leq \varepsilon_2 n/r$, for each $i \in [r]$;

(P2) $P_i(v)$ and $P_j(v')$ are vertex-disjoint, unless $v = v'$, in which case $V(P_i(v)) \cap V(P_j(v')) = \{v\}$;

(P3) $\mathcal{P}$ is rainbow in χ.

Proof. Let G' be the spanning subgraph of G obtained from Lemma 6.22. By Lemma 6.24, R and G' are robust $(\nu/2, 2\tau)$-expanders. For $v \in V_0$, we define $N_R^*(v) = \{V_i \in V(R) : d_{G'}(v, V_i) \geq d_1 n/r\}$. Note that $|N_R^*(v)| \geq (\eta - 2d_1 - 2\varepsilon_1)r \geq \eta r/2$ follows immediately from the regularity lemma. For $X \subseteq V(R)$, we define $J_R(X) := M(RN_{\nu/2}(X))$ and note that $|J_R(X)| \geq |X| + (\nu/2)r$. Thus, $J_R^k(M(N_R^*(v))) = V(R)$.

Now to each $v \in V_0$ we will assign sets $U_{-1}(v), U_1(v), U_2(v), \ldots, U_{2k-1}(v)$ with $U_j(v) \in V(R)$ such that there are many balanced paths $u_{-1}, v, u_1, u_2, \ldots, u_{2k-1}$ with $u_j \in U_j(v)$. Among them, we will find the collection $\mathcal{P}$ of paths satisfying the conditions of the lemma, via an application of the local lemma.

As $|M(N_R^*(v))| \geq \nu r/2$ and $|V_0| \leq \varepsilon_1 n$, we can find a partition $(V_0^{i_0})$ of V_0 such that $V_{i_0} \in N_R^*(v)$ for every $v \in V_0^{i_0}$ and $|V_0^{i_0}| \leq (2\varepsilon_1/\nu)n/r$. For each $v \in V_0^{i_0}$, set $U_{-1}(v) = V_{i_0}$ and $U_{2k-1}(v) = M(V_{i_0})$. Next, we inductively refine this partition. Since $U_{-1}(v) \in J_R^k(M(N_R^*(v)))$ then $U_{2k-1}(v) \in RN_{\nu/2}(J_R^{k-1}(M(N_R^*(v))))$ and there are at least $\nu r/2$ choices of $U_{2k-2} \in J_R^{k-1}(M(N_R^*(v)))$ such that $U_{2k-2}U_{2k-1}(v) \in E(R)$. Hence, there exists a partition $(V_0^{i_0,i_1})$ that refines $(V_0^{i_0})$ satisfying $|V_0^{i_0,i_1}| \leq \left(\frac{2}{\nu r}\right)^2 \varepsilon_1 n$ and we can set

$U_{2k-2}(v) = V_{i_1}$ and $U_{2k-3}(v) = M(V_{i_1})$ for every $v \in V_0^{i_0,i_1}$. Similarly, we proceed to form a partition, $(V_0^{\mathbf{i}})$ of V_0 where $\mathbf{i} = (i_0, i_1, \ldots, i_{k-1})$ such that $i_j \in [r]$ for each j. This partition satisfies $|V_0^{\mathbf{i}}| \leq \left(\frac{2}{\nu r}\right)^k \varepsilon_1 n$ and for each $v \in V_0^{\mathbf{i}}$, $V_{i_0} \in N_R^*(v)$ and $V_{i_j} \in J_R^{k-j}(M(N_R^*(v)))$ for $1 \leq j \leq k-1$.

Finally, for each $v \in V_0^{\mathbf{i}}$, we define $U_{-1}(v) = V_{i_0}$, $U_{2k-1}(v) = M(V_{i_0})$ and for $j \geq 1$, $U_{2(k-j)}(v) = V_{i_j}$ and $U_{2(k-j)-1}(v) = M(V_{i_j})$. This choice of clusters satisfies

(i) $(U_j(v), U_{j+1}(v))$ are (ε_1, d_1)-regular pairs for each $1 \leq j \leq 2k-2$.

(ii) $d_G(v, U_{\pm 1}(v)) \geq d_1 n/r$;

(iii) any path $P = u_{-1} v u_1 u_2 \ldots u_{2k-1}$ with $u_j \in U_j(v)$ is balanced.

We can bound the multiplicity of each cluster V_i:

$$|\{v \in V_0 :\ V_i \in \{U_{-1}(v), U_1(v), U_2(v), \ldots, U_{2k-1}(v)\}\}| \leq 2 \sum_{i=1}^{k} r^{i-1} \left(\frac{2}{\nu r}\right)^i \varepsilon_1 n \leq \frac{\varepsilon_2}{t} \frac{n}{r} \tag{6.7}$$

Consider the following weakening of (P3):

(P3') $\mathcal{P}(v)$ is rainbow in χ, for every $v \in V_0$.

We can greedily construct a collection of paths $\mathcal{P}$ satisfying (P1), (P2) and (P3'). For each $v \in V_0$, we will select t paths P for $\mathcal{P}(v)$ of the form $P = u_{-1} v u_1 u_2 \ldots u_{2k-1}$ with $u_j \in U_j(v)$, so P is balanced of length $2k$. By (6.7), $\mathcal{P}$ satisfies (P1). By (i) and (ii), while constructing a new path, at any time, there are at least $(d_1 - 2\varepsilon_1)n/r$ choices for $u_j \in U_j(v)$ which has degree at least $(d_1 - \varepsilon_1)n/r$ to $U_{j+1}(v)$, for $1 \leq j \leq 2k-2$. By (P1), at most $\varepsilon_2 n/r$ of them have been already used in another path of $\mathcal{P}$, and by the properties of χ, at most $2kt\mu n$ of them would create an edge with a colour already used in another path of $\mathcal{P}(v)$. Since $\varepsilon_2 \ll d_1$ and $k\mu t \ll d_1/r$, we can select $\mathcal{P}$ satisfying (P2) and (P3').

Given the sets $U_{-1}(v), U_1(v), U_2(v), \ldots, U_{2k-1}(v)$ for each $v \in V_0$, let Ω be the uniform probability space over all possible $\mathcal{P} = \cup_{v \in V_0} \mathcal{P}(v)$, where $\mathcal{P}(v) = \{P_1(v), \ldots, P_t(v)\}$ and

$P_i(v)$ is a balanced path P of length $2k$ of the form $P = u_{-1}vu_1u_2\dots u_{2k-1}$ and $u_j \in U_j(v)$, that satisfies (P1), (P2) and (P3'). We will use the lopsided version of the local lemma to find $\mathcal{P} \in \Omega$ satisfying (P3). For the rest of the proof, $\mathcal{P}$ will be a collection of paths chosen uniformly at random from Ω.

A pair (P_1, P_2) of paths is *bad* if their union is not rainbow. For every bad pair, define the event $E(P_1, P_2) = \{P_1, P_2 \in \mathcal{P}\}$. Two events $E(P_1, P_2)$ and $E(P_3, P_4)$ are *dependent* if $V(P_1 \cup P_2) \cap V(P_3 \cup P_4) \neq \emptyset$.

To bound how many events depend on $E(P_1, P_2)$, we count the number of events $E(P_3, P_4)$ such that $w \in V(P_3 \cup P_4)$, for a given $w \in V$. Select first a pair of edges e, f with $\chi(e) = \chi(f)$ that belong to $P_3 \cup P_4$, and note that they cannot both belong to the same path by (P3'). If either e or f are incident to w, then there are at most μn^2 choices for them and we must pick at most $4k-2$ additional vertices to form $P_3 \cup P_4$. Otherwise, there are at most μn^3 choices for e and f but we only need to choose at most $4k-3$ additional vertices. Hence in both cases there are at most μn^{4k} choices for $P_3 \cup P_4$. As any event involves at most $4k+2$ vertices, there are at most $D := 2(4k+2)\mu n^{4k}$ events which depend on $E(P_1, P_2)$.

Next we find $p > 0$ such that for every bad pair (P_1, P_2) we have

$$\mathbb{P}(E(P_1, P_2)| \cap_{E \in S} E^c) \leq p$$

where S is any subset of events which do not depend on $E(P_1, P_2)$ and $\mathbb{P}(\cap_{E \in S} E^c) > 0$. We do this by a simple switching argument. Let $\mathcal{F} = \{\mathcal{P} \in \Omega : \mathcal{P} \in \cap_{E \in S} E^c\}$ and $\mathcal{F}_0 = \{\mathcal{P} \in \mathcal{F} : \mathcal{P} \in E(P_1, P_2)\}$.

If $\mathcal{P}_0 \in \mathcal{F}_0$, we say that $\mathcal{P} \in \mathcal{F} \setminus \mathcal{F}_0$ is obtained by *path-resampling* if there exists $P_1' \neq P_1$ and $P_2' \neq P_2$ such that $\mathcal{P} = (\mathcal{P}_0 \cup \{P_1', P_2'\}) \setminus \{P_1, P_2\}$. Note that P_1' and P_2' have to be chosen so that $\mathcal{P}$ satisfies (P1), (P2) and (P3').

Construct an auxiliary bipartite graph $\mathcal{G}$ with bipartition $(\mathcal{F}_0, \mathcal{F} \setminus \mathcal{F}_0)$. Add an edge from $\mathcal{P}_0 \in \mathcal{F}_0$ to $\mathcal{P} \in \mathcal{F} \setminus \mathcal{F}_0$ for every path-resampling that transforms $\mathcal{P}_0$ into $\mathcal{P}$. As in

Theorem 6.7 we may deduce that

$$\mathbb{P}(E(P_1, P_2)| \cap_{E\in S} E^c) \leq \frac{\Delta(\mathcal{F} \setminus \mathcal{F}_0)}{\delta(\mathcal{F}_0)} := p \ .$$

Thus it suffices to bound the degrees in $\mathcal{G}$. Denote by $v_1 \in V(P_1)\cap V_0$ and $v_2 \in V(P_2)\cap V_0$ the unique vertices in the intersection of the paths and the exceptional set.

Suppose first that $\mathcal{P} \in \mathcal{F} \setminus \mathcal{F}_0$. To add P_1 and P_2 by path-resampling, we need to choose one path in $\mathcal{P}(v_1)$ and one in $\mathcal{P}(v_2)$ to remove. Hence, $\Delta(\mathcal{F} \setminus \mathcal{F}_0) \leq t^2$.

Suppose now that $\mathcal{P}_0 \in \mathcal{F}_0$ and let us count the number of choices for P'_1, P'_2 that give a collection $\mathcal{P}$ in $\mathcal{F} \setminus \mathcal{F}_0$ by path-resampling. To form $P'_1 = u_{-1}v_1u_1 \dots u_{2k-1}$ we must choose $u_j \in U_j(v_1)$. By (ii), for each u_{-1} and u_1 we have at least d_1n/k choices. By (i), for $1 \leq j \leq 2k-3$ and for each choice of u_j, there are at least $(d_1 - 2\varepsilon_1)n/r$ choices for u_{j+1} with degree at least $(d - \varepsilon_1)n/r$ to $U_{j+2}(v_1)$. There are also at least $(d_1 - \varepsilon_1)n/r$ choices for u_{2k-1}. Condition (P1) is clearly satisfied for any choice of P'_1. To verify that we satisfy (P2), P'_1 must intersect $\mathcal{P}_0$ only in v_1, and to satisfy (P3') it should avoid the colours in $\mathcal{P}_0(v_1)$. We have $|V(\mathcal{P}_0)| \leq (2k+1)t|V_0| \leq (2k+1)t\varepsilon_1 n$ and χ has at most $2kt$ different colours in $\mathcal{P}_0(v_1)$ forbidding a total of $2kt\mu n$ vertices for each choice. As $\varepsilon_1, \mu \ll d_1/(krt)$, it follows that there are at least $(d_1n/2r)^{2k}$ choices for P'_1. The argument for P'_2 is analogous. We chose P'_1 and P'_2 such that path-resampling satisfies $\mathcal{P} \in \Omega$, but it also holds that $\mathcal{P} \in \cap_{E\in S}E^c$, as all the paths participating in S are vertex-disjoint with $\{v_1, v_2\}$, but P'_1 and P'_2 are not. So $\delta(\mathcal{F}_0) \geq (\frac{d_1n}{2r})^{4k}$.

We conclude that $p \leq t^2(\frac{2r}{d_1n})^{4k}$ and, as $\mu \ll 1/r, 1/t, d_1$, we have $4pD \leq 1$, and Corollary 3.4 implies that there is collection $\mathcal{P} \in \Omega$ satisfying (P3). $\square$

6.7.3 Proof of Theorem 6.21

Lemma 6.26 provides a rainbow collection of paths that will allow us to attach vertices in the exceptional set to the rest of the graph. However, by using an arbitrary set of paths, we could be using all the colours incident to a vertex. As in the extremal case, we will select a subset of paths such that removing edges with the same colour will have a negligible effect in the degrees of the graph.

With the quantifiers set above, let G be a robust (ν,τ)-expander on n vertices with $\delta(G) \geq \eta n$, $(V_i)_{i\in[r]_0}$ be an (ε_1, d_1)-regular partition which is lower (ε_1, d_1)-super-regular for edges in M the matching (V_{2i-1}, V_{2i}) for $i \in [r/2]$. Let χ be a μn-bounded colouring of $E(G)$. For the clarity of exposition, we split the proof into a number of parts.

<u>The collection of paths $\mathcal{P}^*$:</u> Let $\mathcal{P} = \cup_{v\in V_0}\mathcal{P}(v)$ be the collection of balanced paths of length $2k$ given by Lemma 6.26. Define a new colouring χ' of $E(G)$ by merging some of the colour classes of χ as follows. For each $v \in V_0$ and $i \in [t]$, add a new colour $c(i, v)$. If $e \in E(G)$ satisfies $\chi(e) \in \chi(E(P_i(v)))$ for some $v \in V_0$ and $i \in [t]$, then $\chi'(e) = c(i, v)$; otherwise, $\chi'(e) = \chi(e)$. As $\mathcal{P}$ is rainbow, this gives a well-defined colouring which is $2k\mu n$-bounded.

We will use Lemma 6.8 to select a set of paths $\mathcal{P}^*$ from $\mathcal{P}$, one for each $v \in V_0$. For each $u \in V_{2i-1} \cup V_{2i}$, let C_u be the multiset of colours on edges incident to u in (V_{2i-1}, V_{2i}). Let $N = |V_{2i}| = |V_{2i-1}|$ and note that $N \geq (1-\varepsilon_1)n/r$. As (V_{2i-1}, V_{2i}) is lower (ε_1, d_1)-super-regular, we have $(d_1/2)N \leq |C_u| \leq N$. Moreover, $\sum_u \mathrm{mult}(c, C_u) \leq 4k\mu n \leq 8k\mu r N$ for any colour c. For each $v \in V_0$, let $U_v = \{c(i, v)\}_{i\in[t]}$ and note that $|U_v| = t$ and that the sets U_v are disjoint. Choose $1/t \ll \eta_0 \leq 1$.

We apply Lemma 6.8 to this setup with the following parameters:

Use	$8k\mu r$	$d_1/2$	η_0	$\lvert V_0\rvert \leq 2\varepsilon_1 rN$	t	1	n	N
In place of	μ	ν	η	ℓ	a	b	m	n

So we obtain a set T containing at least one element from each U_v and such that $|C_u \setminus^+ T| \geq (1-\eta_0)|C_u|$ for each $u \in V \setminus V_0$. We may assume that T contains exactly one element from each U_v, as by removing elements $|C_u \setminus^+ T|$ will only increase. Thus, we obtain a subcollection $\mathcal{P}^* = \{P^*(v)\}_{v\in V_0}$ with $P^*(v) \in \mathcal{P}(v)$ satisfying the following. Let G^* be the graph obtained from G by removing the edges $e \notin E(\mathcal{P}^*)$ with $\chi(e) \in \chi(E(\mathcal{P}^*))$. Then $\delta(G^*[V_{2i-1}, V_{2i}]) \geq (1-2\eta)d_1 n/r$ for every $i \in [r/2]$.

The graph $\hat{G}$: Let $\mathcal{P}_*$ be a rainbow collection of edges in G^*, where for every $i \in [r/2]$, we select an arbitrary edge $a_{2i}b_{2i+1}$ from $G^*[V_{2i}, V_{2i+1}]$ (working modulo r) with $a_{2i}, b_{2i+1} \notin V(\mathcal{P}^*)$. This is possible as there are at least $(d_1/2r^2)n^2$ edges in $G[V_{2i}, V_{2i+1}]$, at most $4k\varepsilon_1\mu n^2$ have been deleted in G^* and, by (P1), at most $(2\varepsilon_2/r^2)n^2$ are incident to $V(\mathcal{P}^*)$. Let $\hat{G}$ be the graph obtained from G^* by removing all edges $e \notin E(\mathcal{P}_*)$ with $\chi(e) \in \chi(E(\mathcal{P}_*))$, which satisfies $\delta(\hat{G}[V_{2i-1}, V_{2i}]) \geq (1 - 2\eta_0 - r^2\mu/2)d_1 n/r \geq (d_1/2 - \varepsilon_1)|V_{2i}|$. In particular, (V_{2i-1}, V_{2i}) is lower $(\varepsilon_1, d_1/2)$-super-regular in $\hat{G}$.

Constructing the Hamilton cycle: Let H be a Hamilton cycle on n vertices. We now construct a partition $(\hat{V}_i)_{i\in[r]_0}$ of $V(\hat{G})$ and a copy of H in $\hat{G}$. Consider the exceptional set $\hat{V}_0$ obtained from V_0 by adding all the internal vertices in the paths in $\mathcal{P}^*$. Note that $|\hat{V}_0| \leq 2k\varepsilon_1 n \leq \varepsilon_2 n$. Further, define $\hat{V}_i = V_i \setminus \hat{V}_0$.

The vertices in $V(\mathcal{P}^*) \setminus \hat{V}_0$ come in pairs, corresponding to endpoints of the balanced paths in consecutive sets V_{2i-1} and V_{2i}. For $i \in [r/2]$, let $\ell_i = |(V(\mathcal{P}^*) \setminus \hat{V}_0) \cap V_{2i}|$. For $j \in [\ell_i]$, let a^j_{2i-1}, b^j_{2i} denote the endpoints of the j-th path with endpoints in V_{2i-1} and V_{2i}.

It is not difficult to check that the union of the paths in $\hat{\mathcal{P}}$, $\mathcal{P}^*$ and $\mathcal{P}_*$ forms a copy of H on $\hat{G}$.

The blow-up instance $(\hat{\mathcal{P}}, \hat{G}, M, (X_i)_{i\in[r]_0}, (\tilde{V}_i)_{i\in[r]_0})$: Define a new exceptional set,

$$\tilde{V}_0 = \hat{V}_0 \cup \{a_i, b_i, a^j_i, b^j_i : i \in [r], j \in \mathbb{N}\}.$$

Further, define $\tilde{V}_i = \hat{V}_i \setminus \tilde{V}_0$. All edges in $\mathcal{P}^* \cup \mathcal{P}_*$ are within the exceptional set $\tilde{V}_0$ and, by the way we have constructed each P_i, all edges of $\hat{\mathcal{P}}$ are either in one of the pairs in M or between the exceptional set and one of the clusters. The partition $(\tilde{V}_i)_{i\in[r]_0}$ of $V(\hat{G})$ induces a partition $(X_i)_{i\in[r_0]}$ of $V(H) = V(\hat{\mathcal{P}})$ and $(\hat{\mathcal{P}}, \hat{G}, M, (X_i)_{i\in[r]_0}, (\tilde{V}_i)_{i\in[r]_0})$ is a blow-up instance. Note that we consider $\hat{\mathcal{P}}$ instead of H as X_0 is an independent set in $\hat{\mathcal{P}}$ but not in H.

The blow-up instance is lower (ε_3, d_2)-super-regular: It is sufficient to show that the bipartite graphs $\hat{G}[\tilde{V}_{2i-1}, \tilde{V}_{2i}]$ are lower (ε_3, d_2)-super-regular. This is simply inherited from the $(\varepsilon_1, d_1/2)$-super-regularity of $\hat{G}[V_{2i-1}, V_{2i}]$ by noting that $|V_j \setminus \tilde{V}_j| \leq |V(\mathcal{P}^* \cup \mathcal{P}_*) \cap V_j| \leq \varepsilon_2 n/r + 1$, by (P1).

The pre-embedding ϕ_0 and the candidacy graphs A^i: We consider the identity map $\phi_0 : X_0 \to V_0$ as the pre-embedding of the exceptional set for $\hat{\mathcal{P}}$ into $\hat{G}$. Then we construct the candidacy graphs in accordance with the pre-embedding. For $x \in X_i$, if $x_0 x \in E(H)$ for some $x_0 \in X_0$, we let $N_{A^i}(x) = N_{\hat{G}}(\phi_0(x_0)) \cap \tilde{V}_i$. Otherwise, let $N_{A^i}(x) = \tilde{V}_i$. As no vertex in $V(\hat{\mathcal{P}}) \setminus X_0$ has more than one neighbour in X_0, A^i is well-defined. We check that A^i is lower (ε_3, d_2)-super-regular. Note first that it has minimum degree at least $d_2|V_i|$. As $|V(\mathcal{P}) \cap V_i| \leq \varepsilon_2 n/r$, in A^i we have deleted at most $2\varepsilon_2|\tilde{V}_i|^2$ edges from the complete bipartite graph with support in $(X_i, \tilde{V}_i)$. For any $S \subseteq X_i$, $T \subseteq \tilde{V}_i$ each of size at least $\varepsilon_3|\tilde{V}_i|$, we have

$$e(S,T) \geq |S||T| - 2\varepsilon_2|\tilde{V}_i|^2 \geq |S||T| - \frac{2\varepsilon_2}{(\varepsilon_3)^2}|S||T| \geq (d_2 - \varepsilon_3)|S||T| \, .$$

Hence, the blow-up instance $(\hat{\mathcal{P}}, \hat{G}, M, (X_i)_{i \in [r]_0}, (\tilde{V}_i)_{i \in [r]_0})$ with candidacy graphs A^i is lower-(ε_3, d_2)-super-regular.

The pre-embedding is feasible: Property (F1) follows immediately from the definition of ϕ_0 and A^i. Property (F2) is also satisfied as no vertex in $V(H) \setminus X_0$ has more than one neighbour in X_0.

Applying the rainbow blow-up lemma: We apply Theorem 6.23 with parameters μ, $\Delta = 2$, $\varepsilon = \varepsilon_3$ and $d = d_2$. Conditions (RB1) and (RB3) clearly hold and condition (RB2) holds as $|V_i| = (1 \pm \varepsilon_1)n/r$ and $|V_i \setminus \tilde{V}_i| \leq 2\varepsilon_2|V_i|$. Hence $\hat{G}$ has a rainbow copy of $\hat{\mathcal{P}}$. By construction of $\hat{G}$, the colours in $\mathcal{P}^* \cup \mathcal{P}_*$ are disjoint from the colours used in $E(\hat{G})$. It follows that G contains a rainbow Hamilton cycle.

6.8 Proof of Theorems 6.3 and 6.4 and Corollary 6.5

In this section we give a proof of Theorem 6.3.

Proof of Theorem 6.3. Let $\mu \ll \nu \ll \tau, \gamma < 1$. By Lemma 6.6, G is either a robust (ν, τ)-expander or is γ-close to either $2K_{n/2}$ or to $K_{n/2,n/2}$. Combining Theorems 6.9, 6.14 and 6.21, G has a rainbow Hamilton cycle.

□

Proof of Theorem 6.4. Choose any integer function $k = k(n) \to \infty$ such that $k = o(n)$ and $k(n)$ is even. Consider $G = (V, E)$ a graph on $|V| = n$ vertices with $V = A \cup B$ where $|A| = \lfloor n/2 \rfloor - k$ and $|B| = \lceil n/2 \rceil + k$. The edge set E is constructed by adding all edges between A and B and choosing any k-regular graph in $G[B]$. It is easy to check that G is a Dirac graph.

Consider a colouring of E that assigns $2k - 1$ colours to the edges in $G[B]$, keeping the size of the colour classes as similar as possible, and a distinct colour to each edge in $E(A, B)$. Note that any Hamilton cycle in G must use at least $2k$ edges from $G[B]$, therefore there is no rainbow Hamilton cycle in G. There are $k(\lceil n/2 \rceil + k)/2$ edges in $G[B]$, so, the each colour class has size at most $\lceil \frac{k(\lceil n/2 \rceil + k)}{2(2k-1)} \rceil < \mu n$, for large enough n, concluding the proof. □

Proof of Corollary 6.5. We construct a graph G on $V(H)$ by adding an edge uv if and only if there is an edge in H which contains both u and v. As $\delta_1(H) > \binom{\lceil n/2 \rceil - 1}{r-1}$, then G has minimum degree at least $n/2$ and hence is a Dirac graph. Construct a colouring χ of $E(G)$ by letting $\chi(uv) = e$ for some arbitrary edge $e \in E(H)$ containing both u and v, for each edge $uv \in E(G)$. This colouring is clearly $\binom{r}{2} = o(n)$-bounded. We may apply Theorem 6.3 and deduce that G has a rainbow Hamilton cycle $v_1, v_2 \ldots, v_n$. Then,

$$v_1, e_1 = \chi(v_1 v_2), v_2, e_2 = \chi(v_2 v_3), v_3, \ldots, v_n, e_n = \chi(v_n v_0) ,$$

is a Berge cycle, as the fact that the cycle is rainbow in G implies that all edges are distinct and, by the definition of χ, $\{v_i, v_{i+1}\} \subseteq e_i$. $\square$

6.9 Concluding Remarks

Over the previous three chapters we have shown that at the minimum degree threshold for containing certain subgraphs, we can in fact find rainbow copies of such subgraphs. Furthermore, many other such results can be deduced from the rainbow blow-up lemma of Glock and Joos [44]. However in all of these results, it is the case that the maximum degree is essentially constant in comparison to the size of the host graph. (Our results from Chapter 5 as well as [44] do allow the maximum degree to grow incredibly slowly with the size of the host graph.)

A question of interest would be to ask whether there are any collections of graphs with (quickly) growing maximum degree where we can find rainbow copies of these graphs at their extremal threshold. This would allow one to generalise the work of Böttcher, Kohayakawa and Procacci [14] away from the setting of a complete host graph.

CHAPTER 7

STRONG COMPONENTS IN RANDOM DIGRAPHS

7.1 Introduction

Consider the random digraph model $D(n, p)$ where each of the $n(n-1)$ possible edges is included with probability p independently of all others. This is analogous to the Erdős-Renyi random graph $G(n, p)$ in which each edge is again present with probability p independently of all others. McDiarmid [89] showed that due to the similarity of the two models, it is often possible to couple $G(n, p)$ and $D(n, p)$ to compare the probabilities of certain properties.

In the random graph $G(n, p)$ the component structure is well understood. In their seminal paper [37], Erdős and Rényi proved that for $p = c/n$ the largest component of $G(n, p)$ has size $O(\log(n))$ if $c < 1$, is of order $\Theta(n^{2/3})$ if $c = 1$, and has linear size when $c > 1$ (all with high probability). This threshold behaviour is known as the double jump. If we zoom in further around the critical point, $p = 1/n$ and consider $p = (1 + \varepsilon(n))/n$ such that $\varepsilon(n) \to 0$ and $|\varepsilon(n)|^3 n \to \infty$, Bollobás [10] proved the following theorem for $|\varepsilon| > (2\log(n))^{1/2} n^{-1/3}$,which was extended to the whole range described above by Łuczak [83].

Theorem 7.1 ([10, 83])**.** *Let $np = 1 + \varepsilon$, such that $\varepsilon = \varepsilon(n) \to 0$ but $n|\varepsilon|^3 \to \infty$, and $k_0 = 2\varepsilon^{-2} \log(n|\varepsilon|^3)$.*

i) If $n\varepsilon^3 \to -\infty$ *then a.a.s.* $G(n,p)$ *contains no component of size greater than* k_0*.*

ii) If $n\varepsilon^3 \to \infty$ *then a.a.s.* $G(n,p)$ *contains a unique component of size greater than* k_0*. This component has size* $2\varepsilon n(1+o(1))$*.*

Within the critical window itself i.e. $p = n^{-1} + \lambda n^{-4/3}$ with $\lambda \in \mathbb{R}$, the size of the largest component $\mathcal{C}_1$ is not tightly concentrated as it is for larger p. Instead, there exists a random variable $X_1 = X_1(\lambda)$ such that $|\mathcal{C}_1|n^{-2/3} \to X_1$ in distribution as $n \to \infty$. Much is known about the distribution of X_1, in fact the vector $\mathbf{X} = (X_1, \ldots, X_k)$ of normalised sizes of the largest k components i.e. $X_i = |\mathcal{C}_i|n^{-2/3}$ converges in distribution to the vector of longest excursion lengths of an inhomogeneous reflected Brownian motion by a result of Aldous [3]. In a more quantitative setting where one is more interested about behaviour for somewhat small n, Nachmias and Peres [94] proved the following (similar results may be found in [98, 105]).

Theorem 7.2 ([94])**.** *Suppose* $0 < \delta < 1/10$*,* $A > 8$ *and* n *is sufficiently large with respect to* A, δ*. Then if* $\mathcal{C}_1$ *is the largest component of* $G(n, 1/n)$*, we have*

i) $\mathbb{P}(|\mathcal{C}_1| < \lfloor \delta n^{2/3} \rfloor) \leq 15\delta^{3/5}$

ii) $\mathbb{P}(|\mathcal{C}_1| > An^{2/3}) \leq \frac{4}{A}e^{-\frac{A^2(A-4)}{32}}$

Note we have only stated the version of their theorem with $p = n^{-1}$ for clarity but it holds for the whole critical window. Of course, there are a vast number of other interesting properties of $\mathcal{C}_1$, see [1, 57, 85] for a number of examples.

In the setting of $D(n,p)$, one finds that analogues of many of the above theorems still hold. When working with digraphs, we are interested in the strongly connected components which we will often call the *components*. Note that the weak component structure of $D(n,p)$ is precisely the component structure of $G(n, 2p-p^2)$. For $p = c/n$, Karp [60] and Łuckzak [84] independently showed that for $c < 1$ all components are of size $O(1)$ and when $c > 1$ there is a unique complex component of linear order and every other component is of size $O(1)$ (a component is complex if it has more edges than vertices). The range $p = (1+\varepsilon)/n$ was studied by Łuczak and Seierstad [86] who were able to show the following result which can be viewed as a version of Theorem 7.1 for $D(n,p)$,

Theorem 7.3 ([86])**.** *Let $np = 1+\varepsilon$, such that $\varepsilon = \varepsilon(n) \to 0$.*

i) If $n\varepsilon^3 \to -\infty$ then a.a.s. every component of $D(n,p)$ is an isolated vertex or a cycle of length $O(1/|\varepsilon|)$.

ii) If $n\varepsilon^3 \to \infty$ then a.a.s. $D(n,p)$ contains a unique complex component of size $4\varepsilon^2 n(1+o(1))$ and every other component is an isolated vertex or a cycle of length $O(1/\varepsilon)$.

As a corollary Łuczak and Seierstad obtain a number of weaker results inside the critical window regarding complex components. They showed that there are $O_p(1)$ complex components containing $O_p(n^{1/3})$ vertices combined and that each has spread $\Omega_p(n^{1/3})$ (the *spread* of a complex digraph is the length of its shortest path between vertices of degree at least 3). For a sequence $(X_n)_{n\in\mathbb{N}}$ of random variables with X_n defined on the probability space $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$, the notation $X_n = O_p(g(n))$ means that for any $\varepsilon > 0$ there exist constants $C, N > 0$ such that for all $n \geq N$

$$\mathbb{P}_n\left(\left|\frac{X_n}{g(n)}\right| \geq C\right) \leq \varepsilon.$$

Furthermore, the notation $X_n = o_p(g(n))$ means that for any $\varepsilon > 0$,

$$\mathbb{P}_n\left(\left|\frac{X_n}{g(n)}\right| \geq \varepsilon\right) \to 0 \text{ as } n \to \infty$$

and similarly we may define other asymptotic notation in probability ω_p, Ω_p etc.

Our main result is to give bounds on the tail probabilities of $|\mathcal{C}_1|$ resembling those of Nachmias and Peres for $G(n,p)$.

Theorem 7.4 (Lower Bound). *Let $0 < \delta < 1/800$, $\lambda \in \mathbb{R}$ and $n \in \mathbb{N}$. Let $\mathcal{C}_1$ be the largest component of $D(n,p)$ for $p = n^{-1} + \lambda n^{-4/3}$. Then if n is sufficiently large with respect to δ, λ,*

$$\mathbb{P}(|\mathcal{C}_1| < \delta n^{1/3}) \leq 2e\delta^{1/4}, \tag{7.1}$$

provided that $\delta \leq \min\left(\frac{(\log 2)^2}{4|\lambda|^2}, \frac{1}{800}\right)$.

Note that the constants in the above theorem have been chosen for simplicity and it is possible to give an expression for (7.1) depending on both λ and δ which imposes no restriction on their relation to one another.

Theorem 7.5 (Upper Bound). *Let $\mathcal{C}_1$ be the largest component of $D(n,p)$ for $p = n^{-1} + \lambda n^{-4/3}$. There exist constants, $\zeta, \eta > 0$ such that for any $A > 0, \lambda \in \mathbb{R}$ the following holds. Provided n is sufficiently large with respect to A, λ, and defining $\lambda^+ := \max(\lambda, 0)$,*

$$\mathbb{P}(|\mathcal{C}_1| > An^{1/3}) \leq \zeta e^{-\eta A^{3/2} + \lambda^+ A}.$$

A simple corollary of these bounds is that the largest component has size $\Theta(n^{1/3})$ with high probability. This follows by taking $\delta = o(1)$ in Theorem 7.4 and $A = \omega(1)$ in Theorem 7.5.

Corollary 7.6. *Let $\mathcal{C}_1$ be the largest component of $D(n,p)$ for $p = n^{-1} + \lambda n^{-4/3}$. Then, $|\mathcal{C}_1| = \Theta_p(n^{1/3})$.*

It should be noted that, in contrast to the undirected case, checking whether a set W of vertices constitutes a strongly connected component of a digraph D requires much more than checking only those edges with at least one end in W. In particular, in order for W to be a strongly connected component, it must be strongly connected and there must be no directed path starting and ending in W which contains vertices that are not in W. This precludes us from using a number of methods which have often been used to study $G(n,p)$. We therefore develop novel methods for counting the number of strongly connected components of $D(n,p)$ based upon branching process arguments.

The remainder of this chapter is organised as follows. In Section 7.2 we give a pair of bounds on the number of strongly connected digraphs which have a given excess and number of vertices. Sections 7.3 and 7.4 contain the proofs of Theorems 7.4 and 7.5 respectively in the case that $p = n^{-1}$. The proof of Theorem 7.4 in Section 7.3 is a relatively straightforward application of Janson's inequality. The proof of Theorem 7.5 in Section 7.4 is much more involved. We use an exploration process to approximate the probability that a given subdigraph of $D(n, p)$ is also a component. Using this we approximate the expected number of strongly connected components of size at least $An^{1/3}$ and apply Markov's inequality. The adaptations required to handle the critical window $p = n^{-1} + \lambda n^{-4/3}$ are presented in Section 7.5. We conclude the chapter in Section 7.6 with some open questions and final remarks.

7.2 Enumeration of Digraphs by size and excess

For both the upper and lower bounds on the size of the largest component, we need good bounds on the numbers of strongly connected digraphs with a given excess and number of vertices, where the *excess* of a strongly connected digraph with v vertices and e edges is $e - v$. Let $Y(m, k)$ be the number of strongly connected digraphs with m vertices and excess k. The study of $Y(m, k)$ was initiated by Wright [117] who obtained recurrences for the exact value of $Y(m, k)$. However, these recurrences swiftly become intractable as k grows. This has since been extended to asymptotic formulae when $k = \omega(1)$ and $O(m \log(m))$ [97, 99]. Note that when $k = m \log(m) + \omega(m)$, the fact $Y(m, k) \sim \binom{m(m-1)}{m+k}$ is a simple corollary of a result of Palásti [95]. In this section we give an universal bound on $Y(m, k)$ (Lemma 7.7) as well as a stronger bound for small excess (Lemma 7.9).

Lemma 7.7. *For every $m, k \geq 1$,*

$$Y(m, k) \leq \frac{(m + k)^k m^{2k} (m - 1)!}{k!}.$$

Proof. We will prove this by considering ear decompositions of the strongly connected digraphs in question. An *ear* is a non-trivial directed path in which the endpoints may coincide (i.e. it may be a cycle with a marked start/end vertex). The internal vertices of an ear are those that are not endpoints. An *ear decomposition* of a digraph D is a sequence, $E_0, E_1, \ldots, E_k$ of ears such that:

- E_0 is a directed cycle.
- The endpoints of E_i belong to $\bigcup_{j=0}^{i-1} E_j$.
- The internal vertices of E_i are disjoint from $\bigcup_{j=0}^{i-1} E_j$.
- $\bigcup_{i=0}^{k} E_i = D$.

We make use of the following fact.

Fact 7.8. A digraph D has an ear decomposition with $k+1$ ears if and only if D is strongly connected with excess k.

Thus we count strongly connected digraphs by a double counting of the number of possible ear decompositions. We produce an ear decomposition with m vertices and $k+1$ ears as follows. First, pick an ordering π of the vertices. Then insert k bars between the vertices such that the earliest the first bar may appear is after the second vertex in the order; multiple bars may be inserted between a pair of consecutive vertices. Finally, for each $i \in [k]$, we choose an ordered pair of vertices (u_i, v_i) which appear in the ordering before the ith bar.

This corresponds to a unique ear decomposition. The vertices in π before the first bar are E_0 with its endpoint being the first vertex. The internal vertices of E_i are the vertices of π between the ith and $i+1$st bar. Furthermore, E_i has endpoints u_i and v_i and is directed from u_i to v_i. The orientation of every other edge follows the order π.

Hence, there are at most

$$\binom{m+k-2}{k} m^{2k} m! \leq \frac{(m+k)^k m^{2k} m!}{k!}$$

ear decompositions. Note that each vertex of a strongly connected digraph is contained in a cycle. Therefore each vertex could be the endpoint of E_0 and hence at least m ear decompositions correspond to each strongly connected digraph. Hence the number of strongly connected digraphs of excess k may be bounded by

$$Y(m,k) \leq \frac{(m+k)^k m^{2k} m!}{k! m} = \frac{(m+k)^k m^{2k}(m-1)!}{k!},$$

as claimed. □

Lemma 7.9. *There exists $C > 0$ such that for $1 \leq k \leq \sqrt{m}/3$ and m sufficiently large we have,*

$$Y(m,k) \leq C\frac{m! m^{3k-1}}{(2k-1)!}. \tag{7.2}$$

Note that the above lemma is true for any $k = O(\sqrt{m})$ and the proof remains the same, only changing the final constant. The proof of the above lemma follows similar lines to the proof of Theorem 1.1 in [97] to obtain a bound of a similar order. We then prove that this bound implies the above which is much easier to work with.

First we introduce some definitions and notation from [97]. A random variable X has the zero-truncated Poisson distribution with parameter $\lambda > 0$ denoted $X \sim TP(\lambda)$ if it has probability mass function

$$\mathbb{P}(X = i) = \begin{cases} \frac{\lambda^i}{i!(e^\lambda - 1)} & \text{if } i \geq 1, \\ 0 & \text{if } i < 1. \end{cases}$$

Let $\mathcal{D}$ be the collection of all degree sequences $\mathbf{d} = (d_1^+, \ldots, d_m^+, d_1^-, \ldots, d_m^-)$ such that $d_i^+, d_i^- \geq 1$ for each $1 \leq i \leq m$ and furthermore,

$$\sum_{i=1}^m d_i^+ = \sum_{i=1}^m d_i^- = m + k.$$

A *preheart* is a (not neccesarily simple) digraph with minimum semi-degree at least 1 and no cycle components. The *heart* of a preheart D is the multidigraph $H(D)$ formed by suppressing all vertices of D which have in and out degree precisely 1.

We define the *preheart configuration model*, a two stage variant of the configuration model for digraphs which always produces a preheart, as follows. For $\mathbf{d} \in \mathcal{D}$, define

$$T = T(\mathbf{d}) = \{i \in [m] : d_i^+ + d_i^- \geq 3\}.$$

First we apply the configuration model to T to produce a heart H. That is, assign each vertex $i \in T$ d_i^+ out-stubs and d_i^- in-stubs and pick a uniformly random perfect matching between in- and out-stubs. Next, given a heart configuration H, we construct a preheart configuration Q by assigning $[m] \setminus T$ to $E(H)$ such that the vertices assigned to each arc of H are given a linear order. Denote this assignment including the orderings by q. Then the preheart configuration model, $\mathcal{Q}(\mathbf{d})$ is the probability space of random preheart configurations formed by choosing H and q uniformly at random. Note that each $Q \in \mathcal{Q}(\mathbf{d})$ corresponds to a (multi)digraph with m vertices $m + k$ edges and degree sequence $\mathbf{d}$.

As in the configuration model, each simple digraph with degree sequence $\mathbf{d}$ is produced in precisely $\prod_{i=1}^{m}(d_i^+!d_i^-!)$ ways. So if we restrict to simple preheart configurations, the digraphs we generate in this way are uniformly distributed, where in this case, simple means that there are no multiple edges or loops (however cycles of length 2 are allowed). We now count the number of preheart configurations. Let $m' = m'(\mathbf{d}) = |T(\mathbf{d})|$ be the number of vertices of the heart. Then, we have the following

Lemma 7.10. *Let* $\mathbf{d} \in \mathcal{D}$*. There are*

$$\frac{m'(\mathbf{d}) + k}{m + k}(m + k)!$$

preheart configurations.

Proof. We first generate the heart, and as we are simply working with the configuration model for this part of the model, there are $(m'+k)!$ heart configurations. The assignment of vertices in $[m] \setminus T$ to the arcs of the heart H may be done one vertex at a time by subdividing any already present edge and maintaining orientation. In this way when we add the ith vertex in this stage, there are $m'+k+i-1$ choices for the edge we subdivide. We must add $m-m'$ edges in this stage and so there are

$$\prod_{i=1}^{m-m'} m'+k+i-1 = \frac{(m+k-1)!}{(m'+k-1)!}$$

unique ways to create a preheart configuration from any given heart. Multiplying the number of heart configurations by the number of ways to create a preheart configuration from a given heart yields the desired result. $\square$

The next stage is to pick the degree sequence, $\mathbf{d} \in \mathcal{D}$ at random. We do this by choosing the degrees to be independent and identically distributed zero-truncated Poisson random variables with mean $\lambda > 0$. That is, $d_i^+ \sim TP(\lambda)$ and $d_i^- \sim TP(\lambda)$ such that the family $\{d_i^+, d_i^- : i \in [m]\}$ is independent. Note that this may not give a degree sequence at all, or it may be the degree sequence of a digraph with the wrong number of edges. Thus we define the event $\Sigma(\lambda)$ to be the event that

$$\sum_{i=1}^{m} d_i^+ = \sum_{i=1}^{m} d_i^- = m+k.$$

We shall now prove the following bound,

Lemma 7.11. *For any $\lambda > 0$ we have*

$$Y(m,k) \leq \frac{3k(m+k-1)!(e^{\lambda}-1)^{2m}}{\lambda^{2(m+k)}} \mathbb{P}(\Sigma(\lambda)). \tag{7.3}$$

Proof. Let $\mathbf{D}$ be the random degree sequence generated as above and $\mathbf{d} \in \mathcal{D}$, then

$$\mathbb{P}(\mathbf{D} = \mathbf{d}) = \prod_{i=1}^{m} \frac{\lambda^{d_i^+}}{d_i^+!(e^\lambda - 1)} \frac{\lambda^{d_i^-}}{d_i^-!(e^\lambda - 1)} = \frac{\lambda^{2(m+k)}}{(e^\lambda - 1)^{2m}} \prod_{i=1}^{m} \frac{1}{d_i^+!d_i^-!}. \tag{7.4}$$

By definition of $\Sigma(\lambda)$, we have

$$\sum_{\mathbf{d} \in \mathcal{D}} \mathbb{P}(\mathbf{D} = \mathbf{d}) = \mathbb{P}(\Sigma(\lambda)),$$

as all of the above events are disjoint. Thus, we may rearrange (7.4) to deduce that

$$\sum_{\mathbf{d} \in \mathcal{D}} \prod_{i=1}^{m} \frac{1}{d_i^+!d_i^-!} = \frac{(e^\lambda - 1)^{2m}}{\lambda^{2(m+k)}} \mathbb{P}(\Sigma(\lambda)). \tag{7.5}$$

Lemma 7.10 tells us that for a given degree sequence $\mathbf{d}$, there are

$$\frac{m'(\mathbf{d}) + k}{m + k}(m + k)!$$

preheart configurations. As each simple digraph with degree sequence $\mathbf{d}$ comes from precisely $\prod_{i=1}^{m} d_i^+!d_i^-!$ configurations, and $m'(\mathbf{d}) \leq 2k$ as otherwise the excess would be larger than k, we can deduce that the total number of prehearts with m vertices and excess k is at most

$$\sum_{\mathbf{d} \in \mathcal{D}} \frac{m'(\mathbf{d}) + k}{m + k}(m + k)! \prod_{i=1}^{m} \frac{1}{d_i^+!d_i^-!} \leq \sum_{\mathbf{d} \in \mathcal{D}} (m + k)! \frac{3k}{m + k} \prod_{i=1}^{m} \frac{1}{d_i^+!d_i^-!}. \tag{7.6}$$

Note that any strongly connected digraph is a preheart and so (7.6) is also an upper bound for $Y(m, k)$. Finally, combining (7.5) and (7.6) yields the desired inequality. $\square$

It remains to prove that (7.3) can be bounded from above by (7.2). To this end, we prove the following upper bound on $\mathbb{P}(\Sigma(\lambda))$.

Lemma 7.12. *For $\lambda < 1$,*

$$\mathbb{P}(\Sigma(\lambda)) \leq \frac{147}{\lambda m}.$$

For the proof of this lemma, we will use the Berry-Esseen inequality for normal approximation (see for example [115, Section XX.2].)

Lemma 7.13. *Suppose $X_1, X_2, \ldots, X_n$ is a sequence of independent random variables from a common distribution with zero mean, unit variance and third absolute moment $\mathbb{E}|X|^3 = \gamma < \infty$. Let $S_n = X_1 + X_2 + \ldots + X_n$ and let G_n be the cumulative distribution function of $S_n/\sqrt{n}$. Then for each n we have*

$$\sup_{t\in\mathbb{R}} |G_n(t) - \Phi(t)| \leq \frac{\gamma}{2\sqrt{n}}, \tag{7.7}$$

where Φ is the cumulative distribution function of the standard Gaussian.

Here, the explicit constant $1/2$ in equation (7.7) was obtained by Tyurin [114].

Proof of Lemma 7.12. The in-degrees of the random degree sequence are chosen independently from a truncated poisson distribution with parameter λ. Thus, we want to apply Lemma 7.13 to the sum $S_m = Y_1 + Y_2 + \ldots + Y_m$ where the Y_i are normalised truncated Poisson random variables. So all we must compute are the first three central moments of the truncated poisson distribution. Let $Y \sim TP(\lambda)$, one can easily compute that $\mathbb{E}(Y) = c_\lambda = \frac{\lambda e^\lambda}{e^\lambda - 1}$ and $\mathrm{Var}(Y) = \sigma_\lambda^2 = c_\lambda(1 + \lambda - c_\lambda)$. Note that for $\lambda < 1$ as c_λ is increasing in λ, we have $1 < c_\lambda \leq c_1 < 2$ and so as Y only takes integer values which are at least 1, $\mathbb{E}|Y - \mathbb{E}(Y)|^3 = \mathbb{E}(Y - c_\lambda)^3 + 2(c_\lambda - 1)^3\mathbb{P}(Y = 1)$. Computing this yields

$$\mathbb{E}|Y - \mathbb{E}(Y)|^3 = \lambda + \frac{2\lambda^4 - 5\lambda^3 + 3\lambda^2 - \lambda}{e^\lambda - 1} + \frac{3(2\lambda^4 - 3\lambda^3 + \lambda^2)}{(e^\lambda - 1)^2} + \frac{2(3\lambda^4 - 2\lambda^3)}{(e^\lambda - 1)^3} + \frac{2\lambda^4}{(e^\lambda - 1)^4}$$

One can check that this is bounded above by 2λ for $\lambda < 1$.

The normalised version of Y is $X = (Y - c_\lambda)/\sigma_\lambda$. We have

$$\mathbb{E}|X|^3 = \mathbb{E}\left|\frac{Y - c_\lambda}{\sigma_\lambda}\right|^3 = \frac{1}{\sigma_\lambda^3}\mathbb{E}|Y - c_\lambda|^3 \leq \frac{2\lambda}{\sigma_\lambda^3} = \gamma.$$

For $\lambda < 1$ one can check $c_\lambda < 1 + 2\lambda/3$, which allows us to deduce that $\sigma_\lambda^2 > \lambda/3$ (also using $Y \geq 1$). Hence, $\mathbb{E}|X|^3 \leq 6\sqrt{3}\lambda^{-1/2}$. Substituting into Lemma 7.13 with G_m the distribution function of $S_m/\sqrt{m}$,

$$\sup_{t\in\mathbb{R}} |G_m(t) - \Phi(t)| \leq \frac{3\sqrt{3}}{\sqrt{\lambda m}}.$$

The probability that the sum of the in-degrees is $m + k$ is precisely

$$G_m\left(\frac{m+k-mc_\lambda}{\sigma_\lambda\sqrt{m}}\right) - G_m\left(\frac{m+k-1-mc_\lambda}{\sigma_\lambda\sqrt{m}}\right).$$

Following an application of the triangle inequality, we see that this probability is bounded above by

$$\frac{6\sqrt{3}}{\sqrt{\lambda m}} + \frac{1}{\sqrt{2\pi m}\sigma_\lambda} \leq \frac{7\sqrt{3}}{\sqrt{\lambda m}}.$$

As the event that the in-degrees sum to $m + k$ and the event that the out-degrees sum to $m + k$ are independent and identically distributed events, we may deduce the bound,

$$\mathbb{P}(\Sigma(\lambda)) \leq \frac{147}{\lambda m}.$$

□

Finally, we may prove Lemma 7.9.

Proof of Lemma 7.9. We choose $\lambda = 2k/m < 1$ by assumption, then $\mathbb{P}(\Sigma(\lambda)) \leq 147/2k$ by Lemma 7.12. Combining this with Lemma 7.11 yields

$$Y(m,k) \leq \frac{441(m+k-1)!}{2}\lambda^{-2k}\left(\frac{e^\lambda - 1}{\lambda}\right)^{2m} \leq \frac{441m!m^{3k-1}e^{k^2/m}}{(2k)^{2k}}\left(\frac{e^\lambda - 1}{\lambda}\right)^{2m} \tag{7.8}$$

We use the inequality $e^x \leq 1 + x + x^2/2 + x^3/4$ which holds for all $0 \leq x \leq 1$ to bound $(e^\lambda - 1)/\lambda \leq 1 + \lambda/2 + \lambda^2/4$. Thus,

$$((e^\lambda - 1)/\lambda)^{2m} \leq (1 + \lambda/2 + \lambda^2/4)^{2m} \leq e^{m\lambda + m\lambda^2/2} = e^{2k+2k^2/m}.$$

Then, we can use Stirling's inequality, $e\sqrt{2k-1}(2k-1)^{2k-1}e^{-2k+1} \geq (2k-1)!$, so that

$$\frac{e^{2k}}{(2k)^{2k}} \leq \frac{e^{2k}}{(2k-1)^{2k-1/2}} \leq \frac{e^2}{(2k-1)!},$$

allowing us to rewrite the bound on $Y(m,k)$ as

$$Y(m,k) \leq \frac{441e^3}{2} \frac{m!m^{3k-1}}{(2k-1)!},$$

where we used $e^{k^2/m} \leq e^{1/3}$. This proves the lemma with $C = 441e^3/2$. $\square$

7.3 Proof of Theorem 7.4

In this section we prove a lower bound on component sizes in $D(n,p)$. We give the proof for $p = 1/n$ for simplicity. The proof when $p = n^{-1} + \lambda n^{-4/3}$ is very similar, with more care taken in the approximation of terms involving $(np)^m$. See Section 7.5 for more details.

Theorem 7.14. *Let $0 < \delta < 1/800$, then the probability that $D(n,1/n)$ has no component of size at least $\delta n^{1/3}$ is at most $2\delta^{1/2}$.*

To prove this we will bound from above the probability that there is no cycle of length between $\delta n^{1/3}$ and $\delta^{1/2} n^{1/3}$. Let X be the random variable counting the number of cycles in $D(n,1/n)$ of length between $\delta n^{1/3}$ and $\delta^{1/2} n^{1/3}$. Note that we may decompose X as a sum of dependent Bernoulli random variables, and thus we may apply Janson's Inequality in the following form (see [58, Theorem 2.18 (i)]).

Theorem 7.15. *Let S be a set and $S_p \subseteq S$ chosen by including each element of S in S_p independently with probability p. Suppose that $\mathcal{S}$ is a family of subsets of S and for $A \in \mathcal{S}$, we define I_A to be the event $\{A \subseteq S_p\}$. Let $\mu = \mathbb{E}(X)$ and*

$$\Delta = \frac{1}{2} \sum\sum_{A \neq B, A \cap B \neq \emptyset} \mathbb{E}(I_A I_B).$$

Then,

$$\mathbb{P}(X = 0) \leq e^{-\mu+\Delta}.$$

To apply Theorem 7.15, we define S to be the set of edges of the complete digraph on n vertices. Let $A \in \mathcal{S}$ if and only if $A \subseteq S$ is the set of edges of a cycle of length between $\delta n^{1/3}$ and $\delta^{1/2} n^{1/3}$. Define $X(m)$ to be the number cycles in $D(n, 1/n)$ of length m. We start by approximating the first moment of X.

Lemma 7.16. $\mathbb{E}(X) \geq \log(1/\delta)/2(1 + o(1))$

Proof. Let $a = \delta n^{1/3}$ and $b = \delta^{1/2} n^{1/3}$. Then, we can write X as

$$X = \sum_{m=a}^{b} X(m).$$

Note that

$$\mathbb{E}(X(m)) = \binom{n}{m} \frac{m!}{m} p^m \geq \frac{1}{m}(1 + o(1)). \tag{7.9}$$

So, we may bound the expectation of X as follows

$$\mathbb{E}(X) = \sum_{m=a}^{b} \mathbb{E}(X(m)) \geq (1 + o(1)) \sum_{m=a}^{b} \frac{1}{m} \geq (1 + o(1)) \int_a^b \frac{dx}{x} = (1 + o(1)) \frac{\log(1/\delta)}{2}.$$

□

Let $Z(m, k)$ be the random variable counting the number of strongly connected graphs with m vertices and excess k in $D(n, 1/n)$. Directly computing Δ is rather complicated so we will instead compute an upper bound on Δ that is a linear combination of the first moments of the random variables $Z(m, k)$ for $m \geq a$ and $k \geq 1$. To move from the computation of Δ to the first moments of $Z(m, k)$ we use the following lemma,

Lemma 7.17. *Each strongly connected digraph D with excess k may be formed in at most 27^k ways as the union of a pair of directed cycles C_1 and C_2.*

Proof. Consider the heart $H(D)$ of D. Recall that $H(D)$ is the (multi)-digraph formed by suppressing the degree 2 vertices of D and retaining orientations. As D has excess k, $H(D)$ has at most $2k$ vertices. Furthermore, the excess of $H(D)$ is the same as the excess of D as we only suppress vertices of degree 2. Thus $H(D)$ has at most $3k$ edges.

Then, each edge of $H(D)$ must be a subdigraph of either C_1, C_2 or both. So there are $3^{3k} = 27^k$ choices for the pair C_1, C_2 as claimed. $\square$

We are now in a position to give a bound on Δ.

Lemma 7.18. $\Delta \leq \log(2)$ *for any* $\delta \in (0, 1/800]$

Proof. Let

$$\Gamma(k) := \{E(C) | C \subseteq \overrightarrow{K_n}, C \cong \overrightarrow{C_k}\},$$

where $\overrightarrow{K_n}$ is the complete digraph on $[n]$ and $\overrightarrow{C_k}$ is the directed cycle of length k. For $\alpha \in \Gamma(k)$ let I_α be the indicator function of the event that all edges of α are present in a given realisation of $D(n, 1/n)$. Also, define

$$\Gamma = \bigcup_{k=a}^{b} \Gamma(k).$$

Then, by definition,

$$\Delta = \frac{1}{2} \sum\sum_{\alpha \neq \beta, \alpha \cap \beta \neq \emptyset} \mathbb{E}(I_\alpha I_\beta)$$

Let $\Gamma_\alpha^{m,k}(t)$ be the set of $\beta \in \Gamma(t)$ such that $\alpha \cup \beta$ is a collection of $m+k$ edges spanning m vertices. Then,

$$\begin{aligned} 2\Delta &= \sum_{s=a}^{b} \sum_{t=a}^{b} \sum_{\alpha \in \Gamma(s)} \sum_{m=s}^{\infty} \sum_{k=1}^{\infty} \sum_{\beta \in \Gamma_\alpha^{m,k}(t)} p^{m+k} \\ &\leq \sum_{m=a}^{2b} \sum_{k=1}^{\infty} \sum_{s=a}^{m} \sum_{t=a}^{m} \sum_{\alpha \in \Gamma(s)} \sum_{\beta \in \Gamma_\alpha^{m,k}(t)} p^{m+k} \\ &\leq \sum_{m=a}^{2b} \sum_{k=1}^{\infty} 27^k \mathbb{E}(Z(m,k)), \end{aligned} \tag{7.10}$$

where the last inequality follows from Lemma 7.17. Note that

$$\mathbb{E}(Z(m,k)) = \binom{n}{m} p^{m+k} Y(m,k)$$

by definition. We will use the following two bounds on $Y(m,k)$ which follow immediately from Lemma 7.7.

- If $k \le m$, then $Y(m,k) \le \frac{2^k m^{3k} m!}{k! m}$
- If $k > m$, then $Y(m,k) \le \frac{(2e)^k m^{2k} m!}{m}$

This allows us to split the sum in (7.10) based upon whether $k \le m$ or $k > m$ to obtain

$$\begin{aligned} 2\Delta &\le \sum_{m=a}^{2b} \sum_{k=1}^{m} 27^k \binom{n}{m} \frac{2^k m^{3k} m!}{k! m} p^{m+k} + \sum_{m=a}^{2b} \sum_{k=m+1}^{\infty} 27^k \binom{n}{m} \frac{(2e)^k m^{2k} m!}{m} p^{m+k} \\ &\le \sum_{m=a}^{2b} \frac{1}{m} \sum_{k=1}^{\infty} \frac{(54pm^3)^k}{k!} + \sum_{m=a}^{2b} \frac{1}{m} \sum_{k=m+1}^{\infty} (54em^2 p)^k \\ &\le \frac{\log(4/\delta)}{2} (e^{432\delta^{3/2}} - 1 + 23328 e^2 \delta^2), \end{aligned} \tag{7.11}$$

where the $23328e^2\delta^2$ term comes from noting $k \ge 2$ in the range $k \ge m+1$ and that for $x \le 1/2$,

$$\sum_{k=2}^{\infty} x^k \le 2x^2.$$

As (7.11) is increasing in δ, we simply need to check that the Lemma holds for $\delta = 1/800$ which may be done numerically.

□

Finally, to prove Theorem 7.14 we substitute the values obtained for μ and Δ in Lemmas 7.16 and 7.18 respectively into Theorem 7.15. That is,

$$\mathbb{P}(X = 0) \le e^{-\mu + \Delta} \le e^{-\log(1/\delta)/2 + \log(2)} = 2\delta^{1/2}$$

So the probability there is no directed cycle of length at least $\delta n^{1/3}$ is at most $2\delta^{1/2}$ and, as cycles are strongly connected, this is also an upper bound on the probability there is no strongly connected component of size at least $\delta n^{1/3}$.

7.4 Proof of Theorem 7.5

In this section we prove an upper bound on the component sizes in $D(n,p)$. Again, we only consider the case when $p = 1/n$ to simplify notation and calculations. The reader is referred to Section 7.5 for a sketch of the adaptations to extend the result to the full critical window. The following is a restatement of Theorem 7.5 for $p = 1/n$.

Theorem 7.19. *There exist constants $\zeta, \eta > 0$ such that for any $A > 0$ if n is sufficiently large with respect to A, then the probability that $D(n, 1/n)$ contains any component of size at least $An^{1/3}$ is at most $\zeta e^{-\eta A^{3/2}}$.*

We will use the first moment method to prove this theorem and calculate the expected number of large strongly connected components in $D(n, 1/n)$. Note that it is important to count components and not strongly connected subgraphs as the expected number of strongly connected subgraphs in $D(n, 1/n)$ blows up as $n \to \infty$. Thus for each strongly connected subgraph, we will use an exploration process to determine whether or not it is a component.

The exploration process we will use was initially developed independently by Martin-Löf [88] and Karp [60]. During this process, vertices will be in one of three classes: *active*, *explored* or *unexplored*. At time $t \in \mathbb{N}$, we let X_t be the number of active vertices, A_t the set of active vertices, E_t the set of explored vertices and U_t the set of unexplored vertices.

We will start from a set A_0 of vertices of size X_0 and fix an ordering of the vertices, starting with A_0. For step $t \geq 1$, if $X_{t-1} > 0$ let w_t be the first active vertex. Otherwise, let w_t be the first unexplored vertex. Define η_t to be the number of unexplored out-neighbours of w_t in $D(n, 1/n)$. Change the class of each of these vertices to active and set w_t to explored. This means that $|E_t| = t$ and furthermore, $|U_t| = n - X_t - t$. Let

$N_t = n - X_t - t - \mathbb{1}(X_t = 0)$ be the number of potential unexplored out-neighbours of w_{t+1} i.e. the number of unexplored vertices which are not w_{t+1}. Then, given the history of the process, η_t is distributed as a binomial random variable with parameters N_{t-1} and $1/n$. Furthermore, the following recurrence relation holds.

$$X_t = \begin{cases} X_{t-1} + \eta_t - 1 & \text{if } X_{t-1} > 0, \\ \eta_t & \text{otherwise} \end{cases} \tag{7.12}$$

Let $\tau_1 = \min\{t \geq 1 : X_t = 0\}$. Note that this is a stopping time and at time τ_1 the set E_{τ_1} of explored vertices is precisely the out-component of A_0. If A_0 spans a strongly connected subdigraph D_0 of $D(n, 1/n)$, then D_0 is a strongly connected component if and only if there are no edges from $E_{\tau_1} \setminus A_0$ to A_0. The key idea will be to show that if X_0 is sufficiently large, then it is very unlikely for τ_1 to be small, and consequently it is also very unlikely that there are no edges from $E_{\tau_1} \setminus A_0$ to A_0. This is encapsulated in the following lemma.

Lemma 7.20. *Let X_t be the exploration process defined above with starting set of vertices A_0 of size $X_0 = m$. Suppose $0 < c < \sqrt{2}$ is a fixed constant. Then,*

$$\mathbb{P}(\tau_1 < cm^{1/2}n^{1/2}) \leq 2e^{-\frac{(2-c^2)^2}{8c}m^{3/2}n^{-1/2} + O(m^2n^{-1})}.$$

Proof. Define $\xi = cm^{1/2}n^{1/2}$ and consider the auxiliary process, X'_t which we define recursively by

$$X'_0 = m,$$
$$X'_t = X'_{t-1} - 1 + W_t \text{ for } t \geq 1,$$

where $W_t \sim \text{Bin}(n - t - 10m, p)$. Let τ_2 be the stopping time,

$$\tau_2 = \inf\{t : X_t > 10m\}$$

We note that X'_t stochastically dominated by X_t for $0 \leq t < \tau_2$. That is, there exists a coupling the processes as $(\hat{X}_t, \hat{X}'_t)$ on the same probability space where $\hat{X}_t$ has the same distribution as X_t and $\hat{X}'_t$ has the same distribution as X_t such that $\hat{X}'_t \leq \hat{X}_t$ with probability 1 for $0 \leq t < \tau_2$. The coupling may be explicitly defined by setting $\hat{\eta}_t = \hat{W}_t + \hat{W}'_t$ with $\hat{W}'_t \sim \mathrm{Bin}(10m - \hat{X}_{t-1}, p)$ where $\hat{\eta}_t, \hat{W}_t$ are versions of η_t, W_t but for the random variables $\hat{X}_t, \hat{X}'_t$ instead of X_t and X'_t.

Define another stopping time, $\tau'_1 = \min\{t \geq 1 : X'_t = 0\}$ and consider the following events

$$
\begin{aligned}
\mathcal{E}_1 &= \{\tau_1 < cm^{1/2}n^{1/2}\}, \\
\mathcal{E}_2 &= \{\tau'_1 < cm^{1/2}n^{1/2}\}, \\
\mathcal{E}_3 &= \{\tau_2 < cm^{1/2}n^{1/2}\}.
\end{aligned}
$$

And note that $\mathbb{P}(\mathcal{E}_1) \leq \mathbb{P}(\mathcal{E}_2) + \mathbb{P}(\mathcal{E}_3)$ by our choice of coupling and a union bound (as the coupling guarantees $\mathcal{E}_1 \subseteq \mathcal{E}_2 \cup \mathcal{E}_3$). Thus we only need to bound the probabilities of the simpler events $\mathcal{E}_2$ and $\mathcal{E}_3$. We begin by considering $\mathcal{E}_3$. To bound its probability we consider the upper bound process M_t defined by

$$
\begin{aligned}
M_0 &= m, \\
M_t &= M_{t-1} - 1 + B_t \text{ for } t \geq 1,
\end{aligned}
$$

where $B_t \sim \mathrm{Bin}(n, 1/n)$. It is straightforward to couple (X_t, M_t) such that M_t stochastically dominates X_t. Furthermore, M_t is a martingale. Hence, $\mathbb{P}(\mathcal{E}_3) \leq \mathbb{P}(\tau'_2 < cm^{1/2}n^{1/2})$ where τ'_2 is the stopping time, $\tau'_2 = \min\{t : M_t > 10m\}$. To bound the probability of $\mathcal{E}_2$ consider the process Y_t defined as $Y_t = m - X'_t$. One can check that Y_t is a submartingale.

As $x \mapsto e^{\alpha x}$ is a convex non-decreasing function for any $\alpha > 0$, we may apply Jensen's inequality to deduce that $Z^-_t = e^{\alpha Y_t}$ and $Z^+_t = e^{\alpha M_t}$ are submartingales. Also, $Z^-_t, Z^+_t > 0$ for any $i \in \mathbb{N}$. Starting with Z^-_t, we may apply Doob's maximal inequality [48, Section 12.6] and deduce that

$$\mathbb{P}\left(\min_{0\le t\le \xi} X_i' \le 0\right) = \mathbb{P}\left(\max_{0\le t\le \xi} Z_t^- \ge e^{\alpha m}\right) \le \frac{\mathbb{E}(Z_\xi^-)}{e^{\alpha m}}. \tag{7.13}$$

We may rewrite this by noting that

$$Y_t = m - X_t' = t - \sum_{i=1}^{t} W_i = t - R_t.$$

where R_t is binomially distributed and in particular $R_\xi \sim \mathrm{Bin}(l\xi, p)$ for

$$l\xi = cm^{1/2}n^{3/2} - \frac{c^2mn}{2} - 10cm^{3/2}n^{1/2} + \frac{cm^{1/2}n^{1/2}}{2}.$$

Also, we choose x such that $xl\xi = \xi - m$. Then (7.13) may be rewritten as $e^{-\alpha m}\mathbb{E}(Z_\xi^-) = e^{\alpha x l\xi}\mathbb{E}(e^{-\alpha R_\xi})$. The next stage is to rearrange this into a form which resembles the usual Chernoff bounds (for $x < p$). So, let

$$f(\alpha) = e^{\alpha x l\xi}\mathbb{E}(e^{-\alpha R_\xi}) = \left[e^{\alpha x}(pe^{-\alpha} + 1 - p)\right]^{l\xi}.$$

Then, we choose α^* to minimise f. Solving $f'(\alpha) = 0$, we obtain the solution

$$e^{-\alpha^*} = \frac{x(1-p)}{p(1-x)}.$$

Note $x < p$ so, $e^{-\alpha^*} < 1$ and $\alpha^* > 0$ as desired. Thus,

$$\begin{aligned} f(\alpha^*) = &= \left[\left(\frac{p(1-x)}{x(1-p)}\right)^x \left(x\frac{1-p}{1-x} + 1 - p\right)\right]^{mt} \\ &= \left[\left(x\frac{1-p}{1-x} + 1 - p\right)\left(\frac{p}{x}\right)^x \left(\frac{1-p}{1-x}\right)^x\right]^{mt} \\ &= \left[\left(\frac{p}{x}\right)^x \left(\frac{1-p}{1-x}\right)^{1-x}\right]^{mt}. \end{aligned}$$

Which is the usual expression found in Chernoff bounds. As usual, we bound this by writing

$$f(\alpha^*) = e^{-g(x)l\xi}$$

and bound g, where

$$g(x) = x\log\left(\frac{x}{p}\right) + (1-x)\log\left(\frac{1-x}{1-p}\right).$$

Computing the Taylor expansion of g we find that $g(p) = g'(p) = 0$. So, if $g''(x) \geq \beta$ for all x between p and $p-h$, then $g(p-h) \geq \beta h^2/2$. Furthermore,

$$g''(x) = \frac{1}{x} + \frac{1}{1-x}.$$

As $0 < x < p$, we have $g''(x) \geq 1/x \geq 1/p$. So, we deduce that $g(x) \geq \delta^2 p/2$ where $\delta = 1 - x/p$. All that remains is to compute δ. As defined earlier, we have $xl\xi = \xi - m$ which for convenience we will write as

$$xl\xi = \xi\left(1 - \frac{m^{1/2}}{cn^{1/2}}\right). \tag{7.14}$$

Also, as $p = n^{-1}$, and recalling the definition of $l\xi$ from earlier,

$$\begin{aligned} pl\xi &= cm^{1/2}n^{1/2} - \frac{c^2m}{2} + O(m^{3/2}n^{-1/2}) \\ &= \xi\left(1 - \frac{cm^{1/2}}{2n^{1/2}} + O(mn^{-1})\right). \end{aligned} \tag{7.15}$$

We divide (7.14) by (7.15) and as the Taylor expansion of $1/(1-w)$ is $\sum_{i\geq 0} w^i$,

$$\frac{x}{p} = \frac{1 - \frac{m^{1/2}}{cn^{1/2}}}{1 - \frac{cm^{1/2}}{2n^{1/2}} + O(mn^{-1})} = 1 - \frac{m^{1/2}}{cn^{1/2}} + \frac{cm^{1/2}}{2n^{1/2}} + O(mn^{-1}). \tag{7.16}$$

From which we may deduce

$$\delta = \frac{(2-c^2)m^{1/2}}{2cn^{1/2}} + O(mn^{-1}). \tag{7.17}$$

So,

$$\mathbb{P}(\mathcal{E}_2) \le e^{-\frac{\delta^2 p}{2} l\xi} = e^{-\frac{(2-c^2)^2}{8c} m^{3/2} n^{-1/2} + O(m^2 n^{-1})}. \tag{7.18}$$

We may proceed similarly for Z_t^+, in particular we must still appeal to Doob's maximal inequality as we seek a bound over the entire process. In this case we end up with a $\mathrm{Bin}(n\xi, p)$ distribution and are looking at the upper tail rather than the lower. We find $pn\xi = \xi$ and

$$xn\xi = \xi + 9m = \xi\left(1 + \frac{9m^{1/2}}{cn^{1/2}}\right).$$

Thus,

$$\delta = \frac{x}{p} - 1 = \frac{9m^{1/2}}{cn^{1/2}}.$$

Substituting into the analogous bound,

$$\mathbb{P}(\mathcal{E}_3) \le e^{-\frac{\delta^2 p}{3} n\xi} \le e^{-\frac{27 m^{3/2}}{cn^{1/2}}}. \tag{7.19}$$

Observe that $\mathbb{P}(\mathcal{E}_2) \ge \mathbb{P}(\mathcal{E}_3) e^{O(m^2 n^{-1})}$ for $0 < c < \sqrt{2(1 + 3\sqrt{6})}$. Thus, in the range we are interested in, we may use $2\mathbb{P}(\mathcal{E}_2)$ as an upper bound for $\mathbb{P}(\mathcal{E}_2) + \mathbb{P}(\mathcal{E}_3)$ and this proves the lemma. $\square$

We now compute the probability that any given strongly connected subgraph of $D(n, 1/n)$ is a component. To do so, we use the simple observation that a strongly connected subgraph is a component if it is not contained in a larger strongly connected subgraph.

Lemma 7.21. *There exist $\beta, \gamma > 0$ such that if H is any strongly connected subgraph of $D(n, 1/n)$ with m vertices, then the conditional probability that H is a strongly connected component of $D(n, 1/n)$ is at most $\beta e^{-(1+\gamma) m^{3/2} n^{-1/2} + O(m^2 n^{-1})}$.*

Proof. We compute the probability that H is a component of $D(n, 1/n)$ by running the exploration process X_t starting from $A_0 = V(H)$. So, $X_0 = m$. Once the exploration process dies at time τ_1, any backward edge from $E_{\tau_1} \setminus A_0$ to A_0 gives a strongly connected subgraph of $D(n, 1/n)$ which contains H. Let Y_t be the random variable which counts the number of edges from $E_t \setminus A_0$ to A_0. Note that for $t \geq m$, $Y_t \sim \mathrm{Bin}(m(t-m), p)$. Furthermore, H is a strongly connected component of $D(n, 1/n)$ if and only if $Y_{\tau_1} = 0$.

Let $\varepsilon > 0$ and define the events $\mathcal{A}_i$ for $i = 1, \ldots, r$ (where $r \sim c/\varepsilon$ for some $c > 1$) to be

$$\begin{aligned} \mathcal{A}_i &= \{(i-1)\varepsilon m^{1/2} n^{1/2} \leq \tau_1 < i\varepsilon m^{1/2} n^{1/2}\}, \\ \mathcal{A}_{r+1} &= \{r\varepsilon m^{1/2} n^{1/2} \leq \tau_1\}. \end{aligned}$$

Clearly the family $\{\mathcal{A}_i : i = 1, \ldots, r+1\}$ forms a partition of the sample space. So, by the law of total probability,

$$\mathbb{P}(Y_{\tau_1} = 0) = \sum_{i=1}^{r+1} \mathbb{P}(Y_{\tau_1} = 0 | \mathcal{A}_i) \mathbb{P}(\mathcal{A}_i). \tag{7.20}$$

By applying Lemma 7.20 when $1 \leq i \leq r$ we find

$$\mathbb{P}(\mathcal{A}_i) \leq 2e^{-\frac{(2-i^2\varepsilon^2)^2}{8i\varepsilon} m^{3/2} n^{-1/2} + O(m^2 n^{-1})}.$$

Note that Y_{τ_1} conditioned on $\mathcal{A}_i$ stochastically dominates a $\mathrm{Bin}(m((i-1)\varepsilon m^{1/2} n^{1/2} - m), p)$ distribution. Therefore,

$$\mathbb{P}(Y_{\tau_1} = 0 | \mathcal{A}_i) \leq (1-p)^{m((i-1)\varepsilon m^{1/2} n^{1/2} - m)} \leq e^{-(i-1)\varepsilon m^{3/2} n^{-1/2} + O(m^2 n^{-1})}.$$

Combining the above and substituting into (7.20) yields

$$\begin{aligned} \mathbb{P}(Y_{\tau_1} = 0) &\leq 2 \sum_{i=1}^{r} e^{-((i-1)\varepsilon + \frac{(2-i^2\varepsilon^2)^2}{8i\varepsilon}) m^{3/2} n^{-1/2} + O(m^2 n^{-1})} + e^{-r\varepsilon m^{m/2} n^{-1/2} + O(m^2 n^{-1})}, \quad (7.21) \\ &\leq (2r+1) e^{-(1+\gamma) m^{3/2} n^{-1/2} + O(m^2 n^{-1})}, \quad (7.22) \end{aligned}$$

for some $\gamma > 0$ provided that ε is sufficiently small. The second term in (7.21) is a result of the fact $\mathbb{P}(A_{r+1}) \leq 1$. This proves the lemma and if one wishes for explicit constants, taking $\varepsilon = 0.025$, $r = 45$ works and gives $\beta < 100$, $\gamma > 0.06$. $\square$

The next stage in our proof is to show that a typical instance of $D(n, 1/n)$ has no component of large excess and no exceptionally large components. This will allow us to use the bound from Lemma 7.9 to compute the expected number of large strongly connected components of $D(n, 1/n)$. The first result in this direction is an immediate corollary of a result of Łuczak and Seierstad [86].

Lemma 7.22 ([86]). *The probability that $D(n, 1/n)$ contains a strongly connected component of size at least $n^{1/3} \log\log n$ is $o(1)$.*

The next lemma ensures that there are not too many cycles which enables us to prove that the total excess is relatively small.

Lemma 7.23. *The probability that $D(n, p)$ contains more than $n^{1/6}$ cycles of length at most $n^{1/3} \log\log(n)$ is $o(1)$.*

Proof. In this proof and subsequently we will use the convention that $\log^{(k)} x$ is the logarithm function composed with itself k times, while $(\log x)^k$ is its kth power. We shall show that the expected number of cycles of length at most $n^{1/3} \log^{(2)} n$ is $o(n^{1/6})$ at which point we may apply Markov's inequality. So let C be the random variable which counts the number of cycles of length at most $n^{1/3} \log^{(2)} n$ in $D(n, 1/n)$. We can calculate its expectation as

$$\mathbb{E}(C) = \sum_{k=1}^{n^{1/3} \log^{(2)} n} \binom{n}{k} \frac{k!}{k} p^k \leq \sum_{k=1}^{n^{1/3} \log^{(2)} n} \frac{1}{k}. \tag{7.23}$$

We use the upper bound on the kth harmonic number $H_k \leq \log k + 1$, which allows us to deduce that

$$\mathbb{E}(C) \leq H_{n^{1/3} \log^{(2)} n} \leq \frac{1}{3} \log n + \log^{(3)} n + 1 \leq \log n = o(n^{1/6}). \tag{7.24}$$

Thus the lemma follows by Markov's inequality. $\square$

Corollary 7.24. *The probability that $D(n, 1/n)$ contains a component of excess at least $n^{1/6}$ and size at most $n^{1/3} \log\log n$ is $o(1)$.*

Proof. If D is any strongly connected digraph with m vertices and excess k, then note that it must have at least $k+1$ cycles of length at most m. This can be seen by considering the ear decomposition of D. The first ear must be a cycle, and each subsequent ear adds a path which must be contained in a cycle as D is strongly connected. So as we build the ear decomposition, each additional ear adds at least one cycle. As any ear decomposition of a strongly connected digraph of excess k has $k + 1$ ears, then D must have at least $k + 1$ cycles.

Thus, if D has k cycles, it must have excess at most $k - 1$. So applying Lemma 7.23 completes the proof. $\square$

Finally, we prove the main theorem of this section.

Proof of Theorem 7.19. Let $\mathcal{C}_1$ be the largest strongly connected component of $D(n, 1/n)$ and $L_1 = |\mathcal{C}_1|$. We want to compute $\mathbb{P}(L_1 \geq An^{1/3})$. Define the following three events,

$$\begin{aligned}
\mathcal{E}_1 &= \{L_1 \geq An^{1/3}\}, \\
\mathcal{E}_2 &= \{An^{1/3} \leq L_1 \leq n^{1/3} \log\log(n)\}, \\
\mathcal{E}_3 &= \{L_1 \geq n^{1/3} \log\log(n)\}.
\end{aligned}$$

Clearly, $\mathcal{E}_1 \subseteq \mathcal{E}_2 \cup \mathcal{E}_3$ and by Lemma 7.22, $\mathbb{P}(\mathcal{E}_3) = o_n(1)$. If $\mathcal{F}$ is the event that $\mathcal{C}_1$ has excess at least $n^{1/6}$ then by Corollary 7.24, $\mathbb{P}(\mathcal{E}_2 \cap \mathcal{F}) = o_n(1)$. All that remains is to give a bound on $\mathbb{P}(\mathcal{E}_2 \cap \mathcal{F}^c)$. To this end let $N(A)$ be random variable which counts the number of strongly connected components of $D(n, 1/n)$ which have size between $An^{1/3}$ and $n^{1/3} \log\log n$ and excess bounded above by $n^{1/6}$. By Markov's inequality, we may deduce that $\mathbb{P}(\mathcal{E}_2 \cap \mathcal{F}^c) \leq \mathbb{E}(N(A))$. Computing the expectation of $N(A)$,

$$\mathbb{E}(N(A)) = \sum_{m=An^{1/3}}^{n^{1/3}\log^2(n)} \sum_{k=0}^{n^{1/6}} \binom{n}{m} p^{m+k} Y(m,k) \mathbb{P}(Y_{\tau_1} = 0 | X_0 = m). \tag{7.25}$$

In Lemma 7.21 we showed that $\mathbb{P}(Y_{\tau_1} = 0|X_0 = m) \leq \beta e^{-(1+\gamma)m^{3/2}n^{-1/2}+O(m^2n^{-1})}$. Also, using Lemma 7.9 we can check that

$$\begin{aligned}\sum_{k=0}^{n^{1/6}} Y(m,k)p^k &\leq (m-1)!\left(1 + C\sum_{k=1}^{n^{1/6}} \frac{m^{3k}p^k}{(2k-1)!}\right)\\ &\leq (m-1)!\left(1 + C\sum_{k=1}^{\infty} \frac{m^{3k}p^k}{(2k-1)!}\right)\\ &= (m-1)!\left(1 + C(m^3p)^{1/2}\sinh((m^3p)^{1/2})\right) \qquad (7.26)\end{aligned}$$

where the first term on the right hand side of (7.26) comes from the directed cycles and C is the same constant as in Lemma 7.9. As $\sinh(x) \leq e^x$ we can bound (7.26) by

$$\begin{aligned}\sum_{k=0}^{n^{1/6}} Y(m,k)p^k &\leq (m-1)!(1 + Cm^{3/2}n^{-1/2}e^{m^{3/2}n^{-1/2}})\\ &\leq 2(m-1)!Cm^{3/2}n^{-1/2}e^{m^{3/2}n^{-1/2}}.\end{aligned}$$

Combining these bounds and using $\binom{n}{m} \leq n^m/m!$ we deduce

$$\begin{aligned}\mathbb{E}(N(A)) &\leq \sum_{m=An^{1/3}}^{n^{1/3}\log^2(n)} \frac{(np)^m}{m!} \cdot 2(m-1)!Cm^{3/2}n^{-1/2}e^{m^{3/2}n^{-1/2}} \cdot \beta e^{-(1+\gamma)m^{3/2}n^{-1/2}+O(m^2n^{-1})}\\ &= \sum_{m=An^{1/3}}^{n^{1/3}\log^2(n)} \frac{2\beta C m^{1/2}}{n^{1/2}} e^{-\gamma m^{3/2}n^{-1/2}+O(m^2n^{-1})}\\ &\leq \int_{An^{1/3}}^{n^{1/3}\log^2(n)+1} \frac{2\beta C m^{1/2}}{n^{1/2}} e^{-\frac{\gamma}{2}m^{3/2}n^{-1/2}}\, dm, \qquad (7.27)\end{aligned}$$

where (7.27) holds for all sufficiently large n. Now making the substitution $x = mn^{-1/3}$

we can remove the dependence of (7.27) on both m and n so that

$$\begin{aligned}\mathbb{E}(N(A)) &\leq 2\beta C\int_A^{\log^2(n)+n^{-1/3}} x^{1/2}e^{-\frac{\gamma}{2}x^{3/2}}dx\\ &\leq 2\beta C\int_A^{\infty} x^{1/2}e^{-\frac{\gamma}{2}x^{3/2}}dx\\ &= \frac{8\beta C}{3\gamma}\int_{\frac{\gamma A^{3/2}}{2}}^{\infty} e^{-t}dt = \frac{8\beta C}{3\gamma}e^{-\frac{\gamma A^{3/2}}{2}}.\end{aligned} \tag{7.28}$$

So, by Markov's inequality $\mathbb{P}(\mathcal{E}_2\cap\mathcal{F}^c)\leq\zeta e^{-\eta A^{3/2}}$ where ζ and η are the corresponding constants found in (7.28). So,

$$\mathbb{P}(L_1\geq An^{1/3})\leq\mathbb{P}(\mathcal{E}_2\cap\mathcal{F}^c)+\mathbb{P}(\mathcal{E}_2\cap\mathcal{F})+\mathbb{P}(\mathcal{E}_3)=\zeta e^{-\eta A^{3/2}}+o_n(1).$$

Calculating ζ and γ using the values for C,β and γ in Lemmas 7.9 and 7.21 yields $\zeta<2\times10^7$ and $\eta>0.03$. □

7.5 Adaptations for the Critical Window

In this section we sketch the adaptations one must make to the proofs of Theorems 7.14 and 7.19 such that they hold in the whole critical window, $p=n^{-1}+\lambda n^{-4/3}$ where $\lambda\in\mathbb{R}$.

7.5.1 Lower Bound

For Theorem 7.14, the adaptation is rather simple. We will still apply Janson's inequality and so we only need to recompute μ and Δ. Furthermore, the only difference in these calculations comes from replacing the term n^{-m-k} by p^{m+k}, and in fact the p^k in this turns out to make negligible changes. In this light, Lemma 7.16 changes to

Lemma 7.25.

$$\mathbb{E}(X)\geq\begin{cases}-e^{\frac{\lambda\delta}{2}}\log(\delta)/2 & \textit{if } \lambda\geq 0\\ -e^{2\delta^{1/2}\lambda}\log(\delta)/2 & \textit{otherwise.}\end{cases}$$

The only difference in the proof is to bound $(1+\lambda n^{-1/3})^m$ by its lowest value depending on whether $\lambda \geq 0$ or $\lambda < 0$. We bound this via

$$1 + x \geq \begin{cases} e^{\frac{x}{2}} & \text{if } 0 \leq x \leq 2 \\ e^{2x} & \text{if } -\frac{1}{2} \leq x \leq 0. \end{cases}$$

Furthermore, Lemma 7.18 changes to

Lemma 7.26. *For all sufficiently large n and small enough δ,*

$$\Delta \leq \begin{cases} e^{2\delta^{1/2}\lambda} \log(2) & \textit{if } \lambda \geq 0 \\ e^{\delta\lambda} \log(2) & \textit{otherwise.} \end{cases}$$

The proof again is almost identical with the only change being to approximate the $(np)^m$ term. This time we seek an upper bound so use the approximation $1 + x \leq e^x$ which is valid for any x. We still need to split depending upon the sign of λ as for the above constants we upper bound $(np)^m$ by its largest possible value over the range $\delta n \leq m \leq 2\delta^{1/2} n$. Combining Lemmas 7.25 and 7.26 with the relevant constraints on δ in relation to λ yields Theorem 7.4.

7.5.2 Upper Bound

There is no significant (i.e. of order $e^{\lambda A}$) improvement which can be made with our current method of proof when $\lambda < 0$. This is because the gains we make computing the expectation in the proof of Theorem 7.19 are cancelled out by losses in the branching process considerations of Lemma 7.20.

When $\lambda > 0$ we cannot simply use our bound for $p = n^{-1}$ and thus an adaptation is necessary. Note that by monotonicity in p, the results of Lemmas 7.20 and 7.21 remain true for $p = n^{-1} + \lambda n^{-4/3}$ with $\lambda > 0$. The next adaptation which must be made is in equation (7.23) where now, the expectation becomes

$$\mathbb{E}(\mathcal{C}) \leq \sum_{k=1}^{n^{1/3}\log^{(2)} n} \frac{e^{k\lambda n^{-1/3}}}{k} \leq \sum_{k=1}^{n^{1/3}\log^{(2)} n} \frac{(\log n)^{\lambda}}{k} \leq 2(\log n)^{\lambda+1} = o(n^{1/6}).$$

Thus allowing us to deduce the result of Corollary 7.24 as before. Finally all that remains is to conclude the proof of Theorem 7.5. Ignoring lower order terms, the only difference to the proof compared to that of Theorem 7.19 is in the computation of $\mathbb{E}(N(A))$ where we must change the term $(np)^m$. Thus the integral in (7.27) becomes

$$\int_{An^{1/3}}^{n^{1/3}\log^{(2)} n+1} \frac{2\beta C m^{1/2}}{n^{1/2}} e^{-\frac{\gamma}{2}m^{3/2}n^{-1/2}+\lambda m n^{-1/3}} dm. \tag{7.29}$$

This is much more complex than before due to the extra term in the exponent. However we are still able to give a bound after making the obvious substitution $t = \frac{\gamma}{2}m^{3/2}n^{-1/2} - \lambda m n^{-1/3}$, we obtain

$$\begin{aligned}
\mathbb{E}(N(A)) &\leq \frac{8\beta C}{3\gamma}\int_{\frac{\gamma}{2}A^{3/2}-\lambda A}^{\infty} \frac{m^{1/2}n^{-1/2}}{m^{1/2}n^{-1/2} - \frac{4\lambda n^{-1/3}}{3\gamma}} e^{-t} dt \\
&\leq \frac{10\beta C}{3\gamma}\int_{\frac{\gamma}{2}A^{3/2}-\lambda A}^{\infty} e^{-t} dt = \frac{10\beta C}{3\gamma} e^{-\frac{\gamma}{2}A^{3/2}+\lambda A}
\end{aligned} \tag{7.30}$$

which is of the claimed form. Note the second inequality holds for A sufficiently large compared to λ.

7.6 Concluding Remarks

We have proven that inside the critical window, $p = n^{-1} + \lambda n^{-4/3}$, the largest component of $D(n, p)$ has size $\Theta_p(n^{1/3})$. Furthermore, we have given bounds on the tail probabilities of the distribution of the size of the largest component. Combining this result with previous work of Karp [60] and Łuczak [84] allows us to deduce that $D(n, p)$ exhibits a "double-jump" phenomenon at the point $p = n^{-1}$. However, there are still a large number of open questions regarding the giant component in $D(n,p)$. Perhaps the most obvious such question is to ask for an exact distribution for the size of the giant component.

Question 1. What is the limiting distribution of $n^{-1/3}|\mathcal{C}_1(D(n,p))|$ when $p = n^{-1} + \lambda n^{-4/3}$?

Of course, this has recently been answered by Goldschmidt and Stephenson [46] who in fact showed more. They showed that the sequence of strong components of $D(n,p)$ when rescaled by $n^{-1/3}$ converges to a sequence of distributions on directed multigraphs with edge lengths which are either 3-regular or cycles. However, their limit object is not particularly amenable to computations and given the strong connection between $G(n,p)$ and $D(n,p)$, it seems likely that the limit distributions, $X^\lambda = n^{-2/3}|\mathcal{C}_1(G(n,p))|$ and $Y^\lambda = n^{-1/3}|\mathcal{C}_1(D(n,p))|$ (where $p = n^{-1} + \lambda n^{-4/3}$) are closely related. For larger p, previous work [60, 85] has found that the size of the giant strongly connected component in $D(n,p)$ is related to the size of the square of the giant component in $G(n,p)$. That is, if $|\mathcal{C}_1(G(n,p)| \sim \alpha(n)n$, then $|\mathcal{C}_1(D(n,p)| \sim \alpha(n)^2 n$. Note that the result found in Theorem 7.5 is consistent with this pattern as here we have an exponent of order $A^{3/2}$ while for $G(n,p)$ a similar result is true with exponent A^3 implying that the probability we find a component of size $Bn^{2/3}$ in $G(n,p)$ is similar to the probability of finding a component of size $B^2 n^{1/3}$ in $D(n,p)$ (assuming both bounds are close to tight). As such, we make the following conjecture to explain this pattern.

Conjecture 7.27. *If X^λ and Y^λ are the distributions defined above and X_1^λ, X_2^λ are independent copies of X^λ then, $Y^\lambda = X_1^\lambda X_2^\lambda$.*

Furthermore, let us consider the transitive closure of random digraphs. The *transitive closure* of a digraph D is $cl(D)$ a digraph on the same vertex set as D and such that uv is an edge of $cl(D)$ if and only if there is a directed path from u to v in D. Equivalently, $cl(D)$ is the smallest digraph containing D such that the relation R defined by uRv if and only if uv is an edge is transitive. Karp [60] gave a linear time algorithm to compute the transitive closure of a digraph from the model $D(n,p)$ provided that $p \leq (1-\varepsilon)n^{-1}$

or $p \geq (1+\varepsilon)n^{-1}$. For all other p this algorithm runs in time $O(f(n)(n \log n)^{4/3})$ where $f(n)$ is any $\omega(1)$ function. Now that we know more about the structure of $D(n,p)$ for p close to n^{-1}, it may be possible to adapt Karp's algorithm and obtain a better time complexity.

Question 2. Does there exist a linear time algorithm to compute the transitive closure of $D(n,p)$ when $(1-\varepsilon)n^{-1} \leq p \leq (1+\varepsilon)n^{-1}$?

CHAPTER 8

ZERO-FREE REGIONS IN THE FERROMAGNETIC POTTS MODEL

8.1 Introduction

In statistical physics the Potts model is used to study interacting spins on a graph-like structure. The Potts model is a natural generalisation of both Ising model and bond percolation.

For a graph G we define the partition function of the Potts model on G as follows. Let $k \in \mathbb{N}$ this will be the number of possible spins (or colours). With each edge of G we associate a variable $w_e \in \mathbb{C}$. The k-state partition function of the Potts model is then

$$\mathbf{Z}(G; k, (w_e)_{e \in E(G)}) = \sum_{\phi: V(G) \to [k]} \prod_{\substack{uv \in E \\ \phi(u)=\phi(v)}} w_{uv}.$$

In this chapter we will only be concerned with the univariate case in which $w_e = w$ for all $e \in E(G)$. In this case the partition function is

$$\mathbf{Z}(G; k, w) := \sum_{\phi: V \to [k]} \prod_{\substack{uv \in E \\ \phi(u)=\phi(v)}} w.$$

If we consider only $w \in \mathbb{R}$, the partition function can be viewed as the normalising constant for a family of probability distributions on spin systems over G. That is we define the family of distributions

$$\mu_{G;w}(\phi) = \mathbf{Z}(G;k,w)^{-1} \prod_{\substack{uv \in E \\ \phi(u)=\phi(v)}} w. \tag{8.1}$$

This distribution is interesting in a number of places. In particular the points $w = 0$, $w = 1$ and $w \to \infty$. The distribution $\mu_{G;0}$ is the uniform distribution over proper k-colourings of G while $\mu_{G;1}$ is the uniform distribution over all colourings of G and $\mu_{G;w\to\infty}$ converges to the uniform distribution on $k^{c(G)}$ elements where $c(G)$ is the number of components of G (corresponding to the colourings with monochromatic components).

There are two general regimes for w which are studied separately. These are $w > 1$ and $w < 1$ which are referred to as the ferromagnetic and anti-ferromagnetic Potts model respectively. The names originate from the comparison to magnetism in which particles with the same spin are attracted to one-another. Clearly this is the case if $w > 1$ in 8.1 where states which have many neighbours of the same spin have the greatest weight.

The locations of the complex zeros of partition functions can be related to the existence of phase transitions in the underlying model by a seminal result of Lee and Yang [118]. In particular their work tells us that if there is no complex zero in some domain, then there will also be no phase transition there. Inspired by this Barvinok [8] was able to design efficient approximation algorithms in such domains. This contrasts with the fact that it is usually #P-hard to evaluate the partition function of the Potts model [55]. This has recently been improved from a quasi-polynomial time approximation scheme (QPTAS) to a fully polynomial time approximation scheme (FPTAS) by Patel and Regts [96] for bounded degree graphs.

The main result of this chapter is a zero-free region for the ferromagnetic Potts model. For this we will make a couple of definitions, first we define $\mathbb{N}_{\geq k}$ to be the set of integers which are at least k. We also define for $z \in \mathbb{C}$ and $d \in \mathbb{R}$, the neighbourhood of z to be $\mathcal{N}(z,d) := \{w \in \mathbb{C} : |w - z| \leq d\}$. Also, if instead of z we consider a subset of $D \subseteq \mathbb{C}$, we define the neighbourhood of D to be $\mathcal{N}(D,d) := \cup_{z\in D}\mathcal{N}(z,d)$.

Theorem 8.1. *Let $k, \Delta \in \mathbb{N}_{\geq 2}$ then there exist $\eta > 0$ and $c = c_k$ such that for any*

$$w \in \mathcal{N}\left(\left[1, 1 + \frac{c}{\Delta}\right], \eta\right)$$

and any graph $G = (V, E)$ of maximum degree at most Δ, $\mathbf{Z}(G; k, w) \neq 0$.

Note that by applying the method of Patel and Regts discussed above yields a FPTAS to approximate the partition function $\mathbf{Z}(G; k, w)$ in the interval $[1, 1 + c/\Delta]$. Bencs et al. [9] gave a similar corollary in the antiferromagnetic case and the algorithm we would obtain and the analysis would be essentially identical. As such we shall omit this detail.

When k is large in Theorem 8.1 we will take $c_k = \log(k) - 1$. For smaller choices of k (roughly $k \leq 100$) we can make some improvements to this value. See Table 8.1 for c_k when $k \leq 12$ as well as a parameter α_k which we introduce later and Appendix B for the adjustments to the proof. As we make no improvements to the work of Liu, Sinclair and Srivastava on the Ising model [81] we omit the case $k = 2$

k	3	4	5	6	7	8	9	10	11	12
α_k	1.767	1.803	1.849	1.896	1.944	1.990	2.034	2.076	2.116	2.154
c_k	2.171	2.330	2.472	2.600	2.716	2.820	2.916	3.003	3.084	3.160

Table 8.1: Lower bounds on c_k for small k.

Theorem 8.1 adds to a large body of literature on zero-free regions for the Potts model partition function. In particular there are a number of results for zeros of the partition function on lattice graphs [7, 20, 21, 22, 106] as well as for arbitrary graphs [8, 9, 41, 107]. Most of the results for arbitrary bounded degree graphs require a sufficiently large number of colours and are for the anti-ferromagnetic regime. For example, Sokal [107] proved that for any graph of maximum degree Δ there exists a constant $C < 7.97$ such that if $|w| \leq 1$, then for any $k \geq C\Delta$, $\mathbf{Z}(G; k, w) \neq 0$. This was subsequently improved to $C < 6.91$ by Procacci and Fernández [41] and the condition that w lie in the unit disk was loosened in [54].

Our main result may be viewed as removing the requirement for sufficiently many colours in a slightly different domain, where we take an ν-neighbourhood of a real interval as our domain. This is similar to the approach taken by Bencs et al. [9] who looked at the interval $[0, 1]$. Our proof method is also similar to theirs which is based upon a method of Barvinok [8]. This method involves considering a restricted partition function, where we fix some spins. This gives us a "graph with bounded maximum degree and boundary conditions" which the restricted partition function is the partition function of.

We now define the restricted partition function for the Potts model. For a list $W = w_1 \dots w_m$ of distinct vertices of V and a list $L = \ell_1 \dots \ell_m$ of pre-assigned colours in $[k]$ for the vertices in W which need not be distinct the *restricted partition function* $\mathbf{Z}_L^W(G)$ is defined by

$$\mathbf{Z}_L^W(G) := \sum_{\substack{\phi: V \to [k] \\ \phi \text{ respects } (W,L)}} \prod_{\substack{uv \in E \\ \phi(u)=\phi(v)}} w_{uv},$$

where we say that ϕ respects (W, L) if for all $i = 1 \dots, m$ we have $\phi(w_i) = \ell_i$. We say the vertices $w_1, \dots, w_m$ are *fixed* and refer to the remaining vertices in V as *free* vertices. The length of W (respectively L), written $|W|$ (respectively $|L|$) is the length of the list. Given a list of distinct vertices $W' = w_1 \dots w_m$, and a vertex u (distinct from $w_1, \dots, w_m$) we write $W = W'u$ for the concatenated list $W = w_1 \dots w_m u$ and we use similar notation $L'\ell$ for concatenation of lists of colours.

The remainder of this chapter is organised as follows. In Section 8.2 we collect some preliminary lemmas which will be useful in our proof of Theorem 8.1. Following this we prove a generalisation of our main theorem for restricted partition functions in Section 8.3 modulo one technical lemma about the ratios of very similar restricted partition functions which is proved in Section 8.4. Finally we conclude in Section 8.5 with some remarks and open questions.

8.2 Preliminaries

In this section we will gather a few tools which will come in useful in the proofs during Sections 8.3 and 8.4. First, we state a lemma of Barvinok which is useful for evaluating sums of restricted partition functions.

Lemma 8.2 (Barvinok [8, Lemma 3.6.3])**.** *Let $u_1, \ldots, u_n \in \mathbb{R}^2$ be non-zero vectors such that the angle between any two vectors u_i and u_j is at most α for some $\alpha \in [0, 2\pi/3)$. Then the u_i all lie in a cone of angle at most α and*

$$\left| \sum_{i=1}^{n} u_i \right| \geq \cos(\alpha/2) \sum_{i=1}^{n} |u_i|.$$

Furthermore the following simple corollary of the cosine rule will come in handy.

Lemma 8.3. *Let z, z' be two complex numbers at an angle of at most $\pi/3$, then $|z - z'| \leq \max\{|z|, |z'|\}$.*

Proof. Recall the cosine rule, for a triangle with sides a, b and c; and angles A, B and C where side a is not adjacent to angle A, then

$$|a|^2 = |b|^2 + |c|^2 - 2|b||c| \cos(A)$$

where $|a|$ is the length of side a. Now consider the triangle with vertices in $\mathbb{C}$ at the origin, z and z'. The sides have length $|z|, |z'|$ and $|z - z'|$ and the angle at the origin is the angle $\theta \leq \pi/3$ between z and z'. As $\cos(x) \geq 0.5$ for $x \leq \pi/3$,

$$|z - z'|^2 \leq |z|^2 + |z'|^2 - |z||z'| \leq \max\{|z|, |z'|\}.$$

□

8.3 An induction for k Colours

In this section we prove a generalisation of Theorem 8.1. To see that it implies Theorem 8.1, take $W = L = \emptyset$ in 1.

Lemma 8.4. *Let $\Delta \in \mathbb{N}_{\geq 3}$ and let $k \in \mathbb{N}_{\geq 4}$ such that $\log k \leq \Delta$. Then there exist constants $\pi/(3\Delta) > \theta > \varepsilon > 0$ and $0 < \alpha < c$ such that for any $w \in [1, 1 + c/\Delta]$, there exists $\eta > 0$ such that for any $w' \in \mathbb{C}$ satisfying $|w - w'| \leq \eta$ and $|w'| \leq |w|$ and any graph G of maximum degree at most Δ the following hold for $\mathbf{Z}(G) = \mathbf{Z}(G; k, w')$.*

1. *For all lists W of distinct vertices of G and all lists of pre-assigned colours L of length $|W|$, $\mathbf{Z}_L^W(G) \neq 0$.*

2. *For all lists $W = W'u$ of distinct vertices of G such that u is a leaf and any two lists $L'l$,$L'l'$ of length $|W|$*

   (a) *If the unique neighbour v of u is free,*

      i. *The angle between vectors $\mathbf{Z}_{L'\,l}^{W'\,u}(G)$ and $\mathbf{Z}_{L'\,l'}^{W'\,u}(G)$ is at most θ.*

      ii.
$$\frac{|\mathbf{Z}_{L'\,l}^{W'\,u}(G)|}{|\mathbf{Z}_{L'\,l'}^{W'\,u}(G)|} \leq 1 + \frac{\alpha}{\Delta}.$$

   (b) *If the unique neighbour v of u is fixed,*

      i. *The angle between vectors $\mathbf{Z}_{L'\,l}^{W'\,u}(G)$ and $\mathbf{Z}_{L'\,l'}^{W'\,u}(G)$ is at most ε.*

      ii.
$$\frac{|\mathbf{Z}_{L'\,l}^{W'\,u}(G)|}{|\mathbf{Z}_{L'\,l'}^{W'\,u}(G)|} \leq 1 + \frac{c}{\Delta}.$$

3. *For all lists $W = W'u$ of distinct vertices of G and for all lists of pre-assigned colours L' of length $|W'|$. Let d be the number of free neighbours of u and let $b = \Delta - d$. Then for any pair of colours l, l',*

   (a) *The angle between vectors $\mathbf{Z}_{L'\,l}^{W'\,u}(G)$ and $\mathbf{Z}_{L'\,l'}^{W'\,u}(G)$ is at most $d\theta + b\varepsilon$.*

(b)

$$\frac{|\mathbf{Z}^{W'}_{L'}{}^{u}_{l}(G)|}{|\mathbf{Z}^{W'}_{L'}{}^{u}_{l'}(G)|} \leq (1+\alpha/\Delta)^{d}(1+c/\Delta)^{\Delta-d}$$

To prove this result we need some definitions and an auxiliary lemma.

We define rational functions in two variables z_0, z and respectively $k-1$ variables $z_0, \ldots, z_{k-2}$ by

$$R(z_0, z; w, k) = \frac{wz_0 + (k-2)z + 1}{z_0 + (k-2)z + w},$$

$$R_k(z_0, z_1, \ldots, z_{k-2}; w) = \frac{wz_0 + z_1 + \ldots + z_{k-2} + 1}{z_0 + z_1 + \ldots + z_{k-2} + w}.$$

Consider the cone

$$C(\theta) := \{z = re^{i\vartheta} \mid r \geq 0 \text{ and } |\vartheta| \leq \theta\},$$

and define for $d = 0, \ldots, \Delta$ and $c \geq 0$,

$$\begin{aligned} K(\theta, d, \alpha, c, \varepsilon) &:= C(d\theta + (\Delta - d)\varepsilon) \\ &\cap \left\{z \mid (1+c/\Delta)^{d-\Delta}(1+\alpha/\Delta)^{-d} \leq |z| \leq (1+c/\Delta)^{\Delta-d}(1+\alpha/\Delta)^{d}\right\}. \end{aligned}$$

Lemma 8.5. *Let $\Delta \in \mathbb{N}_{\geq 3}$ and let $k \in \mathbb{N}$ such that $\log(k) \leq \Delta$. Suppose that $\alpha = \log(k)/2 - 1$ and $c = \log(k) - 1$. Then there exist $1 > \theta > \varepsilon > \eta > 0$ such that for each $d = 0, \ldots, \Delta$, and any $z_0, \ldots, z_{k-2} \in K_d := K(\theta, d, \alpha, c, \varepsilon)$ such that for each i, j, $z_i/z_j \in K_d$ and any $w \in [1, 1+c/\Delta]$ and any $w' \in \mathbb{C}$ such that $|w - w'| \leq \eta$ and $|w'| \leq |w|$ the ratio $R = R_k(z_0, z_1, \ldots, z_{k-2}; w')$ satisfies*

$$(1+\alpha/\Delta)^{-1} < |R| < 1 + \alpha/\Delta \quad \textit{and} \quad |\arg(R)| < \theta. \tag{8.2}$$

We will prove this lemma in the next section. We first utilize it to prove Lemma 8.4.

Proof of Lemma 8.4. We prove this theorem by induction on the number of free vertices of G. For the base case, we have no free vertices and so every vertex is fixed. Therefore $\mathbf{Z}^{W}_{L}(G)$ is a product of non-zero terms, hence is non-zero, proving 1. Statement 2.(a) is vacuous as there are no free vertices. Statement 2.(b) follows as the products $\mathbf{Z}^{W'}_{L'}{}^{u}_{\ell}(G)$ and $\mathbf{Z}^{W'}_{L'}{}^{u}_{\ell'}(G)$ differ in at most one term. Thus their ratio is either 1, w or w^{-1}. Similarly we deduce Statement 3. from the fact that the products $\mathbf{Z}^{W'}_{L'}{}^{u}_{\ell}(G)$ and $\mathbf{Z}^{W'}_{L'}{}^{u}_{\ell'}(G)$ differ in at most Δ terms.

Now, we assume that Statements 1., 2. and 3. hold for graphs with $r \geq 0$ free vertices. We prove the statements for $r+1$ free vertices. First, we shall prove 1.

Suppose that u is a free vertex. Note that $\mathbf{Z}^{W}_{L}(G) = \sum_{j=1}^{k} \mathbf{Z}^{W}_{L}{}^{u}_{j}(G)$. As each term in the sum on the right hand side of this expression has one fewer free vertex, we may apply induction to deduce that all of these terms are non-zero by 1. Furthermore, by 3. each pair has angle at most $d\theta + (\Delta - d)\varepsilon$ where d is the number of free neighbours of u. Lemma 8.2 tells us that the $\mathbf{Z}^{W}_{L}{}^{u}_{j}$ all lie in a cone of angle at most $d\theta + (\Delta - d)\varepsilon$ and

$$|\mathbf{Z}^{W}_{L}(G)| = \left| \sum_{j=1}^{k} \mathbf{Z}^{W}_{L}{}^{u}_{j}(G) \right| \geq \cos(d\theta/2 + (\Delta - d)\varepsilon/2) \sum_{j=1}^{k} |\mathbf{Z}^{W}_{L}{}^{u}_{j}(G)| \neq 0.$$

Next, we shall prove 2.(a) so consider the ratios,

$$R_{j,\ell}(G) = \frac{\mathbf{Z}^{W'}_{L'}{}^{u}_{j}(G)}{\mathbf{Z}^{W'}_{L'}{}^{u}_{\ell}(G)}, \qquad\qquad R^{v}_{j,\ell}(G) = \frac{\mathbf{Z}^{W'}_{L'}{}^{v}_{j}(G-u)}{\mathbf{Z}^{W'}_{L'}{}^{v}_{\ell}(G-u)}.$$

As v is the unique neighbour of u and is free, we may write,

$$R_{j,\ell}(G) = \frac{\sum_i \mathbf{Z}^{W}_{L}{}^{uv}_{ji}(G)}{\sum_i \mathbf{Z}^{W}_{L}{}^{uv}_{\ell i}(G)} = \frac{w\mathbf{Z}^{W}_{L}{}^{v}_{j}(G-u) + \sum_{i \notin \{j,\ell\}} \mathbf{Z}^{W}_{L}{}^{v}_{i}(G-u) + \mathbf{Z}^{W}_{L}{}^{v}_{\ell}(G-u)}{\mathbf{Z}^{W}_{L}{}^{v}_{j}(G-u) + \sum_{i \notin \{j,\ell\}} \mathbf{Z}^{W}_{L}{}^{v}_{i}(G-u) + w\mathbf{Z}^{W}_{L}{}^{v}_{\ell}(G-u)}.$$

Dividing both the numerator and denominator by $\mathbf{Z}^{W}_{L}{}^{v}_{\ell}(G-u)$ (which by the inductive hypothesis is non-zero) we obtain,

$$\frac{wR^{v}_{j^*,\ell^*}(G) + \sum_{i \neq j^*,\ell^*} R^{v}_{i,\ell^*}(G) + 1}{R^{v}_{j^*,\ell^*}(G) + \sum_{i \neq j^*,\ell^*} R^{v}_{i,l}(G) + w} = R_k(R^{v}_{j^*,\ell^*}(G), R^{v}_{1,\ell^*}(G), \ldots, R^{v}_{k,\ell^*}(G); w). \tag{8.3}$$

where the function R_k in (8.3) takes as arguments all $R^{v}_{i,\ell^*}(G)$ for $i \neq \ell^*$ precisely once (and so takes precisely $k-1$ arguments as expected.)

Suppose that v has d free neighbours that are not u. Since $G-u$ has one fewer free vertex than G, we may apply the inductive hypothesis. By 3. we find that for any $i \neq \ell^*$, we have $R^v_{i,\ell}(G) \in K(\theta, d, \alpha, c, \varepsilon)$. However, we also have that for any $i, j \neq \ell^*$, that

$$\frac{R^v_{i,\ell^*}(G)}{R^v_{j,\ell^*}(G)} = \frac{\mathbf{Z}^{W'\,v}_{L'\,i}(G-u)}{\mathbf{Z}^{W'\,v}_{L'\,j}(G-u)} = R^v_{i,j}(G) \in K(\theta, d, \alpha, c, \varepsilon).$$

To prove 2.$(a)i$. observe that the angle between $\mathbf{Z}^{W'\,u}_{L'\,j}$ and $\mathbf{Z}^{W'\,u}_{L'\,\ell}$ is precisely the angle of $R_{j,\ell}(G)$ from the real axis in $\mathbb{C}$ and so is bounded by the absolute value of the argument of $R_{j,\ell}(G)$, which by Lemma 8.5 bounded by θ as desired. Statement 2.$(a)ii$. also follows immediately from Lemma 8.5.

For the proof of 2.(b), we note that as v is fixed, then

$$\mathbf{Z}^{W'\,u}_{L'\,j}(G) \in \{w^{-1}\mathbf{Z}^{W'\,u}_{L'\,j}(G), \mathbf{Z}^{W'\,u}_{L'\,j}(G), w\mathbf{Z}^{W'\,u}_{L'\,j}(G)\}$$

from which both i. and ii. follow.

Finally, we prove 3.. To do so we consider the graph $G \star u$ which is formed as follows. Let $v_1, \ldots, v_r$ be the neighbours of u ordered arbitrarily. Let $u_1, \ldots, u_r$ be r new vertices which will be copies of u. Then $G \star u$ is the graph obtained by deleting u and its incident edges, adding the vertices $u_1, \ldots, u_r$ and edges $u_1v_1, \ldots, u_rv_r$. Furthermore, $G \star u$ inherits any colouring of G and if u is coloured, all of the new vertices inherit this colour. Note that if u is coloured, then the graph $G \star u$ has the same partition function as G. Also, in this case $G \star u$ has the same number of free vertices as G. This allows us to prove 3. from 2. by changing the colour of one copy of u at a time. That is,

$$\frac{\mathbf{Z}^{W'\,u}_{L'\,j}(G)}{\mathbf{Z}^{W'\,u}_{L'\,l}(G)} = \frac{\mathbf{Z}^{W'\,u_1\ldots u_r}_{L'\;\,j\,\ldots\,j}(G \star u)}{\mathbf{Z}^{W'\,u_1\ldots u_r}_{L'\;\,l\,\ldots\,l}(G \star u)} = \prod_{i=1}^{r} \frac{\mathbf{Z}^{W'\,u_1\ldots u_{i-1}\,u_i\ldots u_r}_{L'\;\,j\,\ldots\;\,j\;\;\,l\,\ldots\,l}(G \star u)}{\mathbf{Z}^{W'\,u_1\ldots u_i\,u_{i+1}\ldots u_r}_{L'\;\,j\,\ldots\,j\;\;\,l\;\;\,\ldots\,l}(G \star u)} \tag{8.4}$$

By 2. each of the terms in the product in (8.4) has angle at most θ and absolute value at most $1 + \alpha/\Delta$ (if u_i is free) or angle at most ε and absolute value at most $1 + c/\Delta$ (if u_i is fixed). As u has d free neighbours and at most $\Delta - d$ fixed neighbours, this allows us to conclude 3.(a) and 3.(b) completing the induction. $\square$

8.4 Proof of Lemma 8.5

To prove the lemma, we first note that c and α satisfy the following inequality,

$$\frac{ce^c}{e^c + k - 1} < \alpha. \tag{8.5}$$

We will also require a technical lemma concerning the real and imaginary parts of the ratios $R(z_1, z_2; w, k)$.

Lemma 8.6. *Let $z_1, z_2 \in \mathbb{C}$ be defined as $z_1 = xe^{i\theta_x}$, $z_2 = ye^{i\theta_y}$ with $x, y \in \mathbb{R}^+$ and $\theta_x, \theta_y \in [0, 2\pi)$ and suppose $w \in [1, 1 + \frac{c}{\Delta}]$ is real. Then, the real and imaginary parts of $R(z_1, z_2; w, k)$ are as follows where N is a non-zero constant,*

$$\begin{aligned} \Re(R(z_1, z_2; w, k)) =& N(wx^2 + (w+1)(k-2)xy\cos(\theta_x - \theta_y) + (k-2)^2y^2 \\ &+ (w^2+1)x\cos(\theta_x) + (w+1)(k-2)y\cos(\theta_y) + w), \end{aligned} \tag{8.6}$$

$$\begin{aligned} \Im(R(z_1, z_2; w, k)) =& N(w-1)((k-2)xy\sin(\theta_x - \theta_y) \\ &+ (1+w)x\sin(\theta_x) + (k-2)y\sin(\theta_y)). \end{aligned} \tag{8.7}$$

Hence, provided that θ_x and θ_y are small and setting $\theta = \max(|\theta_x|, |\theta_y|, |\theta_x - \theta_y|)$,

$$\left|\frac{\Im(R(z_1, z_2; w, k))}{\Re(R(z_1, z_2; w, k))}\right| \le \frac{\frac{c}{\Delta}\left((k-2)xy|\theta_x - \theta_y| + (2 + \frac{c}{\Delta})x|\theta_x| + (k-2)y|\theta_y|\right)}{(x + (k-2)y + 1 + \frac{c}{\Delta})(x + (k-2)y + 1 + x\frac{c}{\Delta}) - O(\theta^2)}. \tag{8.8}$$

Where one can compute the $O(\theta^2)$ term to be $\theta^2((w+1)(k-2)(x+1)y + (w^2+1)x)/2$.

Proof. We may write $z_1 = x\cos(\theta_x) + ix\sin(\theta_x)$ and $z_2 = y\cos(\theta_y) + iy\sin(\theta_y)$. Hence,

$$\begin{aligned}R(z_1, z_2; w, k) &= \frac{w(x\cos(\theta_x) + ix\sin(\theta_x)) + (k-2)(y\cos(\theta_y) + iy\sin(\theta_y)) + 1}{x\cos(\theta_x) + ix\sin(\theta_x) + (k-2)(y\cos(\theta_y) + iy\sin(\theta_y)) + w} \\ &= \frac{wx\cos(\theta_x) + (k-2)y\cos(\theta_y) + 1 + i(wx\sin(\theta_x) + (k-2)y\sin(\theta_y))}{x\cos(\theta_x) + (k-2)y\cos(\theta_y) + w + i(x\sin(\theta_x) + (k-2)y\sin(\theta_y))}.\end{aligned} \tag{8.9}$$

Rationalising the denominator in (8.9), we obtain the following in which we write c_x for $\cos(\theta_x)$ and similarly define c_y, s_x and s_y to simplify notation.

$$\begin{aligned}R(z_1, z_2; w, k) =& N^{-1}\left(wxc_x + (k-2)yc_y + 1 + i(wxs_x + (k-2)ys_y)\right) \\ & \times \left(xc_x + (k-2)yc_y + w - i(xs_x + (k-2)ys_y)\right)\end{aligned} \tag{8.10}$$

where $N = |xc_x + (k-2)yc_y + w + i(xs_x + (k-2)ys_y)|^2$. Expanding the expression in (8.10), the real and imaginary parts are given by the following expressions.

$$\begin{aligned}\Re(R(z_1, z_2; w, k)) =& N^{-1}(wx^2c_x^2 + (w+1)(k-2)xyc_xc_y + (k-2)^2c_y^2 \\ & + wx^2s_x^2 + (w+1)(k-2)xys_xs_y + (k-2)^2s_y^2 \\ & + (w^2+1)xc_x + (w+1)(k-2)yc_y + w) \\ \Im(R(z_1, z_2; w, k)) =& N^{-1}((k-2)xy(c_xs_y + ws_xc_y) - (k-2)xy(wc_xs_y + s_xc_y) \\ & + (w^2-1)xs_x + (w-1)(k-2)ys_y).\end{aligned}$$

Combining these expressions with the trigonometric identities

$$\begin{aligned}\cos^2(\vartheta) + \sin^2(\vartheta) &= 1 \\ \sin(\alpha - \beta) &= \sin(\alpha)\cos(\beta) - \sin(\beta)\cos(\beta) \\ \cos(\alpha - \beta) &= \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta)\end{aligned}$$

yields the expressions (8.6) and (8.7) as claimed.

By an application of the triangle law combined with an application of the approximations $|\sin(\theta)| \leq |\theta|$ and $\cos(\theta) \geq 1 - \theta^2/2$, we obtain

$$
\begin{aligned}
|\Im(R(z_1, z_2; w, k))| \leq & N^{-1}(w-1)\left((k-2)xy|\theta_x - \theta_y| + (1+w)x|\theta_x| + (k-2)y|\theta_y|\right), \\
& \hfill (8.11) \\
\Re(R(z_1, z_2; w, k)) \geq & N^{-1}\left((wx + (k-2)y + 1)(x + (k-2)y + w)\right. \\
& \left. - ((w+1)(k-2)(x+1)y + (w^2+1)x)\theta^2/2\right). \qquad (8.12)
\end{aligned}
$$

Dividing (8.11) by (8.12), noting that when $\theta \to 0$ this is maximised when $w = 1 + \frac{c}{\Delta}$ and regrouping some terms yields the bound (8.8). $\square$

We can now give a proof of Lemma 8.5

Proof of Lemma 8.5. It suffices to prove the lemma for w' real. Indeed if it holds for these w', then by continuity and since the inequalities are strict it follows that there exists a small enough $\eta > 0$ such that the lemma still holds for w' in an η-neighbourhood of any real $w' \in [1, 1 + c/\Delta]$ for some small enough $\eta > 0$.

Fix $d \in \{0, \ldots, \Delta\}$. First we observe that we may assume that $|R| \geq 1$. Indeed, if $|R| < 1$, then

$$1/R = \frac{z_0 + \sum_{i=1}^{k-2} z_i + w}{wz_0 + \sum_{i=1}^{k-2} z_i + 1} = \frac{1 + \sum_{i=1}^{k-2} z_i/z_0 + w/z_0}{w + \sum_{i=1}^{k-2} z_i/z_0 + 1/z_0}$$

and $|1/R| > 1$. Since for each $i, j \geq 0$, the pairs z_i/z_0 and z_j/z_0 also satisfy our assumptions this shows our claim.

Next define $z = \frac{1}{k-2} \sum_{i=1}^{k-2} z_j$. Then by construction note that $R_k(z_0, z_1, \ldots, z_{k-2}; w) = R(z_0, z; w, k)$. We will therefore analyse $R(z_0, z; w, k)$. Note that $z \in C(d\theta + (\Delta - d)\varepsilon)$ and by Barvinok's lemma (Lemma 8.2) we have

$$\cos(d\theta/2 + (\Delta - d)\varepsilon/2)(1 + c/\Delta)^{d-\Delta}(1 + \alpha/\Delta)^{-d} \leq |z| \leq (1 + c/\Delta)^{\Delta - d}(1 + \alpha/\Delta)^d.$$

Next we observe that

$$|R(z_0, z; w, k)| = \left|1 + \frac{(w-1)z_0 + (1-w)}{z_0 + (k-2)z + w}\right| \leq 1 + \frac{\frac{c}{\Delta}|z_0 - 1|}{|z_0 + (k-2)z + w|}. \tag{8.13}$$

Lower bounding the denominator of (8.13) may be done with another application of Barvinok's lemma. For the numerator we apply Lemma 8.3 as the angle between z_0 and 1 may be assumed to be less than $\pi/3$. This allows us to deduce that

$$|R(z_0, z; w, k)| \leq 1 + \frac{\frac{c}{\Delta}\max\{|z_0|, 1\}}{\cos(d\theta/2 + (\Delta - d)\varepsilon/2))(|z_0| + (k-2)|z| + 1)}.$$

To maximize the above quantity clearly one should take $|z|$ as small as possible, so if $|z_0| < 1$, we take $|z| = \cos(d\theta/2 + (\Delta - d)\varepsilon/2)(1 + c/\Delta)^{d-\Delta}(1 + \alpha/\Delta)^{-d}$ and rearrange to deduce that $|R(z_0, z; w, k)| < 1 + \alpha/\Delta$ as θ is small and applying (8.5).

Otherwise, we take $|z| = \cos(d\theta/2 + (\Delta - d)\varepsilon/2)(1 + c/\Delta)^{d-\Delta}(1 + \alpha/\Delta)^{-d}|z_0|$ and conclude similarly. Note that the fact $|z| \geq \cos(d\theta/2 + (\Delta - d)\varepsilon/2)(1 + c/\Delta)^{d-\Delta}(1 + \alpha/\Delta)^{-d}|z_0|$ follows by applying Barvinok's lemma to the terms z_i/z_0 for $i = 1, \ldots, k-2$.

To prove the bound on the argument of $R(z_0, z; w, k)$ we use the inequality, $|\beta| \leq |\tan(\beta)|$. It therefore suffices to bound the ratio $\frac{|\Im R(z_0,z;w,k)|}{|\Re R(z_0,z;w,k)|} = \tan(\arg(R(z_0, z; w, k)))$, which by Lemma 8.6 is bounded by,

$$\frac{\frac{c}{\Delta}\left((k-2)|z_0 z||\theta_0 - \theta_z| + (2 + \frac{c}{\Delta})|z_0\theta_0| + (k-2)|z\theta_z|\right)}{(|z_0| + (k-2)|z| + 1 + \frac{c}{\Delta})(|z_0| + (k-2)|z| + 1 + |z_0|\frac{c}{\Delta}) - \frac{\theta^2}{2}f(w, k, z_0, z)}. \tag{8.14}$$

Here we have $f(w, k, z_0, z) = (w+1)(k-2)(|z_0|+1)|z| + (w^2+1)|z_0|$. Now suppose that we can prove that

$$\frac{\left((k-2)|z_0 z||\theta_0 - \theta_z| + (2 + \frac{c}{\Delta})|z_0\theta_0| + (k-2)|z\theta_z|\right)}{(|z_0| + (k-2)|z| + 1 + \frac{c}{\Delta})(|z_0| + (k-2)|z| + 1 + |z_0|\frac{c}{\Delta})} \leq \frac{\Delta\theta}{c} - \tau, \tag{8.15}$$

for some fixed constant τ. Then in the bound (8.14) the influence of the term $\frac{\theta^2}{2}f(w, k, z_0, z)$ in the denominator will decrease as $\theta \to 0$ (and hence $\varepsilon \to 0$) while maintaining the inequality (8.15). So for θ small enough, we have that (8.14) is at most θ and hence $|\arg(R(z_0, z; w, k))| \leq \theta$, as desired.

We will now show that (8.15) holds. So, first note that

$$\begin{aligned}&\frac{\big((k-2)|z_0 z||\theta_0-\theta_z| + (2+\frac{c}{\Delta})|z_0\theta_0| + (k-2)|z\theta_z|\big)}{(|z_0| + (k-2)|z| + 1 + \frac{c}{\Delta})(|z_0| + (k-2)|z| + 1 + |z_0|\frac{c}{\Delta})}\\ &\leq \frac{((k-2)|z_0 z||\theta_0-\theta_z| + 2|z_0\theta_0| + (k-2)|z\theta_z|)}{(|z_0| + (k-2)|z| + 1)^2},\end{aligned} \tag{8.16}$$

which can be observed by computing the derivative of the left hand side of (8.15) with respect to $\frac{c}{\Delta}$ and noting it is strictly negative. Now, we maximize (8.16) so first we show that there is a maximum point where exactly two of $|\theta_0 - \theta_z|, |\theta_0|, |\theta_z|$ are as large as possible and one is zero. To see this, first note that clearly at least one of $|\theta_0 - \theta_z|, |\theta_0|, |\theta_z|$ must be as large as possible i.e. equal to $d\theta + (\Delta - d)\varepsilon$. In fact exactly two of these must be maximised as the maximisation with respect to the θ terms only is of the form $g(\theta_0, \theta_z) = a|\theta_0 - \theta_z| + b|\theta_0| + c|\theta_z|$ for constants $a, b, c > 0$. So if $|\theta_0 - \theta_z| = d\theta + (\Delta - d)\varepsilon$ for example, then if $b \geq c$ we may set $\theta_0 = d\theta + (\Delta - d)\varepsilon$, $\theta_z = 0$ increasing $g(\theta_0, \theta_z)$. Similar logic allows one to conclude that two of $|\theta_0 - \theta_z|, |\theta_0|, |\theta_z|$ are equal to $d\theta + (\Delta - d)\varepsilon$ and one is 0 in every other case.

This leaves us with three maximisation problems over $R_d \subseteq \mathbb{R}^2$ defined by

$$\begin{aligned} R_d = \{(x,y) | (1+c/\Delta)^{d-\Delta}(1+\alpha/\Delta)^{-d} &\leq x \leq (1+c/\Delta)^{\Delta-d}(1+\alpha/\Delta)^d, \\ \cos(d\theta/2 + (\Delta-d)\varepsilon/2)(1+c/\Delta)^{d-\Delta}(1+\alpha/\Delta)^{-d} &\leq y \leq (1+c/\Delta)^{\Delta-d}(1+\alpha/\Delta)^d, \\ \cos(d\theta/2 + (\Delta-d)\varepsilon/2)(1+c/\Delta)^{d-\Delta}(1+\alpha/\Delta)^{-d} &\leq y/x \leq (1+c/\Delta)^{\Delta-d}(1+\alpha/\Delta)^d\}. \end{aligned}$$

We enlarge the region slightly obtaining the region $\widetilde{R_d} \subseteq \mathbb{R}^2$ defined by

$$\begin{aligned} \widetilde{R_d} &= \left\{(x,y) \middle| \exp\left(-(\frac{d}{\Delta}\alpha + (1-\frac{d}{\Delta})c)\right) \leq x, y, y/x \leq \exp\left(\frac{d}{\Delta}\alpha + (1-\frac{d}{\Delta})c\right)\right\} \\ &= \{(x,y) | e k^{\frac{d}{2\Delta}-1} \leq x, y, y/x \leq k^{1-\frac{d}{2\Delta}}/e\}. \end{aligned}$$

The functions to maximise are,

$$\begin{aligned}
f_1(x,y) &= \frac{(k-2)(xy+y)}{(x+(k-2)y+1)^2},\\
f_2(x,y) &= \frac{(k-2)xy+2x}{(x+(k-2)y+1)^2},\\
f_3(x,y) &= \frac{2x+(k-2)y}{(x+(k-2)y+1)^2}.
\end{aligned}$$

First we look at f_1, it has critical points along the line $x+1=(k-2)y$ where it attains its maximum value of $1/4$. However, note that due to our choice of c and α, this line does not lie inside of $\widetilde{R_d}$, hence the maximum must be attained at a boundary point. Furthermore both f_2 and f_3 have no critical points strictly inside the first quadrant, so again their maxima must be attained at a boundary point. This allows us to reduce the problem to 18 univariate maximisation problems, each of which has maximum at most $3k^{-\frac{d}{2\Delta}}/e$ over $\widetilde{R_d}$ (see Appendix B for details).

Provided that $d \geq 1$ and $\varepsilon < \theta/6\Delta$, it is always the case that $d\theta + (\Delta - d)\varepsilon \leq 7d\theta/6$. Substituting this back in, we find that (8.16) is upper bounded by

$$\frac{7d\theta}{2e}k^{-\frac{d}{2\Delta}}. \tag{8.17}$$

We divide (8.17) by Δ, write $x = \frac{d}{\Delta}$ and maximise the resulting expression for $x > 0$. The maximum is attained at $x = 2/\log(k)$ and is equal to $\frac{7\theta}{e^2\log(k)} < \frac{\theta}{\log(k)-1}$. If $d = 0$, then as f_1, f_2 and f_3 are all bounded above by 1, provided $\varepsilon < \frac{\theta}{\log(k)}$, the left hand side of (8.15) at most $\varepsilon\Delta < \theta\Delta/c$. This completes the proof of (8.15) and hence of the lemma.

□

8.5 Concluding Remarks

We have proven that in any graph with maximum degree Δ the partition function of the ferromagnetic Potts model with k colours is zero-free in an open set in $\mathbb{C}$ containing the interval $[1, 1+\frac{\log(k)-1}{\Delta}]$. This is approximately half way to a hardness threshold [43] which can be found at the point

$$\mathfrak{B}_o = \frac{k-2}{(k-1)^{1-2/\Delta}-1} = 1 + \frac{2(k-1)\log(k-1)}{k-2} \cdot \frac{1}{\Delta} + O\left(\frac{1}{\Delta^2}\right). \tag{8.18}$$

It would be interesting to know if it is possible to get closer to this threshold in general. For k very large compared to Δ (roughly $k \geq \Delta^{\Omega(\Delta)}$) a very recent preprint of Borgs et al. [12, Theorem 2.4] manages to get further and gives a zero-free region when the parameter is at most $1 + \frac{3\log(k)}{2\Delta}$. Of course this is a very large choice of k whereas our results hold for any $k \geq 3$. It would be of interest to investigate whether the dependence of k on Δ in this theorem could be reduced or removed. Of course if it were possible to completely remove this dependence then this would also improve our result.

Finally, note that in our proof we look at a graph vertex by vertex and look at their neighbourhoods. Possibly exploring further and looking at second or third neighbourhoods would be able to improve the results which we obtained.

APPENDIX A

PROOF OF EXISTENCE FOR $\delta_\ell^*(H)$

Lemma A.1. *The limit, $\delta_\ell^* = \lim_{m\to\infty} \delta_\ell(H, mh)$ exists.*

Proof. To prove the existence of this limit, we show convergence to the lim inf. So, let $\delta_\ell^- = \liminf_{m\to\infty} \delta_\ell(H, mh)$. Let $\varepsilon > 0$, by definition of the lim inf there exists m_ε such that $\delta_\ell(H, m_\varepsilon h) \leq \delta_\ell^- + \varepsilon$. That is every graph with $m_\varepsilon h$ vertices and minimum ℓ-degree at least $(\delta_\ell^- + \varepsilon)(m_\varepsilon h)^{r-\ell}$ has an H-factor. Without loss of generality, we may assume that m_ε is large enough that random subgraphs picked as in Lemma 5.3 lose at most $\varepsilon|V_i|^{r-\ell}$ from their minimum degree. Now, pick $n \gg m_\varepsilon h^2$ such that $h|n$ and let G be any r-graph on n vertices with minimum ℓ-degree at least $(\delta_\ell^- + 2\varepsilon)n^{r-\ell}$. We shall show that G has an H-factor.

If G has an H-factor, then we are done. So suppose for a contradiction that G has no H-factor. Let F_0 be a largest H-factor of G. Extend F_0 to F_0^* in G by arbitrarily adding vertex disjoint copies of any graph on h vertices such that F_0^* spans $V(G)$. Let H_0 be any $H' \in F_0^*$ which is not a copy of H. Let $\mathcal{P} = (V_1, \dots, V_h)$ be the random transverse partition obtained from Lemma 5.3 with $m = m_\varepsilon h$. With positive probability, for all $i \in [h]$ we have $|V_i| = m$ and $\delta_\ell(G[V_i]) \geq (\delta_\ell^- + \varepsilon)m^{r-\ell}$. So there exists a set X with $|X| = m$ whose corresponding partition satisfies these properties. By definition of m_ε, we may pick an H-factor in each $G[V_i]$. By removing X from F_0 and adding these new H-factors we obtain an H-factor that is strictly larger than F_0, as H_0 was not a copy of H. This is a contradiction with the maximality of F_0. Hence G has an H-factor.

Thus, for all $\epsilon > 0$ we have $\delta_\ell(H, mh) \leq \delta_\ell^- + 2\varepsilon$ provided that m is sufficiently large. In particular, $\limsup_{m\to\infty} \delta_\ell(H, mh) \leq \delta_\ell^- + 2\varepsilon$ for any $\varepsilon > 0$. Taking $\varepsilon \to 0$ allows us to deduce that the $\liminf$ and $\limsup$ are both the same and hence the limit in (5.1) exists.

□

APPENDIX B

OPTIMISING PARAMETERS IN THE POTTS MODEL

B.1 18 Maximisation Problems

We look at the maximisation problems coming from 8.4 and claim that each has an upper bound of at most $3k^{-\frac{d}{2\Delta}}/e$. We find 18 of them, one for each of the 3 functions with either x, y or y/x fixed to one of the two corresponding boundary values. This allows us to reduce to the univariate maximisation problems detailed below. To simplify the expressions we will let $k-2=r$, $ek^{\frac{d}{2\Delta}-1}=s$ and $k^{1-\frac{d}{2\Delta}}/e=t$.

	f_1	f_2	f_3
$x=s$	$p_1(y)=\frac{ry(1+s)}{(s+ry+1)^2}$	$p_2(y)=\frac{rsy+2s}{(s+ry+1)^2}$	$p_3(y)=\frac{2s+ry}{(s+ry+1)^2}$
$x=t$	$p_4(y)=\frac{ry(1+t)}{(t+ry+1)^2}$	$p_5(y)=\frac{rty+2t}{(t+ry+1)^2}$	$p_6(y)=\frac{2t+ry}{(t+ry+1)^2}$
$y=s$	$p_7(x)=\frac{rs(1+x)}{(x+rs+1)^2}$	$p_8(x)=\frac{rxs+2x}{(x+rs+1)^2}$	$p_9(x)=\frac{2x+rs}{(x+rs+1)^2}$
$y=t$	$p_{10}(x)=\frac{rt(1+x)}{(x+rt+1)^2}$	$p_{11}(x)=\frac{rxt+2x}{(x+rt+1)^2}$	$p_{12}(x)=\frac{2x+rt}{(x+rt+1)^2}$
$y/x=s$	$p_{13}(x)=\frac{rs(1+x^{-1})}{(x^{-1}+rs+1)^2}$	$p_{14}(x)=\frac{rs+2x^{-1}}{(x^{-1}+rs+1)^2}$	$p_{15}(x)=\frac{2x^{-1}+rx^{-1}s}{(x^{-1}+rs+1)^2}$
$y/x=t$	$p_{16}(x)=\frac{rt(1+x^{-1})}{(x^{-1}+rt+1)^2}$	$p_{17}(x)=\frac{rt+2x^{-1}}{(x^{-1}+rt+1)^2}$	$p_{18}(x)=\frac{2x^{-1}+rx^{-1}t}{(x^{-1}+rt+1)^2}$

To begin the maximisation, first observe that under the map $x \mapsto x^{-1}$, each of the functions $p_j(x)$ is the same as some function $p_l(x)$ for some $13 \le j \le 18$ and $7 \le l \le 12$. Furthermore, $y=s$ yields the bounds $s \le x \le 1$ and $y/x = s$ gives $1 \le x \le t$. Similarly we may compare $y = t$ and $y/x = t$. Thus the ranges for x are identical after inverting x. Hence we may ignore p_{13} through p_{18} leaving us with 12 problems.

Next, consider p_{10}, p_{11} and p_{12}, each of which can be bounded above by

$$\frac{2rtx}{(x+rt+1)^2} \leq \frac{2rtx}{r^2t^2} \leq \frac{2}{r}$$

where the final inequality follows as $x \leq t$.

Similarly, we can bound p_4, p_5 and p_6. As it must be the case that $y \geq 1$, the numerator of each is bounded above by $2try$. Thus an upper bound for all three is $2t/ry$. Furthermore, $r \geq 2k/3$ so we are left with an upper bound of $3k^{-\frac{d}{2\Delta}}/e$.

The remaining problems are similar. The numerators may all be bounded above by $rs(1+x) \leq 2rs$ (or for p_1, p_2 and p_3 by $2ry$.) The denominators are all bounded from below by r^2s^2 and r^2y^2 respectively. Thus all six of these are upper bounded by $2/rs$ which is at most $3k^{-\frac{d}{2\Delta}}/e$.

Hence an upper bound on all of the problems p_1 through p_{18} is $3k^{-\frac{d}{2\Delta}}/e$ as claimed.

B.2 Small k

When k is small, then the parameter $c = \log(k) - 1$ is also very small. In fact we do not obtain a better constant than for the Ising model until $k \geq 21$. However it is possible to do better, we can choose different values for α and c which work better in these cases. In this section we will show how to derive the vales in table 8.1.

First, we note that we may do the the analysis in an identical way until we find ourselves with the maximisation problems f_1, f_2 and f_3. Now we maximise these more carefully than in section B.1. First, for f_1 we apply AM-GM to the denominator to deduce that $f_1(x, y) \leq \frac{1}{4}$ for any x, y. This allows us to take any $c < 4$ and as k is small this is all we need and so we may ignore this constraint. This leaves us to maximise f_2 and f_3. A similar argument to the one in the proof of Lemma 8.5 allows us to deduce that the maxima are on the boundary of R_d and hence we need only consider the boundary of $\widetilde{R_d}$.

Now, we proceed as in section B.1 with different choices of s and t where this time we will take $t = e^{d/\Delta\alpha+(1-d/\Delta)c}$ and $s = t^{-1}$. We start with 12 maximisation problems which we reduce to 8 by symmetry as before. Furthermore, $f_2 > f_3$ if and only if $x > 1$ which allows us to half the number of problems left to consider leaving us with 4 problems. That is, we are left with p_3, p_5, p_9 and p_{11}. All of these are of the form $f(x) = (ax+b)(x+d)^{-2}$ which has a maximum at $x = d - 2b/a$. See the following table for the maximisation of these 4 functions.

Function	a	b	c	x^*	$f(x^*)$
p_3	$\frac{1}{k-2}$	$\frac{2s}{(k-2)^2}$	$\frac{s+1}{k-2}$	$\frac{1-3s}{k-2}$	$\frac{1}{4(1-s)}$
p_5	$\frac{t}{k-2}$	$\frac{2t}{(k-2)^2}$	$\frac{t+1}{k-2}$	$\frac{t-3}{k-2} \leq 1$	$\frac{kt}{(t+k-1)^2}$
p_9	2	$(k-2)s$	$1+(k-2)s$	1	$\frac{1}{(2+(k-2)s)}$
p_{11}	$2+(k-2)t$	0	$1+(k-2)t$	$1+(k-2)t > t$	$\frac{(k-2)t^2+2t}{((k-1)t+1)^2}$

Note that in the cases of p_5 and p_{11} the maximum value x_* is outside the domain which we are maximising over and thus we maximise at the endpoints of the domain instead.

Now, recall that the maximum values obtained above must also satisfy (8.5). Also, when $s = e^{-\alpha}$ it must be the case that $(2+(k-2)e^{-\alpha})^{-1} < 1$ (from p_9). Combining these after rearrangement yields the inequity

$$\frac{ce^c}{e^c+k-1} \leq \alpha \leq \log\left(\frac{k-2}{c-2}\right) \tag{B.1}$$

We may solve this inequality computationally for c, and deduce that there is a choice of α, c which satisfies (B.1) provided that $c \leq c_k$ for some c_k which can be found in the following table. The corresponding value of α_k is also provided. We give both c_k and α_k rounded to three decimal places.

k	3	4	5	6	7	8	9	10	11	12
α_k	1.767	1.803	1.849	1.896	1.944	1.990	2.034	2.076	2.116	2.154
c_k	2.171	2.330	2.472	2.600	2.716	2.820	2.916	3.003	3.084	3.160

Now, we check that these are indeed the maximum values. To do this, we first note that we have $p_3 \leq 1/4$ and applying AM-GM to the denominator of the maximum for p_5 similarly yields a result which is smaller than the values from p_9. Finally, for p_{11}, the denominator is at least $(k-1)(k-2)t^2 + 2t(k-1)$. Thus, after cancellations we are left with $p_{11} \leq 1/(k-1)$ which suffices for $k \geq 4$. For $k = 3$ we can easily check that $(t^2 + 2t)(2t+1)^{-2}$ is maximised when $t = 1$ and hence is certainly at most $1/3 < 1/2.17$.

Recall when computing the maximum of p_9, we took s as large as possible where one would expect that we should do the opposite to maximise p_9. We now justify this choice. So recall that we must ensure $d\theta p_9(x) \leq \Delta\theta/c$. Furthermore, s may be considered as a function of d and as such is equal to $\exp(-d/\Delta\alpha - (1 - d/\Delta)c)$. Thus we must ensure that

$$g(d) = \frac{dc/\Delta}{2 + (k-2)s} \leq 1.$$

Writing λ for d/Δ gives the function with domain $[0, 1]$

$$G(\lambda) = \frac{\lambda c}{2 + (k-2)e^{-\lambda\alpha - (1-\lambda)c}} \leq 1.$$

Differentiating this with respect to λ, we see that either $c - \alpha < 1$ and G is increasing on $[0, 1]$ or there is a maximum with $\lambda > 1$ which is not inside the domain. Thus, we maximise G at one of its boundary points and it is easy to see that $\lambda = 1$ is the maximum point rather than $\lambda = 0$ where $G(\lambda) = 0$.

LIST OF REFERENCES

[1] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probability Theory and Related Fields*, 152(3-4):367–406, 2012.

[2] M. Albert, A. Frieze, and B. Reed. Multicoloured Hamilton cycles. *The Electric Journal of Combinatorics*, 2(1):R10, 1995.

[3] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *The Annals of Probability*, pages 812–854, 1997.

[4] N. Alon. Combinatorial nullstellensatz. *Combinatorics, Probability and Computing*, 8(1-2):7–29, 1999.

[5] N. Alon and E. Fischer. 2-factors in dense graphs. *Discrete Mathematics*, 152(1):13 – 23, 1996.

[6] N. Alon and R. Yuster. H-factors in dense graphs. *Journal of combinatorial theory, Series B*, 66(2):269–282, 1996.

[7] P.D. Alvarez, F. Canfora, S.A. Reyes, and S. Riquelme. Potts model on recursive lattices: some new exact results. *The European Physical Journal B*, 85(3):99, 2012.

[8] A. Barvinok. *Combinatorics and complexity of partition functions*, volume 276. Springer, 2016.

[9] F. Bencs, E. Davies, V. Patel, and G. Regts. On zero-free regions for the anti-ferromagnetic Potts model on bounded-degree graphs. *arXiv preprint arXiv: 1812.07532*, 2018.

[10] B. Bollobás. The evolution of random graphs. *Transactions of the American Mathematical Society*, 286(1):257–274, 1984.

[11] B. Bollobás and S. E. Eldridge. Packings of graphs and applications to computational complexity. *Journal of Combinatorial Theory, Series B*, 25(2):105–124, 1978.

[12] C. Borgs, J. Chayes, T. Helmuth, W. Perkins, and P. Tetali. Efficient sampling and counting algorithms for the Potts model on $\mathbb{Z}^d$ at all temperatures. *arXiv preprint arXiv: 1909.09298*, 2019.

[13] J. Böttcher. Large-scale structures in random graphs. *Surveys in Combinatorics*, 440:87–140, 2017.

[14] J. Böttcher, Y. Kohayakawa, and A. Procacci. Properly coloured copies and rainbow copies of large graphs with small maximum degree. *Random Structures & Algorithms*, 40(4):425–436, 2012.

[15] J. Böttcher, M. Schacht, and A. Taraz. Proof of the bandwidth conjecture of Bollobás and Komlós. *Mathematische Annalen*, 343(1):175–205, 2009.

[16] R.L. Brooks. On colouring the nodes of a network. *Mathematical Proceedings of the Cambridge Philosophical Society*, 37(2):194–197, 1941.

[17] R. A. Brualdi and H. J. Ryser. *Combinatorial matrix theory*, volume 39. Cambridge University Press, 1991.

[18] P. Cano, G. Perarnau, and O. Serra. Rainbow spanning subgraphs in bounded edge–colourings of graphs with large minimum degree. *Electronic Notes in Discrete Mathematics*, 61:199–205, 2017.

[19] P. A. Catlin. Subgraphs of graphs, I. *Discrete Mathematics*, 10(2):225–233, 1974.

[20] S.C. Chang, J.L. Jacobsen, J. Salas, and R. Shrock. Exact Potts model partition functions for strips of the triangular lattice. *Journal of statistical physics*, 114(3-4):763–823, 2004.

[21] S.C. Chang, J. Salas, and R. Shrock. Exact Potts model partition functions for strips of the square lattice. *Journal of Statistical Physics*, 107(5-6):1207–1253, 2002.

[22] S.C. Chang and R. Shrock. Exact Potts model partition functions on strips of the honeycomb lattice. *Physica A: Statistical Mechanics and its Applications*, 296(1-2):183–233, 2001.

[23] V. Chvátal. On Hamilton's ideals. *Journal of Combinatorial Theory, Series B*, 12(2):163–168, 1972.

[24] D. Clemens, J. Ehrenmüller, and Y. Person. A Dirac-type theorem for Hamilton Berge cycles in random hypergraphs. *Electronic Notes in Discrete Mathematics*, 54:181–186, 2016.

[25] M. Coulson. The critical window in random digraphs. *arXiv preprint arXiv: 1905.00624*, 2019.

[26] M. Coulson, E. Davies, A. Kolla, V. Patel, and G. Regts. Statistical physics approaches to Unique Games. *arXiv preprint arXiv: 1911.01504*, 2019.

[27] M. Coulson, P. Keevash, G. Perarnau, and L. Yepremyan. Rainbow factors in hypergraphs. *Journal of Combinatorial Theory, Series A*, 172:105184, 2020.

[28] M. Coulson and G. Perarnau. A rainbow Dirac's theorem. *arXiv preprint arXiv: 1809.06392*, 2018.

[29] M. Coulson and G. Perarnau. Rainbow matchings in Dirac bipartite graphs. *Random Structures & Algorithms*, 55(2):271–289, 2019.

[30] B. Csaba. On the Bollobás–Eldridge conjecture for bipartite graphs. *Combinatorics, Probability and Computing*, 16(5):661–691, 2007.

[31] B. Csaba, D. Kühn, A. Lo, D. Osthus, and A. Treglown. *Proof of the 1-factorization and Hamilton decomposition conjectures*, volume 244. American Mathematical Society, 2016.

[32] L. DeBiasio, D. Kühn, T. Molla, D. Osthus, and A. Taylor. Arbitrary orientations of Hamilton cycles in digraphs. *SIAM Journal on Discrete Mathematics*, 29(3):1553–1584, 2015.

[33] G.A. Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, 3(1):69–81, 1952.

[34] A. Dudek and M. Ferrara. Extensions of results on rainbow Hamilton cycles in uniform hypergraphs. *Graphs and Combinatorics*, 31(3):577–583, 2015.

[35] A. Dudek, A. Frieze, and A. Ruciński. Rainbow Hamilton cycles in uniform hypergraphs. *The Electronic Journal of Combinatorics*, 19(1):46, 2012.

[36] P. Erdős. A problem on independent r-tuples. *Ann. Univ. Sci. Budapest*, 8:93–95, 1965.

[37] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 5(1):17–60, 1960.

[38] P. Erdős and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. In *Colloquia mathematica societatis Janos Bolyai: Infinite and finite sets*, volume 10, pages 609–627, 1975.

[39] P. Erdős and J. Spencer. Lopsided Lovász local lemma and Latin transversals. *Discrete Applied Mathematics*, 30(2-3):151–154, 1991.

[40] A. B. Evans. Latin squares without orthogonal mates. *Designs, Codes and Cryptography*, 40(1):121–130, 2006.

[41] R. Fernández and A. Procacci. Regions without complex zeros for chromatic polynomials on graphs with bounded degree. *Combinatorics, Probability and Computing*, 17(2):225–238, 2008.

[42] J. L. Fouquet and A. P. Wojda. Mutual placement of bipartite graphs. *Discrete Mathematics*, 121(1):85 – 92, 1993.

[43] A. Galanis, D. Štefankovič, E. Vigoda, and L. Yang. Ferromagnetic Potts model: Refined #BIS-hardness and related results. *SIAM Journal on Computing*, 45(6):2004–2065, 2016.

[44] S. Glock and F. Joos. A rainbow blow-up lemma. *arXiv preprint arXiv: 1802.07700*, 2018.

[45] L.A. Goldberg and M. Jerrum. Approximating the partition function of the ferromagnetic Potts model. *Journal of the ACM (JACM)*, 59(5):25, 2012.

[46] C. Goldschmidt and R. Stephenson. The scaling limit of a critical random directed graph. *arXiv preprint arXiv: 1905.05397*, 2019.

[47] R. Graham and N. Sloane. On additive bases and harmonious graphs. *SIAM Journal on Algebraic Discrete Methods*, 1(4):382–404, 1980.

[48] G. Grimmett and D. Stirzaker. *Probability and random processes*. Oxford University Press, 2001.

[49] G. Hahn. Un jeu de colouration. In *Actes du Colloque de Cerisy*, volume 12, pages 18–18, 1980.

[50] G. Hahn and C. Thomassen. Path and cycle sub-Ramsey numbers and an edge-colouring conjecture. *Discrete Mathematics*, 62(1):29–33, 1986.

[51] A. Hajnal and E. Szemerédi. Proof of a conjecture of P. Erdős. *Combinatorial theory and its applications*, 2:601–623, 1970.

[52] P. Hall. On representatives of subsets. *Journal of the London Mathematical Society*, 1(1):26–30, 1935.

[53] P. Hatami and P. W. Shor. A lower bound for the length of a partial transversal in a Latin square. *J. Combin. Theory Ser. A*, 115(7):1103–1113, 2008.

[54] B. Jackson, A. Procacci, and A.D. Sokal. Complex zero-free regions at large $|q|$ for multivariate Tutte polynomials (alias Potts-model partition functions) with general complex edge weights. *Journal of Combinatorial Theory, Series B*, 103(1):21–45, 2013.

[55] F. Jaeger, D.L. Vertigan, and D.J.A. Welsh. On the computational complexity of the Jones and Tutte polynomials. *Mathematical Proceedings of the Cambridge Philosophical Society*, 108(1):35–53, 1990.

[56] S. Janson. New versions of Suen's correlation inequality. *Random Structures and Algorithms*, 13(3-4):467–483, 1998.

[57] S. Janson, D.E. Knuth, T. Łuczak, and B. Pittel. The birth of the giant component. *Random Structures & Algorithms*, 4(3):233–358, 1993.

[58] S. Janson, T. Luczak, and A. Rucinski. *Random graphs*, volume 45. John Wiley & Sons, 2011.

[59] N. Kamčev, B. Sudakov, and J. Volec. Bounded colorings of multipartite graphs and hypergraphs. *European Journal of Combinatorics*, 66:235–249, 2017.

[60] R.M. Karp. The transitive closure of a random digraph. *Random Structure & Algorithms*, 1(1):73–93, 1990.

[61] P. Katerinis. Minimum degree of a graph and the existence of k-factors. *Proceedings of the Indian Academy of Sciences - Mathematical Sciences*, 94(2):123–127, 1985.

[62] H. Kaul, A. Kostochka, and G. Yu. On a graph packing conjecture by Bollobás, Eldridge and Catlin. *Combinatorica*, 28(4):469–485, 2008.

[63] P. Keevash. Hypergraph Turán problems. *Surveys in combinatorics*, 392:83–140, 2011.

[64] P. Keevash, D. Mubayi, B. Sudakov, and J. Verstraëte. Rainbow Turán problems. *Combinatorics, Probability and Computing*, 16(1):109–126, 2007.

[65] J. Kim, D. Kühn, A. Kupavskii, and D. Osthus. Rainbow structures in locally bounded colorings of graphs. *Random Structures & Algorithms*, 2018.

[66] D.J. Kleitman and K.J. Winston. The asymptotic number of lattices. *Combinatorical Mathematics, Optimal Designs and their Applications (J. Srivastava, ed.), Ann. Discrete Math*, 6:243–249, 1980.

[67] D.J. Kleitman and K.J. Winston. On the number of graphs without 4-cycles. *Discrete Mathematics*, 41(2):167–172, 1982.

[68] J. Komlós, G. Sárközy, and E. Szemerédi. Blow-up lemma. *Combinatorica*, 17(1):109–123, 1997.

[69] J. Komlós, G. Sárközy, and E. Szemerédi. Proof of the Seymour conjecture for large graphs. *Annals of Combinatorics*, 2(1):43–60, 1998.

[70] J. Komlós, G. Sárközy, and E. Szemerédi. Proof of the Alon–Yuster conjecture. *Discrete Mathematics*, 235(1-3):255–269, 2001.

[71] M. Krivelevich, C. Lee, and B. Sudakov. Robust Hamiltonicity of Dirac graphs. *Transactions of the American Mathematical Society*, 366(6):3095–3130, 2014.

[72] M. Krivelevich, C. Lee, and B. Sudakov. Compatible Hamilton cycles in Dirac graphs. *Combinatorica*, 37(4):697–732, 2017.

[73] D. Kühn, J. Lapinskas, and D. Osthus. Optimal packings of Hamilton cycles in graphs of high minimum degree. *Combinatorics, Probability and Computing*, 22(3):394–416, 2013.

[74] D. Kühn, A. Lo, D. Osthus, and K. Staden. The robust component structure of dense regular graphs and applications. *Proceedings of the London Mathematical Society*, 110(1):19–56, 2014.

[75] D. Kühn and D. Osthus. Critical chromatic number and the complexity of perfect packings in graphs. In *Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm*, pages 851–859. Society for Industrial and Applied Mathematics, 2006.

[76] D. Kühn and D. Osthus. Embedding large subgraphs into dense graphs. *Surveys in Combinatorics*, 365:137–167, 2009.

[77] D. Kühn and D. Osthus. The minimum degree threshold for perfect graph packings. *Combinatorica*, 29(1):65–107, 2009.

[78] D. Kühn and D. Osthus. Hamilton cycles in graphs and hypergraphs: an extremal perspective. *Proceedings of the International Congress of Mathematicians 2014*, 4:381–406, 2014.

[79] D. Kühn, D. Osthus, and A. Treglown. Hamiltonian degree sequences in digraphs. *Journal of Combinatorial Theory, Series B*, 100(4):367–380, 2010.

[80] C. Lee and B. Sudakov. Dirac's theorem for random graphs. *Random Structures & Algorithms*, 41(3):293–305, 2012.

[81] J. Liu, A. Sinclair, and P. Srivastava. Fisher Zeros and Correlation Decay in the Ising Model. In *10th Innovations in Theoretical Computer Science Conference (ITCS 2019)*, volume 124 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 55:1–55:8, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[82] L. Lu and L. Székely. Using Lovász local lemma in the space of random injections. *the electronic journal of combinatorics*, 14(1):R63, 2007.

[83] T. Łuczak. Component behavior near the critical point of the random graph process. *Random Structures & Algorithms*, 1(3):287–310, 1990.

[84] T. Łuczak. The phase transition in the evolution of random digraphs. *Journal of Graph Theory*, 14(2):217–223, 1990.

[85] T. Łuczak, B. Pittel, and J.C. Wierman. The structure of a random graph at the point of the phase transition. *Transactions of the American Mathematical Society*, 341(2):721–748, 1994.

[86] T. Łuczak and T. Seierstad. The critical behavior of random digraphs. *Random Structures & Algorithms*, 35(3):271–293, 2009.

[87] M. Maamoun and H. Meyniel. On a problem of G. Hahn about coloured Hamiltonian paths in K_{2t}. *Discrete Mathematics*, 51(2):213–214, 1984.

[88] A. Martin-Löf. Symmetric sampling procedures, general epidemic processes and their theshold limit theorems. *Journal of Applied Probability*, 23(2):265–282, 1986.

[89] C. McDiarmid. Clutter percolation and random graphs. In *Combinatorial Optimization II*, pages 17–25. Springer, 1980.

[90] M. Molloy and B. Reed. *Graph colouring and the probabilistic method*, volume 23. Springer Science & Business Media, 2013.

[91] R. Montgomery, A. Pokrovskiy, and B. Sudakov. Embedding rainbow trees with applications to graph labelling and decomposition. *arXiv preprint arXiv: 1803.03316*, 2018.

[92] R. Montgomery, A. Pokrovskiy, and B. Sudakov. Decompositions into spanning rainbow structures. *Proceedings of the London Mathematical Society*, 119(4):899–959, 2019.

[93] J. Moon and L. Moser. On Hamiltonian bipartite graphs. *Israel Journal of Mathematics*, 1(3):163–165, 1963.

[94] A. Nachmias and Y. Peres. The critical random graph, with martingales. *Israel Journal of Mathematics*, 176(1):29–41, 2010.

[95] I. Palásti. On the strong connectedness of directed random graphs. *Studia Sci. Math. Hungar.*, 1:205–214, 1966.

[96] V. Patel and G. Regts. Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM Journal on Computing*, 46(6):1893–1919, 2017.

[97] X. Pérez-Giménez and N. Wormald. Asymptotic enumeration of strongly connected digraphs by vertices and edges. *Random Structures & Algorithms*, 43(1):80–114, 2013.

[98] B. Pittel. On the largest component of the random graph at a nearcritical stage. *Journal of Combinatorial Theory, Series B*, 82(2):237–269, 2001.

[99] B. Pittel. Counting strongly-connected, moderately sparse directed graphs. *Random Structures & Algorithms*, 43(1):49–79, 2013.

[100] A. Pokrovskiy and B. Sudakov. A counterexample to Stein's equi-n-square conjecture. *Proceedings of the AMS*, 147:2281–2287, 2019.

[101] G. Ringel. Problem 25. In *Theory of graphs and its applications*, page 162, 1963.

[102] V. Rödl and A. Ruciński. Dirac-type questions for hypergraphs–a survey (or more problems for Endre to solve). In *An irregular mind*, pages 561–590. Springer, 2010.

[103] H. J. Ryser. Neuere probleme der kombinatorik. *Vortrageber Kombinatorik, Oberwolfach*, 1967.

[104] N. Sauer and J. Spencer. Edge disjoint placement of graphs. *Journal of Combinatorial Theory, Series B*, 25(3):295–302, 1978.

[105] A.D. Scott and G.B. Sorkin. Solving sparse random instances of max cut and max 2-CSP in linear expected time. *Combinatorics, Probability and Computing*, 15(1-2):281–315, 2006.

[106] R. Shrock. Exact Potts/Tutte polynomials for polygon chain graphs. *Journal of Physics A: Mathematical and Theoretical*, 44(14):145002, 2011.

[107] A.D. Sokal. Bounds on the complex zeros of (di) chromatic polynomials and Potts-model partition functions. *Combinatorics, Probability and Computing*, 10(1):41–77, 2001.

[108] J. Spencer. Asymptotic lower bounds for Ramsey functions. *Discrete Mathematics*, 20:69–76, 1977.

[109] S. K. Stein. Transversals of Latin squares and their generalizations. *Pacific J. Math.*, 59(2):567–575, 1975.

[110] B. Sudakov. Robustness of graph properties. *Surveys in combinatorics*, 440:372, 2017.

[111] B. Sudakov and J. Volec. Properly colored and rainbow copies of graphs with few cherries. *Journal of Combinatorial Theory, Series B*, 122:391–416, 2017.

[112] B. Sudakov and V. Vu. Local resilience of graphs. *Random Structures & Algorithms*, 33(4):409–433, 2008.

[113] E. Szemerédi. Regular partitions of graphs. 1975.

[114] I.S. Tyurin. Refinement of the upper bounds of the constants in Lyapunov's theorem. *Russian Mathematical Surveys*, 65(3):586–588, 2010.

[115] S.S. Venkatesh. *The theory of probability: Explorations and applications*. Cambridge University Press, 2013.

[116] I. M. Wanless and B. S. Webb. The existence of Latin squares without orthogonal mates. *Designs, Codes and Cryptography*, 40(1):131–135, 2006.

[117] E.M. Wright. Formulae for the number of sparsely-edged strong labelled digraphs. *The Quarterly Journal of Mathematics*, 28(3):363–367, 1977.

[118] C.N. Yang and T.D. Lee. Statistical theory of equations of state and phase transitions. I. Theory of condensation. *Physical Review*, 87(3):404, 1952.

[119] R. Yuster. Rainbow H-factors. *The electronic journal of combinatorics*, 13(1):13, 2006.

[120] Y. Zhao. Recent advances on Dirac-type problems for hypergraphs. In *Recent trends in combinatorics*, pages 145–165. Springer, 2016.