Continuous Optimisation in Extremal Combinatorics

Matthew Jenssen

A thesis submitted for the degree of
Doctor of Philosophy

Department of Mathematics

London School of Economics
and Political Science

26/07/2017

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

Abstract

In this thesis we explore instances in which tools from continuous optimisation can be used to solve problems in extremal graph and hypergraph theory.

We begin by introducing a generalised notion of hypergraph Lagrangian and use tools from the theory of nonlinear optimisation to explore some of its properties. As an application we find the Turán density of a small family of hypergraphs.

We determine the exact k-colour Ramsey number of an odd cycle on n vertices when n is large. This resolves a conjecture of Bondy and Erdős for large n. The first step of our proof is to use the regularity method to relate this problem in Ramsey theory to one in nonlinear optimisation. We establish a correspondence between extremal constructions in the Ramsey setting and optimal points in the continuous setting. We thereby uncover a correspondence between extremal constructions and perfect matchings in the k-dimensional hypercube. This allows us to prove a stability type result around these extremal constructions.

We consider two models from statistical physics, the hard-core model and the monomer-dimer model. Using tools from linear programming we give tight upper bounds on the logarithmic derivative of the independence and matching polynomials of a d-regular graph. For independent sets, this is a strengthening of a sequence of results of Kahn, Galvin and Tetali, and Zhao that a disjoint union of $K_{d,d}$'s maximises the independence polynomial and total number of independent sets among all d-regular graphs on the same number of vertices. For matchings, the result implies that disjoint unions of $K_{d,d}$'s also maximise the matching polynomial and total number of matchings. Moreover we prove the Asymptotic Upper Matching Conjecture of Friedland, Krop, Lundow, and Markström.

Through our study of the hard-core model, we also prove lower bounds on the average size and the number of independent sets in a triangle-free graph of maximum degree d. As a consequence we obtain a new proof of Shearer's celebrated upper bound on the Ramsey number $R(3, k)$.

Acknowledgements

I would like to thank my supervisors Julia Böttcher and Jozef Skokan for providing me with support, encouragement and fascinating mathematical problems throughout my time as a PhD student.

I would like to thank the staff and faculty of the LSE Mathematics department for providing a welcoming and wonderful work environment. In particular I would like to thank Rebecca Lumb and Kate Barker for their endless help and motivational chocolates.

I would like to thank my PhD colleagues, Ewan Davies and Barnaby Roberts, for the constant source of mathematical discussion and coffee time puzzles. I would also like to thank Will Perkins for inspiring me with his energetic and creative approach to mathematics and for being a wonderful collaborator.

I am grateful for the funding provided by my LSE PhD Studentship.

I would like to thank my parents, whose unwavering support and faith in me made it possible for me to pursue my passions. Without them, this would not have been possible.

Finally I would like to thank my grandfather Osmund Dorenfeldt Jenssen who introduced me to the wonders of mathematics. This thesis is dedicated to him.

Contents

1

Introduction

Extremal graph theory concerns itself with problems of the following type:

> **Question A**: Given a real-valued graph parameter P and a class of graphs $\mathcal{C}$, how large (or small) can P be for an element of $\mathcal{C}$?

One of the earliest results in extremal graph theory is Mantel's theorem [67] from 1907, which considers the case where $\mathcal{C}$ is the class of all triangle-free graphs on n vertices and P is the number of edges of a graph. Indeed, Mantel's theorem asserts that any triangle-free graph on n vertices must have at most $\lfloor n^2/4 \rfloor$ edges. If we can answer Question A, we may want to go further and determine precisely which elements of $\mathcal{C}$ optimise P. In the example of Mantel's theorem, one can show that the complete, balanced, bipartite graph on n vertices, is the *unique* triangle-free graph on n vertices with $\lfloor n^2/4 \rfloor$ edges. We call the graphs in $\mathcal{C}$ which optimise P *extremal* graphs. Once we have determined these extremal graphs, a curious phenomenon often occurs. Often one can show that elements of $\mathcal{C}$ which almost optimise P, must be *close* (in some combinatorial sense) to one of our extremal elements of $\mathcal{C}$. We call this phenomenon *combinatorial stability.*

Extremal graph theory can be viewed as 'discrete optimisation' where a natural continuous analogue might be a question of the following form:

> **Question B**: Given a function $f : \mathbb{R}^n \to \mathbb{R}$ and a subset $S \subseteq \mathbb{R}^n$, how large (or small) can $f(x)$ be for an element $x \in S$?

We will call a question of this type a question in *continuous optimisation.* If we can answer Question B, again we may want to know precisely which elements of S optimise f. We call elements of S which optimise f *optimal points.* Similarly to the discrete case, one can often show that elements of S that almost optimise f must be close (in Euclidean distance say) to a genuine optimal point of S. We refer to this phenomenon as *analytic stability.*

Questions of the form of Question B date back at least as far as 300 BC when Euclid considered the minimal distance between a point and a line, and proved that a square has the greatest area among all rectangles with a given total length of edges. With the invention of the Calculus in the 17^{th}

century by Newton and Leibniz, many powerful new techniques to attack problems in continuous optimisation emerged. Euler and Lagrange were early pioneers in the general theory of continuous optimisation and by now there are many sophisticated tools to deal with questions of type B.

In this thesis, we explore instances in which problems in extremal graph (and hypergraph) theory can be related to problems in continuous optimisation. We investigate a range of questions of type A along with its extensions and take advantage of the parallels that we saw with questions of type B. The aim is to attack problems in graph theory by importing powerful analytic tools which are not usually at one's disposal in a discrete setting.

We outline how the rest of this chapter is arranged. In Section 1.1 we collect some common notation and terminology that we make use of throughout this thesis.

In Section 1.2 we introduce the notion of a hypergraph (a generalisation of graphs) and discuss the extremal theory of hypergraphs. In particular we will discuss a relatively recent and powerful tool known as *hypergraph Lagrangians*. This will give the relevant background and preparation for Chapter 2.

In Section 1.3 we discuss graph Ramsey theory in order to provide the relevant background and preparation for Chapter 3.

In Section 1.4 we introduce notions from the intersection of graph theory and statistical physics in preparation for Chapters 4 and 5.

Finally, in Section 1.5 we introduce the tools that we will need to borrow from the theory of continuous optimisation.

1.1 Notation and Terminology

Most of the notation introduced here is standard but we include it for completeness.

For a natural number k, we let $[k]$ denote the set $\{1, \ldots, k\}$. For a set S, we let $\binom{S}{k}$ denote the set of all unordered k-tuples of distinct elements of S.

A *graph* is a pair $G = (V, E)$ where $V = V(G)$ is some fixed set and $E = E(G) \subseteq \binom{V}{2}$. We call $V(G)$ the set of *vertices* of G and we refer to $E(G)$ as the set of *edges*. All graphs in this thesis can be assumed to be *finite* meaning that $V(G)$ is a finite set. For a finite graph G we let $v(G) = |V(G)|$ and $e(G) = |E(G)|$. If there is no ambiguity, we may slightly abuse notation by writing $v \in G$ and $\{x, y\} \in G$ in lieu of $v \in V(G)$ and $\{x, y\} \in E(G)$ respectively. For $u, v \in V(G)$ we may write $u \sim v$ to indicate that $\{u, v\} \in E(G)$. We may also denote an edge $\{u, v\}$ simply by uv and refer to u and v as the *endpoints* of the edge uv. For two edges $e, f \in E(G)$ we may write $e \sim f$ to indicate that e and f are incident i.e. they share an endpoint.

For disjoint subsets $A, B \subseteq V(G)$, we denote by $G[A, B]$ the graph with vertex set $A \cup B$ and edge set $\{\{a, b\} \in E : a \in A, b \in B\}$, and we let $e_G(A, B)$ denote the size of this set. In the case where $A = \{v\}$ a singleton, we write $G[v, B]$ instead of $G[\{v\}, B]$.

For $v \in V(G)$, we let $N_G(v) = \{u \in V(G) : u \sim v\}$ denote the *neighbourhood* of v in G and let $d_G(v) = |N_G(v)|$, the *degree* of v. We let $\delta(G) = \min_{v \in G} d_G(v)$ and $\Delta(G) = \max_{v \in G} d_G(v)$, the *minimum* and *maximum degree* of G respectively.

Subscripts in the above notation may be suppressed if they are clear from the context.

For two graphs F, G, we say that F is a *subgraph* of G if there exists an injective function $f : V(F) \to V(G)$ such that $f(e) \in E(G)$ for all $e \in E(F)$ (for a set $S \subseteq V(F)$, $f(S)$ denotes the set $\{f(v) : v \in S\}$). For a subset $U \subseteq V(G)$, we let $G[U]$ denote the graph with vertex set U and edge set $E(G) \cap \binom{U}{2}$ and call $G[U]$ the subgraph of G *induced* by U.

Throughout this thesis we omit the use of floor and ceiling symbols where they are not crucial. We will use standard asymptotic notation with a subscript indicating that the implied constant may depend on that subscript. All other notation will be explained in the relevant sections.

1.2 Extremal Hypergraph Theory

A hypergraph is a generalisation of the notion of a graph. An *r-uniform hypergraph* (or r-graph for short) is a pair $H = (V, E)$ where $V = V(H)$ is some fixed set and $E = E(H) \subseteq \binom{V}{r}$. We call $V(H)$ the set of *vertices* of H and we refer to $E(H)$ as the set of *(hyper)edges*. All hypergraphs we consider will be *finite*, meaning that they have finite vertex set. Note that a graph is simply a 2-uniform hypergraph. Many extremal problems in graph theory can be generalised to the setting of hypergraphs and often the problems become significantly more difficult. This is certainly the case for the extremal problem that we focus on in this section, the study of Turán numbers of hypergraphs (to be defined shortly). For two hypergraphs F and H, we say that F is a *subgraph* of H if there exists an injective function $f : V(F) \to V(H)$ such that $f(e) \in E(H)$ for all $e \in E(F)$. If F is not a subgraph of H we say that H is *F-free*.

A natural question asked by Turán, first for graphs and then for hypergraphs, is the following: given a fixed r-graph F, what is maximum number of edges attained by an F-free r-graph on n vertices? We denote this number by $\mathrm{ex}(n, F)$ and call it the *Turán number* of F. We refer to F-free r-graphs on n vertices with $\mathrm{ex}(n, F)$ edges as *extremal*. Turán famously determined the extremal graphs (and hence also the Turán number) in the case where $F = K_t$, the *complete graph* on t vertices, that is the graph on t vertices with all edges present. The result is known as Turán's Theorem and it extends Mantel's Theorem which we introduced at the start of this chapter. Before stating Turán's Theorem we introduce some notation and definitions. We make these definitions in the more general context of hypergraphs. We say an r-graph H is *ℓ-partite* if there exists a partition $V(H) = V_1 \cup \ldots \cup V_\ell$ such that

$$E(H) \subseteq \left\{ e \in \binom{V(H)}{r} : |e \cap V_i| \leq 1 \text{ for } i = 1, \ldots, \ell \right\}.$$

We call H *complete* ℓ-partite if we have equality in the above inclusion and we call H *balanced* ℓ-partite if $||V_i| - |V_j|| < 1$ for all $i, j \in [\ell]$.

Let $Tr(n, \ell)$ denote the complete, balanced ℓ-partite graph on n vertices.

These graphs are called *Turán graphs*, and we write $tr(n,\ell)$ for the number of edges in $Tr(n,\ell)$.

Theorem 1.1 (Turán [84]). *Let $\ell \geq 1$. Then for $n \geq \ell+1$,*

$$ex(n, K_{\ell+1}) = tr(n,\ell)\,.$$

Moreover $Tr(n,\ell)$ is the unique extremal $K_{\ell+1}$-free graph on n vertices.

For most graphs and hypergraphs H, the exact determination of $\mathrm{ex}(n,H)$ is extremely difficult. Instead we might ask for the asymptotic behaviour of $\mathrm{ex}(n,H)$. By using a simple averaging argument Katona, Nemetz, and Simonovits [57] showed that for any fixed r-graph H the following limit always exists

$$\pi(H) := \lim_{n\to\infty} \frac{\mathrm{ex}(n,H)}{\binom{n}{r}}.$$

We call $\pi(H)$ the *Turán density* of H and determining these densities is one of the central problems in extremal hypergraph theory. Rather remarkably the Turán density is known for any graph G and it depends only on its *chromatic number*. The chromatic number $\chi(G)$ of G is the least k such that there exists a function $f: V(G) \to [k]$ with $f(u) \neq f(v)$ for all edges $uv \in G$ (we call such a function a *proper vertex colouring* of G with k colours). Erdős and Simonovits [29] discovered the following corollary of a theorem of Erdős and Stone [30].

Theorem 1.2 (Erdős-Stone-Simonovits). *If G is a graph with at least one edge, then*

$$\pi(G) = 1 - \frac{1}{\chi(G)-1}.$$

Given this result, it may seem surprising that as soon as $r \geq 3$, the Turán density is unknown for most r-graphs. Let $K_n^{(r)}$ denote the complete r-graph on n vertices (that is the r-graph on n vertices with all possible hyperedges present). A natural first question would be to ask for the Turán density of $K_4^{(3)}$, however this remains a major open problem. Turán showed that $\pi(K_4^{(3)}) \geq 5/9$ and conjectured that this is the right value. Erdős famously offered \$500 for a verification of this conjecture. Currently the best known upper bound is due to Razborov [73] who applied the recently developed method of flag algebras to show that $\pi(K_4^{(3)}) \leq 0.561666$. For an excellent survey on progress on hypergraph Turán problems up until 2011 see Keevash [58].

A huge variety of tools and techniques have now been developed for the purpose of determining the Turán densities of graphs. We will focus on just one of them, known as the method of *hypergraph Lagrangians*.

1.2.1 Hypergraph Lagrangians

First let us introduce the notion of *homomorphism* between hypergraphs. Given two r-graphs F, H we say that $f : V(F) \to V(H)$ is a homomorphism if $f(e) \in E(H)$ for all $e \in E(F)$. We say that H contains a *homomorphic copy* of F in this case. Note that f is not necessarily injective and so F may or may not be a subgraph of H. We will say that H is *F-hom-free* if H contains no homomorphic copy of F. In analogy to $\mathrm{ex}(n, F)$ we may define $\mathrm{ex}_{\mathrm{hom}}(n, F)$ to be the maximum number of edges attained by an F-hom-free r-graph on n vertices. Although $\mathrm{ex}(n, F)$, $\mathrm{ex}_{\mathrm{hom}}(n, F)$ can be different, it will be useful to recall (see e.g. [58]) that they are asymptotically equal i.e. for any r-graph F

$$\lim_{n\to\infty} \frac{\mathrm{ex}_{\mathrm{hom}}(n, H)}{\binom{n}{r}} = \lim_{n\to\infty} \frac{\mathrm{ex}(n, H)}{\binom{n}{r}}. \tag{1.1}$$

Let H be an r-graph on vertex set $[n]$ and let $t = (t_1, \ldots, t_n)$ be a vector of positive integers. The *t-blowup* of H, denoted by $H(t)$, is the r-graph with vertex set $V_1 \cup \ldots \cup V_n$, where each V_i is a set of size t_i, and edge set

$$\big\{\{v_1, \ldots, v_r\} : v_i \in V_{x_i} \text{ for } i = 1, 2, \ldots, r \text{ where } \{x_1, \ldots, x_r\} \in E(H)\big\}.$$

In other words, we replace each vertex i with a set of size t_i and replace each edge with the corresponding complete r-partite r-graph. A useful observation is that for r-graphs F and H, H is F-hom-free if and only if $H(t)$ is F-free for all t.

Note that for an r-graph H on n vertices and a vector $t = (t_1, \ldots, t_n)$ of positive integers we have

$$e(H(t)) = \sum_{e\in E(H)} \prod_{i\in e} t_i\,.$$

The right hand side is a homogeneous polynomial of degree r in the variables $t_1, \dots, t_n$ and we denote this polynomial by $p_H(t)$.

Suppose now that H is an n-vertex F-hom-free r-graph so that $H(t)$ is F-free on $|t| := \sum_{i=1}^n t_i$ vertices for any vector of positive integers $t = (t_1, \dots, t_n)$. It follows that

$$\pi(F) \geq \limsup_{|t|\to\infty} \frac{p_H(t)}{\binom{|t|}{r}} = \limsup_{|t|\to\infty} r! p_H(t/|t|)\,, \tag{1.2}$$

where for the last equality we used that p_H is homogeneous of degree r. In view of (1.2) it is natural to ask for the maximum of p_H over the set

$$S = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1 \text{ and } x_i \geq 0 \text{ for all } i \right\}.$$

Note that since p_H is continuous and S is compact, the maximum is indeed attained. S is often referred to as the *standard simplex* in $\mathbb{R}^n$. Since p_H is continuous and any point in S can be arbitrarily approximated by vectors of the form $t/|t|$ where $t \in \mathbb{N}^n$, it follows from (1.2) that

$$\pi(F) \geq r! \sup_{x\in S} p_H(x). \tag{1.3}$$

We call $\sup_{x\in S} p_H(x)$ the *Lagrangian* of H and denote it by $\lambda(H)$. The following simple lower bound on $\lambda(H)$ is often handy.

$$r!\lambda(H) \geq r! p_H(\tfrac{1}{n}, \dots, \tfrac{1}{n}) = r! \frac{e(H)}{n^r} = \frac{e(H)}{\binom{n}{r}} - O(\tfrac{1}{n})\,. \tag{1.4}$$

Note that by (1.1), $\pi(F)$ is the limit supremum of $\binom{n}{r}^{-1} e(H)$ over all F-hom-free H and so by (1.3) and (1.4), $\pi(F)$ is the limit supremum of $r!\lambda(H)$ over all F-hom-free H as well. We can in fact say a bit more, but first we require one more observation regarding the Lagrangian.

We say that an r-graph H *covers pairs* if every pair of vertices in H is contained in some edge of H. Suppose that H is an r-graph on vertex set $[n]$ that doesn't cover pairs i.e. there exists $i, j \in [n]$ such that no edge of H contains both i and j. It follows that $p_H(x) = Ax_i + Bx_j + C$ where A, B and C do not depend on x_i and x_j. Suppose now that $x \in S$ and suppose without loss of generality that $A \geq B$. Let $x' \in \mathbb{R}^n$ be the vector with coordinates $x'_i = x_i + x_j$, $x'_j = 0$ and $x'_k = x_k$ for $k \neq i, j$, then clearly $x' \in S$ and $p_H(x') \geq p_H(x)$. It follows that there is a subgraph H' of H such that H' covers pairs and $\lambda(H') = \lambda(H)$. For an r-graph F, let $\mathcal{C}(F)$ be the set of all F-hom-free r-graphs that cover pairs. It follows that

$$\pi(F) = \sup_{H \in \mathcal{C}(F)} r!\lambda(H). \tag{1.5}$$

To illustrate the use of this formalism let us show how it can be used to prove some of the classical results we have already seen in this chapter. First note that the only graphs that cover pairs are complete graphs and so the Lagrangian of any graph is equal to the Lagrangian of the largest complete graph it contains. It is not difficult to show (see e.g. Chapter 2) that the symmetry of the complete graph K_t means that $p_{K_t}(x)$ is maximised over S when all coordinates are equal and so $\lambda(K_t) = \frac{1}{2}(1 - \frac{1}{t})$. Suppose now that G is a K_t-free graph so that $\lambda(G) = \lambda(K_s)$ for some $s < t$. If G has n vertices it follows that

$$\frac{2}{n^2} e(G) = 2p_G(\tfrac{1}{n}, \ldots, \tfrac{1}{n}) \leq 2\lambda(G) = 1 - \frac{1}{s} \leq 1 - \frac{1}{t-1}\,.$$

This is Turán's Theorem in the case where $t - 1$ divides n. This argument is due to Motzkin and Straus [68] and it is one of the earliest appearances of the method of Lagrangians. Note also that the Erdős-Stone-Simonovits Theorem is an immediate corollary of (1.5) since a complete graph K_t is F-hom-free if and only if $t < \chi(F)$.

The development of the theory of Lagrangians for hypergraphs is attributed to Sidorenko [78] and Frankl and Füredi [41]. In Chapter 2 we present a generalised notion of hypergraph Lagrangian and use tools from continuous optimisation to exploit some of its properties. As an application we calculate the Turán densities of a new small class of hypergraphs.

1.3 Ramsey Theory

Ramsey theory is a central area of research in combinatorics whose philosophy can be summarised by the following epithet: 'large structures, no matter how disordered, must contain ordered substructures.' In order to make this more formal, let us introduce some common language in terms of which almost all results in Ramsey theory are phrased. Given a set X and positive integer k, a *k-colouring* of X is any map $\chi : X \to [k]$ where $[k] = \{1, \ldots, k\}$ is the set of colours. Given such a colouring χ, we call a subset $Y \subseteq X$ *monochromatic*, if it is contained in the set $\chi^{-1}(\{i\})$ for some $i \in [k]$ (i.e. all elements of Y are given the same colour).

Although Ramsey theory owes its name to the seminal paper of Frank Ramsey [72] from 1930, arguably the first result in Ramsey theory was proved by Hilbert in 1892. Given natural numbers $a, d_1, \ldots, d_m$, define

$$H(a; d_1, \ldots, d_m) = \left\{ a + \sum_{i \in I} d_i \;\middle|\; I \subseteq [m] \right\}.$$

We call such a set a *Hilbert cube* of dimension m. Hilbert [53] proved that, given positive integers k, m, there exists a number $H = H(k, m)$ such that any k-colouring of $[H]$ contains a monochromatic Hilbert cube of dimension m. Another early and seminal result in Ramsey theory is due to van der Waerden [85], who showed in 1927 that any colouring of the natural numbers with a finite number of colours, contains monochromatic arithmetic progressions of arbitrary length.

In this thesis we will be concerned with graph Ramsey theory which concerns itself with studying the *Ramsey numbers* of graphs, defined as follows. Given graphs $G_1, G_2, \ldots, G_k$, the *Ramsey number* $R(G_1, \ldots, G_k)$ is the least integer N such that any k-colouring of the edges of the complete graph K_N on N vertices contains a monochromatic copy of G_i in the i-th colour for some i, $1 \leq i \leq k$. In the case where $G_1, \ldots, G_k$ are all isomorphic to the graph G, we call $R(G_1, \ldots, G_k)$ the *k-colour Ramsey number* of G and denote it by $R_k(G)$. In the case of two colours we write $R(G)$ in place of $R_2(G)$. We call $R_k(G)$ a *diagonal* Ramsey number and we refer to $R(G_1, \ldots, G_k)$ as *off-diagonal* if G_i is not isomorphic to G_j for some pair i, j. In Ramsey's celebrated paper [72], he showed that Ramsey numbers always exist i.e. for any collection of finite graphs $G_1, \ldots, G_k$, $R(G_1, \ldots, G_k)$ is finite.

The oldest and most famous examples of Ramsey numbers are those involving complete graphs. For positive integers $t_1, \ldots, t_k$, we write $R(t_1, \ldots, t_k)$ as a shorthand for the Ramsey number $R(K_{t_1}, \ldots, K_{t_k})$, we use $R_k(t)$ as a shorthand for the case where all the t_i are equal to t and we let $R(t)$ denote $R_2(t)$. The systematic study of such Ramsey numbers began with a paper of Erdős and Szekeres [31] (1935) who established the bound

$$R(s,t) \leq \binom{s+t-1}{s-1}, \tag{1.6}$$

for all $s, t \geq 2$. The exact value of $R(s,t)$ is only known in a small handful of cases (see Radziszowski [58] for an excellent survey of such exact results) and the problem of improving the known bounds on these quantities is notoriously difficult. Particular notoriety has been attached to the case where $s = t$, where not even the value of $R(5)$ is known. The bound (1.6) shows that $R(t) = O(4^t/\sqrt{t})$ and despite considerable effort over the past 80 years no improvement has been made to the base of the exponent in this bound. The current best upper bound is due to Conlon [19] who gave the first superpolynomial improvement showing that there exists a positive constant c such that

$$R(t) \leq t^{-c\log t/\log\log t}4^t .$$

It wasn't until a decade after the discovery of the bound (1.6), that a significant lower bound on the quantity $R(t)$ was established. In 1947 Erdős [27] pioneered the use of the probabilistic method, producing one of the most well-known proofs in all of combinatorics, in order to establish the bound

$$R(t) \geq (1 - o(1))\frac{t}{\sqrt{2}e}\sqrt{2}^t . \tag{1.7}$$

In a similar manner to (1.6), this bound has stubbornly resisted improvement. In fact, since 1947 the only significant improvement is due to Spencer [82] who improved (1.7) by a factor of 2 using the Lovász local lemma.

Extensive research has also been dedicated to the study of the Ramsey numbers $R(s,t)$ in the case where s is fixed and t is growing. In this case (1.6) shows that $R(s,t) \leq t^{s-1}$. In 1980, Ajtai, Komlós and Szemerédi [1] improved this by a polylogarithmic factor showing that for s fixed

$$R(s,t) = O\left(\frac{t^{s-1}}{\log^{s-2} t}\right) .$$

The proof is an induction on s and the main effort is in establishing the base case where $s = 3$. The authors in fact show that a triangle-free graph G

(i.e. a graph which does not have the complete graph K_3 as a subgraph) on n vertices with average degree d has an independent set (that is a collection of vertices with no edges between them) of size $0.01\frac{\log d}{d}n$. This implies the bound $R(3,t) \leq 100\frac{\log^2 t}{t}$. In 1983 Shearer [77] gave an elegant and short proof of the improved bound

$$R(3,t) \leq (1+o(1))\frac{\log^2 t}{t}. \tag{1.8}$$

Perhaps surprisingly, in 1995 Kim [60] showed that this is the correct order of magnitude for $R(3,t)$ i.e. he showed that $R(3,t) = \Omega\left(\frac{\log^2 t}{t}\right)$. Kim's proof was a pioneering use of what has become known as the semi-random method. Recently Bohman [6] gave an alternative proof of Kim's result by analysing a stochastic graph process called the triangle-free process. The process starts with a graph with no edges and step by step adds edges uniformly at random from the collection of edges whose addition would not create a triangle. The process stops when the addition of any new edge would create a triangle. More recently still, Bohman and Keevash [7] and independently Fiz Pontiveros, Griffiths and Morris [40] analysed the running time of the triangle-free process more carefully to show that

$$R(3,t) \geq \left(\frac{1}{4}+o(1)\right)\frac{\log^2 t}{t}. \tag{1.9}$$

It is already rather remarkable that we know $R(3,t)$ to such accuracy, however it is a major open problem to reduce the gap between (1.8) and (1.9) further still. In Chapter 4 we give a new proof of Shearer's bound (1.8) and suggest new strategies for improving this bound.

Returning to the diagonal case, the difficulty in improving the bounds for $R(t)$ motivated the study of Ramsey numbers of graphs with a 'simpler' structure, where the problem may be more tractable. In stark contrast to the exponential behaviour of $R(t)$, Chvatál, Rödl, Szemerédi and Trotter [16] showed that bounded degree graphs have *linear* Ramsey number. Formally they showed that for all d there exists a constant c_d such that if G is a graph on n vertices with maximum degree d then

$$R(G) \leq c_d n\,.$$

The authors remark that their argument extends easily to the k-colour case i.e. for all k, d there exists a constant $c_{k,d}$ such that $R_k(G) \leq c_{k,d}n$ for any graph G on n vertices with maximum degree d. Recently Lee [63] greatly generalised this result showing that the same conclusion holds if 'bounded degree' is replaced by 'bounded degeneracy' (again the result is stated for two colours, but extends to k-colours). Lee's result settled the famous Burr-Erdős conjecture from 1973 [14].

In Chapter 3 we will discuss the Ramsey theory of 'sparse' graphs in more detail and see examples where the Ramsey numbers of graphs can even be determined exactly. In particular we will focus on a conjecture of Bondy and Erdős from 1973 which asserts that for all k and odd $n > 3$

$$R_k(C_n) = 2^{k-1}(n-1) + 1 \,.$$

Here C_n denotes the cycle on n vertices (see Chapter 3 for a formal definition). In particular we prove that the conjecture holds for any fixed k and n sufficiently large. The first step of the proof is to relate the problem in Ramsey theory to the problem of maximising a linear function over a region in 3^k-dimensional Euclidean space bounded by quadratic constraints.

1.4 Statistical Physics Models on Graphs

Many important graph polynomials, such as the independence polynomial, matching polynomial and chromatic polynomial, can be viewed in terms of partition functions of statistical physics models on graphs.

In this section we introduce some examples of these models and present a general approach for bounding their partition functions. This will help prepare us for Chapters 4 and 5. To begin with we introduce the hard-core model from statistical physics. Recall that an independent set in a graph is simply a collection of vertices with no edges between them. For a graph G, we let $\mathcal{I}(G)$ denote the set of all independent sets in G. The *hard-core* model on a graph G at *fugacity* λ is a random independent set I drawn according to the distribution

$$\mathbb{P}_{G,\lambda}[I] = \frac{\lambda^{|I|}}{P_G(\lambda)}, \quad \text{where} \quad P_G(\lambda) = \sum_{I \in \mathcal{I}(G)} \lambda^{|I|}.$$

Here $|I|$ denotes the number of vertices in I. Note that we indicate the graph and value of the fugacity in the subscript of probabilities and expectations but drop it from the notation when they are clear from the context. The function $P_G(\lambda)$ is the partition function of the hard-core model, or in the language of graph theory, the independence polynomial. Note that evaluating $P_G(1)$ counts the total number of independent sets of G which we will denote by $i(G)$.

The hard-core model is relevant in statistical physics as a simple model of a gas consisting of particles of non-negligible size. In this context the host graph is usually a lattice, the vertices of which may or may not be occupied by a gas particle. The constraint that the gas particles form an independent set in the lattice can be interpreted as the condition that these particles are non-overlapping.

For a positive integer d, we say that a graph G is *d-regular* if every vertex in G has degree d. Let $K_{d,d}$ denote the complete bipartite graph with d vertices in each part. A classical result of Kahn [55] states that if G is a d-regular bipartite graph then

$$i(G) \le i(K_{d,d})^{v(G)/2d}.$$

In particular, if $2d$ divides n then the bipartite d-regular graph on n vertices with the most independent sets is a disjoint union of $K_{d,d}$'s on n vertices. Kahn's argument makes elegant use of the information theoretic notion of entropy to study the hard-core model. Galvin and Tetali [46] gave a broad generalisation of Kahn's result to counting homomorphisms from a d-regular, bipartite G to any graph H. The case where H is formed of two connected vertices, one with a self-loop, is that of counting independent sets. Via a modification of H and a limiting argument, they proved that if G is a d-regular bipartite graph and $\lambda > 0$ then we in fact have

$$P_G(\lambda) \le P_{K_{d,d}}(\lambda)^{v(G)/2d}. \tag{1.10}$$

Zhao [86] then discovered a way to remove the bipartite restriction showing that (1.10) in fact holds for any d-regular graph G. This resolved a conjecture of Alon [3] whose original motivation was a problem in combinatorial group theory.

In Chapter 4 we will present a new approach to bounding $P_G(\lambda)$ for regular graphs. Instead of dealing with the partition function directly we study a related parameter known as the *occupancy fraction*. The occupancy fraction of the hard-core model on a graph G is simply the expected fraction of vertices of G in a random independent set drawn according to the model i.e.

$$\alpha_G(\lambda) := \frac{1}{v(G)} \sum_{I \in \mathcal{I}(G)} |I| \cdot \mathbb{P}[I] \,.$$

In Chapter 4, we study the occupancy fraction from the perspective of an extremal combinatorialist, asking which graphs maximise or minimise the occupancy fraction under certain constraints on the graph class. We then deduce extremal information on the partition function of a graph via the interpretation of the occupancy fraction as the scaled logarithmic derivative of the partition function:

$$\alpha_G(\lambda) = \frac{1}{v(G)} \frac{\sum_{I \in \mathcal{I}(G)} |I| \lambda^{|I|}}{P_G(\lambda)} = \frac{1}{v(G)} \frac{\lambda P'_G(\lambda)}{P_G(\lambda)} = \frac{\lambda}{v(G)} \cdot (\log P_G(\lambda))' \,.$$

In Chapter 4 we will prove that for any $\lambda > 0$, the d-regular graph which maximises the occupancy fraction is $K_{d,d}$. This strengthens the results of Kahn, Galvin-Tetali and Zhao mentioned above. Via the same method, we provide a lower bound for the occupancy fraction of a bounded degree graph with no triangles. As a result we obtain new lower bounds for the average size and the number of independent sets in triangle-free graphs. As a further corollary we obtain a new proof of (1.8), Shearer's upper bound on the Ramsey number $R(3, t)$.

Unlike the partition function, the occupancy fraction is the expected value of a physical observable of our model. This probabilistic interpretation is crucial for our proof method. Our proof method has proven sufficiently general that it can be used to analyse a variety of statistical physics models. In Chapter 5 we study the *monomer-dimer* model. A matching in a graph

G is simply a collection of vertex disjoint edges and we let $\mathcal{M}(G)$ denote the set of all matchings in G. In the language of statistical physics, the edges of a matching in G are referred to as 'dimers' and the unmatched vertices are the 'monomers'. The monomer-dimer model is a probability distribution over matchings M in a graph G, where

$$\mathbb{P}_{G,\lambda}[M] = \frac{\lambda^{|M|}}{M_G(\lambda)}, \quad \text{and} \quad M_G(\lambda) = \sum_{M \in \mathcal{M}(G)} \lambda^{|M|}.$$

Here $|M|$ denotes the number of edges in the matching M. In graph theory M_G is known as the *matching polynomial* of G. The monomer-dimer model dates back to 1935 when Roberts [74] considered the problem of adsorption of oxygen and hydrogen on a tungsten surface.

We remark that the monomer-dimer model is simply the hard-core model run on the *line graph* of G (the line graph of G is the graph on vertex set $E(G)$ where two vertices e, f are adjacent if and only if they are incident as edges in G).

As in the hard-core model, we can define the *edge* occupancy fraction, the expected fraction of edges occupied by a random matching:

$$\alpha_G^M(\lambda) := \frac{1}{e(G)} \sum_{M \in \mathcal{M}(G)} |M| \cdot \mathbb{P}[M] = \frac{\lambda}{e(G)} \left(\log M_G(\lambda)\right)'.$$

In analogy to our result on the hard-core model we prove that for any $\lambda > 0$, the d-regular graph which maximises the edge occupancy fraction is $K_{d,d}$. We get as a corollary that for any d-regular graph G we have

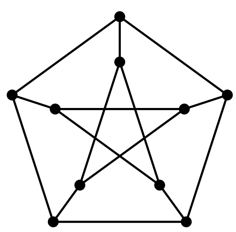
$$M_G(\lambda) \le M_{K_{d,d}}(\lambda)^{v(G)/2d}. \tag{1.11}$$

This resolves a conjecture of Galvin (Conjecture 7.1 in [45]). In the case where $2d$ divides $v(G)$ it is natural to conjecture that (1.11) holds coefficient by coefficient, that is, over all d-regular graphs on n vertices, a disjoint union of $K_{d,d}$'s maximises the number of matchings of any given size. This is known as the Upper Matching Conjecture of Friedland, Krop and Markström [44]. In Chapter 5 we prove new upper bounds on the number of matchings of a given size in regular graphs. Although we do not resolve the Upper Matching Conjecture, our bounds are sufficient to prove a weakened version known as the Asymptotic Upper Matching Conjecture [43].

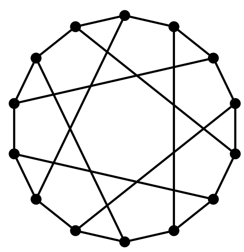
The proof method that unifies Chapters 4 and 5 can be summarised as follows. We choose a random vertex or edge from our graph and a random sample from our model (i.e. an independent set or a matching). We then look at the way in which our random sample intersects the neighbourhood of our vertex or edge (we call this a *local view*). Each local view occurs with some probability and we can place consistency constraints on these probabilities that must hold for all regular graphs. We then relax the extremal problem on graphs to an optimisation problem on probability distributions on local views and pose the relaxation as a linear program (see the next section). We then use techniques from linear programming to solve this optimisation problem and show that the optimal distribution matches the distribution obtained from our conjectured extremal graph.

To end this section we mention a couple of further applications of this method that do not appear in this thesis. By applying this method to the Potts model (a generalisation of the famous Ising model), Davies, Perkins, Roberts and the current author [23] showed that over all 3-regular graphs on n vertices, a disjoint union of $K_{3,3}$'s maximises the number of proper q-colourings.

So far all of the extremal graphs have been complete bipartite. Perkins and Perarnau [70] showed that by forbidding certain local structures one can obtain a richer class of extremal graphs. In particular they show that for $\lambda > 0$, over all cubic graphs G of girth at least 5, the occupancy fraction $\alpha_G(\lambda)$ is maximised by the Heawood graph (see below). They also show that for $0 < \lambda \leq 1$, over all triangle-free cubic graphs, the occupancy fraction is minimised by the Peterson graph.



Peterson graph



Heawood graph

1.5 Tools from Continuous Optimisation

In this section we collect some standard tools and results from the theory of continuous optimisation that we use throughout this thesis. For us, a continuous optimisation problem will be a problem of the form

$$\begin{aligned} &\text{maximise} && f(x) \\ &\text{subject to} && g_i(x) \leq 0, \quad i = 1, \ldots, m. \end{aligned} \tag{1.12}$$

Here $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is a vector, and we refer to the coordinates x_i as the *decision variables*. The function $f : \mathbb{R}^n \to \mathbb{R}$ is the *objective function* and the functions $g_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$ are called the *constraint functions*. A vector x is called *feasible* if it satisfies each of the constraints $g_i(x) \leq 0$, $i = 1, \ldots, m$. A vector x^* is called *optimal* if it has the largest objective value among all feasible vectors i.e. for any z with $g_1(z) \leq 0, \ldots, g_m(z) \leq 0$, we have $f(z) \leq f(x^*)$.

In the special case where the objective function and the constraint functions are all linear, we call a problem of the form (1.12) a *linear program* (or LP for short), otherwise we call the problem *nonlinear*. In the case where the objective function and the constraint functions are all convex we call a problem of the form (1.12) a *convex optimisation problem*. There is an enormous amount of literature and deep theory in the study of continuous optimisation and much of this is dedicated to the special case of linear or convex problems. Here we only borrow some standard tools from this theory.

1.5.1 Linear Programming

Let us introduce some standard vector notation that we use throughout this thesis. Let $\mathbb{R}^{m \times n}$ denote the space of all $m \times n$ matrices with real entries. For $A \in \mathbb{R}^{m \times n}$ we let A^T denote the transpose of A. For two vectors $z, w \in \mathbb{R}^n$ we write $z \leq w$ if the inequality holds componentwise. If $f(x)$ is a scalar function of $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, then we let $\nabla f(x)$ be the gradient of f at x i.e. the vector in $\mathbb{R}^n$ whose ith coordinate is $\frac{\partial}{\partial x_i} f(x)$.

Now, let $b, c \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m\times n}$ and consider the linear program

$$\begin{aligned} &\text{maximise} \quad c^T x \\ &\text{subject to} \quad Ax \leq b,\ x \geq 0. \end{aligned} \tag{1.13}$$

The *dual* linear program to (1.13) is

$$\begin{aligned} &\text{minimise} \quad b^T y \\ &\text{subject to} \quad A^T y \geq c,\ y \geq 0. \end{aligned} \tag{1.14}$$

We refer to (1.13) as the *primal* linear program.

We will make use of the following two standard tools from the theory of linear programming in Chapters 4 and 5 (see [12, p. 244] for a detailed account).

Theorem 1.3 (Weak LP Duality)**.** *Suppose x and y are feasible for the primal and dual linear programs* (1.13) *and* (1.14)*, then $c^T x \leq b^T y$. In particular, if $c^T x = b^T y$, then x and y must be optimal for the primal and the dual.*

Theorem 1.3 is simply the observation that if you multiply the inequality $Ax \leq b$ on the left by y^T and you multiply the inequality $A^T y \geq c$ on the left by x^T, then you obtain $c^T x \leq y^T Ax \leq b^T y$. Despite its simplicity, Theorem 1.3 is extremely useful. In particular it shows that if you manage to find feasible solutions for a linear program and its dual whose objective values match, then they must both be optimal solutions.

Theorem 1.4 (Complementary Slackness)**.** *Suppose x and y are feasible for the primal and dual linear programs* (1.13) *and* (1.14)*. Then x and y are optimal if and only if $(b - Ax)^T y = 0$ and $(A^T y - c)^T x = 0$.*

The proof of Theorem 1.4 is not complicated although we omit it here. In conjunction, Theorems 1.3 and 1.4 furnish us with the following strategy for solving linear programs: suppose that we believe x^* is an optimal solution to the linear program (1.13). The support of x^* (i.e. the coordinates i for which $x^*_i \neq 0$) then tells us, by Theorem 1.4, which of the dual constraints

$A^T y \geq c$ should hold with equality. The hope is that by solving these equality constraints, one finds a dual feasible solution y^* whose objective value matches that of the original guess x^*. If this is the case, then Theorem 1.3 tells us that x^* and y^* are both indeed optimal.

We remark that the linear programs we will come across in this thesis have the following form:

$$\begin{aligned} &\text{maximise} \quad c^T x \\ &\text{subject to} \quad Ax = b,\ x \geq 0\,. \end{aligned}$$

Note that we have equality constraints here rather than inequality constraints. Of course this can be manipulated into the same form as (1.13) (sometimes referred to as symmetric form) by replacing the equality constraint $Ax = b$ with the pair of inequality constraints $Ax \leq b$ and $-Ax \leq -b$. The dual linear program can then be written as

$$\begin{aligned} &\text{minimise} \quad b^T y \\ &\text{subject to} \quad A^T y \geq c\,. \end{aligned}$$

Note that y is no longer constrained to be non-negative.

1.5.2 Karush-Kuhn-Tucker Conditions

In this section we introduce a very general tool from the theory of continuous optimisation known as the Karush-Kuhn-Tucker (KKT) optimality conditions. We only discuss a version of this theory that is sufficiently general to suit our needs. As before, suppose we have an optimisation problem of the form of

$$\begin{aligned} &\text{maximise} \quad f(x) \\ &\text{subject to} \quad g_i(x) \leq 0, \quad i = 1, \ldots, m\,, \end{aligned} \tag{1.15}$$

where $x \in \mathbb{R}^n$ and f and g_i are *differentiable* functions $\mathbb{R}^n \to \mathbb{R}$ for all i. For $x^* \in \mathbb{R}^n$ we will say that the *KKT conditions* hold at x^* if there exist $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ such that

(i) $\nabla f(x^*) = \sum_{i=1}^m \lambda_i \nabla g_i(x^*)$,

(ii) $\lambda_i \geq 0, \ \ i = 1, \ldots, m,$

(iii) $\lambda_i g_i(x^*) = 0, \ \ i = 1, \ldots, m.$

The KKT conditions are of particular interest since under certain 'regularity conditions', an optimal point of (1.15) must satisfy the KKT conditions (see Theorem 1.5 below). These regularity conditions are often quite general and only ask for the constraint functions to satisfy certain mild properties. These are often referred to as *constraint qualifications.* Many different types of constraint qualification appear in the literature. Here we will make use of the following two well-known constraint qualifications (see [12, p.146] for a detailed account).

- **Slater's Condition:** $f, g_1, \ldots, g_m$ are all convex and there exists $z \in \mathbb{R}^n$ such that $g_i(z) < 0$ for $i = 1, \ldots, m$.
- **Linearity Constraint Qualification (LCQ):** $g_1, \ldots, g_m$ are all affine functions.

Theorem 1.5. *If Slater's condition or LCQ holds, then any optimal point of* (1.15) *must satisfy the KKT conditions.*

We remark that when applied to the linear program (1.13), the KKT conditions are precisely the assertion that there exists a feasible solution to the dual program which satisfies the complimentary slackness conditions of Theorem 1.4.

2

A hypergraph Turán theorem via a generalised notion of hypergraph Lagrangian

2.1 Introduction

Recall that for an r-graph H, the *Turán number* $\mathrm{ex}(n, H)$ is the maximum number of edges attained by an r-graph on n vertices containing no copy of H as a subgraph and the *Turán density* of H is the limit

$$\pi(H) = \lim_{n\to\infty} \frac{\mathrm{ex}(n, H)}{\binom{n}{r}}.$$

A valuable tool in the arsenal of methods that has been developed over the years to attack hypergraph Turán problems is the hypergraph Lagrangian which we introduced in Section 1.2.1. To describe one of the early successes of the method of Lagrangians we begin with a question of Katona. In an attempt to generalise Mantel's Theorem [67], Katona [56] asked for the largest number of edges an n-vertex 3-graph can have under the constraint that there is no edge that contains the symmetric difference of two other edges. Bollobás [8] settled this question by showing that the maximum is achieved uniquely by the complete balanced 3-graph on n vertices and went on to conjecture that the same should hold for arbitrary r, not just $r = 2, 3$. With an early use of the method of Lagrangians, Sidorenko [79] settled the $r = 4$ case of this conjecture. In fact he showed that the extremal construction is the same under the weaker constraint that there are no three edges e, f, g such that $|f \cap g| = 3$ and $f \triangle g \subseteq e$ (here $\triangle$ denotes the symmetric difference operator).

Let us define a *k-avoiding* r-graph to be an r-graph H with the property that for all edges $e, f \in E(H)$ we have $|e \cap f| \neq k$. As a key lemma in Sidorenko's proof in [79], he shows that the maximum Lagrangian over all 3-avoiding 4-graphs is attained by the hypergraph formed by a single edge (the method shows that the same is true for $(r-1)$-avoiding r-graphs where $r = 2, 3$). In [42], Frankl and Füredi extend Sidorenko's method to show that for $r = 5, 6$ the maximum Lagrangian over all $(r-1)$-avoiding r-graphs is attained by the Steiner systems $S(11, 5, 4)$ and $S(12, 6, 5)$ respectively (a Steiner system, $S(n, r, q)$, is an r-graph H on n vertices in which every element of $\binom{V(H)}{q}$ is contained in exactly one hyperedge). As a result one can determine the Turán density of the *generalised triangle*, the graph on vertex set $[2r-1]$

and edges

$$\{1, 2 \ldots, r\}, \{1, 2, \ldots, r-1, r+1\}, \{r, r+1, \ldots, 2r-1\}$$

for $r = 2, 3, 4, 5$ and 6. In a similar spirit, Hefetz and Keevash [51] asked for the maximum Lagrangian attained by an intersecting r-graph (an r-graph whose edges have pairwise non-empty intersection). They proved that for $r = 3$ the maximum is attained by $K_5^{(3)}$ and as a consequence they obtain the Turán density of a related 3-graph. The authors then go on to propose a more general direction of investigation in extremal combinatorics, namely to determine the maximum Lagrangian of a hypergraph satisfying a given property. Natural properties to consider are those which restrict edge intersection sizes as in the results mentioned above.

In this chapter we prove

Theorem 2.1. *Let H be an $(r-2)$-avoiding r-graph. Then*

$$\lambda(H) \leq \lambda(K_{r+1}^{(r)}) = \frac{1}{(r+1)^{(r-1)}}$$

for $r = 3, 4, 5, 6$ and 7.

As a result we determine the Turán density of what we shall call the 'r-uniform generalised K_4' for these values of r. More precisely the generalised K_4, denoted by $\mathcal{K}_4^{(r)}$, is the r-graph on $5r-6$ vertices with the 6 edges

$$\{x_1, \ldots, x_r\}, \{y_1, y_2, x_3, \ldots, x_r\} \text{ and } \{x_i, y_j, z_{ij1}, \ldots, z_{ij(r-2)}\} \text{ for } i, j \in \{1, 2\}.$$

(In words, $\mathcal{K}_4^{(r)}$ is the graph obtained by taking two edges with intersection size $r-2$ and for each pair of vertices not in an edge, adding $(r-2)$ new vertices to form an edge with that pair).

Theorem 2.2.

$$\pi(\mathcal{K}_4^{(r)}) = \frac{r!}{(r+1)^{(r-1)}}$$

for $r = 3, 4, 5, 6$ and 7.

We note that $\mathcal{K}_4^{(2)} = K_4$, the complete graph on 4 vertices. We believe that the method of proof of the above theorems is of independent interest.

We introduce a generalised notion of hypergraph Lagrangian and use the Karush-Kuhn-Tucker conditions introduced in Section 1.5.2 to derive some of its properties.

It is tempting to conjecture that Theorem 2.1 (and therefore Theorem 2.2) holds for all r. However, the following theorem, which determines the order of the maximum Lagrangian attained by an $(r-2)$-avoiding r-graph (as a function of r), shows that this is not the case.

Theorem 2.3. *Let A_r denote the set of all $(r-2)$-avoiding r-graphs. Then*

$$\sup_{H \in A_r} \lambda(H) = \Theta\left(\frac{1}{r^4 r!}\right).$$

The layout of this chapter is as follows: in Section 2.2 we introduce the generalised hypergraph Lagrangian. In Section 2.3 we introduce some standard techniques for explicitly calculating the Lagrangian of an r-graph and use the results of Section 2.2 to prove Theorem 2.1. We do this by first bounding the generalised Lagrangian over the much simpler class of 1-avoiding 3-graphs. In Section 2.4, we show how Theorem 2.1 can be used to prove Theorem 2.2. Finally in Section 2.5 we prove Theorem 2.3 and suggest avenues of future research.

We note that after completing this chapter, the author discovered that the some of its contents (in particular the cases $r = 3, 4$ in Theorem 2.1) are implicit in a previous paper of Sidorenko [80].

2.2 The Generalised Lagrangian

In this section we introduce a generalised notion of hypergraph Lagrangian and explore some of its properties. Let H be an r-graph on $[n]$ and let w, t be positive reals. We define

$$S_{w,t} = \left\{ x \in \mathbb{R}^n : 0 \leq x_i \leq t \text{ for } 1 \leq i \leq n \text{ and } \sum_{i=1}^{n} x_i \leq w \right\}$$

and

$$\lambda_{w,t}(H) = \sup_{x \in S_{w,t}} p_H(x).$$

Note that this supremum is attained as p_H is continuous and $S_{w,t}$ is compact. This should be compared to the Lagrangian defined in Section 1.2.1 where the only difference is in the modification to the standard simplex. Note that $\lambda_{1,1}(H) = \lambda(H)$ for all r-graphs H and so we use the term *generalised Lagrangian* to refer to quantities of the form $\lambda_{w,t}(H)$.

For an r-graph $H = (V, E)$ and subset $X \subseteq V$, we let $H(X)$ denote the $(r - |X|)$-graph with vertex set $V - X$ and edge set $\{e - X : e \in E, X \subseteq e\}$ (for $x \in V$ we write $H(x)$ in place of $H(\{x\})$). Using the Karush-Kuhn-Tucker conditions (Theorem 1.5) we prove the following result which allows us to bound the generalised Lagrangian of an r-graph H in terms of a related generalised Lagrangian of the hypergraph $H(x)$ for some $x \in V(H)$. The advantage of this approach is that $H(x)$ may have a simpler structure and so its generalised Lagrangian may be more amenable to analysis.

Theorem 2.4. *Let H be an r-graph and let $w, t > 0$. Then there exists an $x \in V(H)$ and $s \leq t, w$ such that*

$$r\lambda_{w,t}(H) \leq w\lambda_{w-s,s}(H(x)).$$

Proof. Let $V(H) = [n]$ and choose $a \in S_{w,t}$ such that $p_H(a) = \lambda_{w,t}(H)$. Note that we are maximising p_H over $S_{w,t}$ which is defined by affine constraints. We may therefore apply Theorem 1.5 (with the LCQ condition) to find constants $\Lambda, \mu_i, \theta_i \geq 0$ such that for all $i \in [n]$

$$\frac{\partial p_H}{\partial x_i}(a) = \Lambda - \mu_i + \theta_i \tag{2.1}$$

and

$$a_i\mu_i = 0, \qquad a_i\theta_i = t\theta_i, \text{ and } \Lambda \sum_i a_i = \Lambda w. \tag{2.2}$$

Note that since p_H is a homogeneous polynomial of degree r we have that

$$r\lambda_{w,t}(H) = \sum_{i=1}^{n} a_i \frac{\partial p_H}{\partial x_i}(a).$$

By (2.1) and (2.2) we then have

$$r\lambda_{w,t}(H) = \sum_{i=1}^{n} a_i(\Lambda - \mu_i + \theta_i) = w\Lambda + t\sum_{i=1}^{n} \theta_i. \tag{2.3}$$

Noting that $\theta_i > 0$ only if $a_i = t$ and that $\sum_i a_i \leq w$ we see that $\theta_i > 0$ for at most $\lfloor w/t \rfloor$ values of $i \in [n]$. It follows by averaging that we can find $j \in [n]$ such that

$$\theta_j \geq \frac{t}{w} \sum_{i=1}^{n} \theta_i. \tag{2.4}$$

We consider two cases:

Case 1: $\theta_j > 0$. Without loss of generality, let $j = n$. Since $\theta_n > 0$ we have $a_n = t > 0$ and so $\mu_n = 0$. Thus, by (2.1), (2.3) and (2.4) we have

$$\frac{\partial p_H}{\partial x_n}(a) = \Lambda + \theta_n \geq \Lambda + \frac{t}{w} \sum_{i=1}^{n} \theta_i = \frac{r}{w} \lambda_{w,t}(H).$$

Case 2: $\theta_j = 0$. In this case, noting that $\theta_i \geq 0$ for all $i \in [n]$, (2.4) shows that in fact $\theta_i = 0$ for all $i \in [n]$. Without loss of generality let $a_n = \max\{a_1, \ldots, a_n\}$. Clearly we must have that $a_n > 0$ and so $\mu_n = 0$. Thus, as in Case 1 we have

$$\frac{\partial p_H}{\partial x_n}(a) = \Lambda = \frac{r}{w} \lambda_{w,t}(H).$$

Finally note that in both cases $\frac{\partial p_H}{\partial x_n}(a) = p_{H(n)}(a')$ where $a' = (a_1, \ldots, a_{n-1})$ and $a_n = \max\{a_1, \ldots, a_n\}$ so that $a' \in S_{w-s,s} \subseteq \mathbb{R}^{n-1}$ for some $s \leq t, w$. The result follows. $\square$

Note that successive applications of Theorem 2.4 allows one to repeatedly simplify the hypergraph whose Lagrangian we are trying to bound. Starting with the Lagrangian $\lambda(H) = \lambda_{1,1}(H)$ of a hypergraph H and repeatedly applying Theorem 2.4 leads to the following.

Theorem 2.5. *Let H be an r-graph, then for $1 \leq m \leq r$ there exists $X \in \binom{V(H)}{m}$ and $s \leq 1/m$ such that*

$$\frac{r!}{(r-m)!} \lambda(H) \leq \lambda_{1-ms,s}(H(X)) \prod_{i=0}^{m-1} (1 - is).$$

Proof. We proceed by induction on m. Noting that $\lambda(H) = \lambda_{1,1}(H)$, Theorem 2.4 gives an $x \in V(H)$ and an $s \leq 1$ such that $r\lambda(H) \leq \lambda_{1-s,s}(H(x))$.

This gives the base case $m = 1$. For fixed $2 \leq m < r$, suppose that there exists $X \in \binom{V(H)}{m}$ and $t \leq 1/m$ such that

$$\frac{r!}{(r-m)!}\lambda(H) \leq \lambda_{1-mt,t}\big(H(X)\big)\prod_{i=0}^{m-1}(1-it). \tag{2.5}$$

By Theorem 2.4 applied to $H(X)$ there exists an $x \in V(H(X))$ and $s \leq t, 1-mt$ (and so $s \leq 1/(m+1)$) such that

$$(r-m)\lambda_{1-mt,t}\big(H(X)\big) \leq (1-mt)\lambda_{1-mt-s,s}\big(H(X\cup\{x\})\big). \tag{2.6}$$

Since $s \leq t$ and since $\lambda_{w,s}\big(H(X\cup\{x\})\big)$ is an increasing function of w, the right hand side of (2.6) is at most $(1-ms)\lambda_{1-(m+1)s,s}\big(H(X\cup\{x\})\big)$. In view of (2.5) this completes the induction.

$\square$

In many cases we expect equality in the statement of Theorem 2.5, an example of which we see in the next section.

2.3 1-avoiding 3-graphs and a proof of Theorem 2.1.

In this section we consider the class of 1-avoiding 3-graphs and bound the generalised Lagrangian of such a hypergraph. We will then use Theorem 2.5 to bound the Lagrangian of an $(r-2)$-avoiding r-graph. We first need to introduce some tools that are useful for explicitly calculating the generalised Lagrangian of a given hypergraph. Recall that for r-graphs F and H, a *homomorphism* from F to H is a map $f : V(F) \to V(H)$ such that $f(e) \in E(H)$ for all $e \in E(F)$. We call a bijective homomorphism from H to itself an *automorphism* of H and let $Aut(H)$ denote the group of all automorphisms of H under composition.

Definition 2.6. *Given an r-graph H on vertex set $[n]$, let $\sim_H$ denote the equivalence relation on $[n]$ given by $i \sim_H j$ if and only if $Aut(H)$ contains the transposition (ij).*

The following lemma can be found as Lemma 2.8 in [51]. It will be useful to replicate the proof here and to mimic a corollary (Corollary 2.9 of [51]) of that lemma in our modified setting.

Lemma 2.7. *Let $H = ([n], E)$ be a hypergraph and let $i, j \in [n]$ be such that $i \sim_H j$. Suppose $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ with $a \geq 0$ and let $a' \in \mathbb{R}^n$ have coordinates $a'_k = a_k$ for $k \neq i, j$ and $a'_i = a'_j = (a_i + a_j)/2$. Then we have $p_H(a') \geq p_H(a)$.*

Proof. Since (ij) is an automorphism of H it is easy to see that

$$p_H(a') - p_H(a) = \sum_{\substack{e \in E \\ \{i,j\} \subseteq e}} ((a_i + a_j)^2/4 - a_i a_j) \prod_{k \in e \setminus \{i,j\}} a_k \geq 0$$

. □

Corollary 2.8. *If H is an r-graph and $w, t > 0$ then there exists $a \in S_{w,t}$ such that $p_H(a) = \lambda_{w,t}(H)$ and $a_i = a_j$ whenever $i \sim_H j$.*

Proof. Suppose $V(H) = [n]$. Let $\{P_i : i \in I\}$ be the set of equivalence classes of $\sim_H$ on $[n]$. For $x \in \mathbb{R}^n$ and $P \subseteq [n]$, let $x_P := \frac{1}{|P|} \sum_{j \in P} x_j$. Choose an $a \in R := \{x \in S_{w,t} : p_H(x) = \lambda_{w,t}(H)\}$ which minimises the sum $T(a) = \sum_{m \in I} \sum_{\ell \in P_m} |a_\ell - a_{P_m}|$ (note that T is continuous and R is compact and so we may choose such an $a \in R$). We wish to show that $T(a) = 0$. Suppose not, then we can find $m \in I$ and $i, j \in P_m$ such that $a_i < a_{P_m} < a_j$. Let a' have coordinates $a'_k = a_k$ for $k \neq i, j$ and $a'_i = a'_j = (a_i + a_j)/2$ and note that $a' \in S_{w,t}$. Since $i \sim_H j$ we have that $p_H(a') = \lambda_{w,t}(H)$ by Lemma 2.7. This contradicts the choice of a since $T(a') < T(a)$. □

Recall that we say a hypergraph H is k-avoiding if $|e \cap f| \neq k$ for all $e, f \in E(H)$. The following basic lemma gives a full characterisation of 1-avoiding 3 graphs. We go on to use this characterisation to bound the generalised Lagrangian of such a hypergraph. We let $K_4^{(3)-}$ denote the unique 3-graph on 4 vertices with 3 edges. For $k \geq 3$, we let S_k be the 3-graph with vertex set $[k]$ and with edge set $\{\{12j\} : 3 \leq j \leq k\}$. S_k is sometimes called a sunflower with $k - 2$ petals and kernel of size 2. For $k = 1, 2$ we let $S_k = ([k], \emptyset)$. For a hypergraph H and subset $X \subseteq V(H)$, we let $d(X)$ denote the number of edges of H that contain X.

Lemma 2.9. *Let H be a 1-avoiding 3-graph. Then H is a vertex disjoint union of copies of $K_4^{(3)}, K_4^{(3)-}$ and S_k where $k \geq 1$.*

Proof. We proceed by induction on N, the number of pairs $\{u, v\} \in H$ with $d(\{u, v\}) \geq 2$. If $N = 0$ then H is a matching (i.e. a disjoint union of copies of S_3's) so we're done. If $N > 0$, then select $\{u, v\} \subseteq V(H)$ with $d(\{u, v\}) = k \geq 2$. Let $\{x_1, \ldots, x_k\}$ be the set of vertices of H such that $\{u, v, x_i\}$ in an edge of H. If $k \geq 3$, the vertices $\{u, v, x_1, \ldots, x_k\}$ induce an isolated copy of S_k in H since H is 1-avoiding. Similarly, if $k = 2$, the vertices $\{u, v, x_1, x_2\}$ induce an isolated copy of $S_4, K_4^{(3)-}$ or $K_4^{(3)}$. In each case, removing this isolated subgraph of H and applying the induction hypothesis completes the proof. $\square$

We now calculate the generalised Lagrangian of some specific 1-avoiding 3-graphs and show that given $w, t > 0$, the 1-avoiding 3-graph H that maximises the quantity $\lambda_{w,t}(H)$ is a vertex disjoint union of many copies of $K_4^{(3)}$. In the following we let $mK_4^{(3)}$ denote the disjoint union of m copies of $K_4^{(3)}$.

Lemma 2.10. *Let $w, t > 0$, then*

$$\lambda_{w,t}(mK_4^{(3)}) \leq 4t^3 \left(\left\lfloor \frac{w}{4t} \right\rfloor + \left(\frac{w}{4t} - \left\lfloor \frac{w}{4t} \right\rfloor \right)^3 \right)$$

with equality if $m > w/4t$.

Proof. Let $H = mK_4^{(3)}$. If $m \leq w/4t$ then $\lambda_{w,t}(H) = 4mt^3$ as we may assign the maximum value of t to each variable in the polynomial $p_H(x)$. Since m is an integer we in fact have that $m \leq \lfloor w/4t \rfloor$ and so the inequality in the statement of the lemma is clearly satisfied.

Suppose then that $m > w/4t$. By considering each copy of $K_4^{(3)}$ in $mK_4^{(3)}$ separately, we may write

$$\lambda_{w,t}(mK_4^{(3)}) = \lambda_{w_1,t}(K_4^{(3)}) + \ldots + \lambda_{w_m,t}(K_4^{(3)})$$

for some w_i satisfying $\sum_i w_i \leq w$ and $0 \leq w_i \leq 4t$ (the total weight on each copy of $K_4^{(3)}$ is at most $4t$). By Corollary 2.8 we have $\lambda_{w_i,t}(K_4^{(3)}) = 4(w_i/4)^3 = w_i^3/16$ and so

$$\lambda_{w,t}(mK_4^{(3)}) = (w_1^3 + w_2^3 + \ldots + w_m^3)/16.$$

Pick $i \neq j$ and let $u = w_i + w_j$. Maximising the function $f(x, y) = x^3 + y^3$ subject to the constraints $x + y = u$ and $0 \leq x, y \leq 4t$ we find that one of x and y is equal to 0 or $4t$. It follows that $w_k = 0$ or $4t$ for all but at most one value of $k \in [m]$. Letting $l = \lfloor w/4t \rfloor$ we may assume wlog that $w_k = 4t$ for $k = 1, 2, \ldots l$, $w_{l+1} = w - 4tl$ and $w_k = 0$ for $k = l + 2, \ldots, m$. The result follows.

□

Lemma 2.11. *Suppose $k \geq 3$ and $w, t > 0$ then*

$$\lambda_{w,t}(S_k) \leq \begin{cases} w^3/27 & \text{if } w \leq 3t \\ t^2(w - 2t) & \text{if } w > 3t\,. \end{cases}$$

Proof. By Corollary 2.8 we have that $\lambda_{w,t}(S_k) = (k-2)x^2y$ for some $0 \leq x, y \leq t$ such that $2x + (k-2)y \leq w$. It follows that $\lambda_{w,t}(S_k) \leq x^2(w - 2x)$ which, subject to the constraints $0 \leq x \leq t$, is maximised when $x = \min\{t, w/3\}$. The result follows. □

Lemma 2.12. *Let $w, t > 0$, $m \geq w/4t$ and $k \geq 3$. Then*

$$\lambda_{w,t}(S_k) \leq \lambda_{w,t}(mK_4^{(3)})\,.$$

Proof. By Lemmas 2.10 and 2.11 we have the following: If $w \leq 3t$ then $\lambda_{w,t}(S_k) \leq w^3/27$, $\lambda_{w,t}(mK_4^{(3)}) = w^3/16$ and so we're done. Suppose then that $w > 3t$ and let $\mu = w/4t$. We then have

$$\lambda_{w,t}(S_k) \leq 4t^3\left(\mu - \frac{1}{2}\right) \leq 4t^3(\lfloor \mu \rfloor + (\mu - \lfloor \mu \rfloor)^3) = \lambda_{w,t}(mK_4^{(3)})\,.$$

The second inequality can be seen by letting $x = \mu - \lfloor \mu \rfloor$ and noting that $x - 1/2 \leq x^3$ for $x \geq 0$. □

Lemma 2.13. *Let $w, t > 0$ and let H be a 1-avoiding 3-graph. Then*

$$\lambda_{w,t}(H) \leq 4t^3\left(\left\lfloor \frac{w}{4t} \right\rfloor + \left(\frac{w}{4t} - \left\lfloor \frac{w}{4t} \right\rfloor\right)^3\right).$$

Proof. By Lemma 2.9 we may write $H = H_1 \cup \ldots \cup H_s$ a disjoint union where each H_i is isomorphic to $K_4^{(3)}, K_4^{(3)-}$ or S_k for some $k \geq 1$. It follows that

$$\lambda_{w,t}(H) \leq \lambda_{w_1,t}(H_1) + \ldots + \lambda_{w_s,t}(H_s) \tag{2.7}$$

for some $w_i \geq 0$ satisfying $\sum_i w_i \leq w$. Note that $K_4^{(3)-} \subseteq K_4^{(3)}$ and so $\lambda_{w,t}(K_4^{(3)-}) \leq \lambda_{w,t}(K_4^{(3)})$ for all $w, t > 0$. Applying this observation and Lemma 2.12 to (2.7) gives, for m_i suitably large,

$$\lambda_{w,t}(H) \leq \lambda_{w_1,t}(m_1 K_4^{(3)}) + \ldots + \lambda_{w_s,t}(m_s K_4^{(3)}) \leq \lambda_{w,t}(m K_4^{(3)})$$

where $m = \sum_i m_i$. The result now follows from Lemma 2.10. $\square$

We are now in a position to address one of the main results of this chapter.

Proof of Theorem 2.1. If $r = 3$, setting $w = t = 1$ in Lemma 2.13 tells us that $\lambda(H) = \lambda_{1,1}(H) \leq 1/16$ which is the desired bound. Assume therefore that $r \geq 4$. By Theorem 2.5 there exists $X \in \binom{V(H)}{r-3}$ and $s \leq 1/(r-3)$ such that

$$r!\lambda(H) \leq 6\lambda_{1-(r-3)s,s}(H(X)) \prod_{i=1}^{r-4}(1 - is)\,.$$

As H is an $(r-2)$-avoiding r graph, $H(X)$ is a 1-avoiding 3-graph. By Lemma 2.13

$$\lambda_{1-(r-3)s,s}(H(X)) \leq 4s^3(\lfloor \mu \rfloor + (\mu - \lfloor \mu \rfloor)^3)$$

where $\mu = (1 - (r-3)s)/4s$. Combining the above two inequalities yields

$$r!\lambda(H) \leq 24s^3(\lfloor \mu \rfloor + (\mu - \lfloor \mu \rfloor)^3) \prod_{i=1}^{r-4}(1 - is) =: f_r(s).$$

To complete the proof it suffices to show that $f_r(s)$ attains its maximum value over the interval (0,1] at $s = 1/(r+1)$ for $r = 4, 5, 6$ and 7. The proof of this is left to the Appendix (Claim A in Section A). $\square$

2.4 The Generalised K_4

In this section we prove Theorem 2.2 showing how the above results can be used to compute the Turán density of $\mathcal{K}_4^{(r)}$ for $r = 3, 4, 5, 6$ and 7. First let us recall a result from our introductory discussion of hypergraph Lagrangians in Chapter 1 (Section 1.2.1). Recall that we say that an r-graph H *covers pairs* if every pair of vertices in H is contained in some edge of H. For an r-graph F, letting $\mathcal{C}(F)$ be the set of all F-hom-free r-graphs that cover pairs, recall that

$$\pi(F) = \sup_{H \in \mathcal{C}(F)} r!\lambda(H). \tag{2.8}$$

Proof of Theorem 2.2. The lower bound $\pi(\mathcal{K}_4^{(r)}) \geq r!/(r+1)^{(r-1)}$ can be established by observing that blowups $K_{r+1}^{(r)}(m)$ of the complete r-graph on $(r+1)$ vertices are $\mathcal{K}_4^{(r)}$-free. Indeed consider a pair of edges in $K_{r+1}^{(r)}(m)$ intersecting in exactly $r-2$ vertices. This pair of edges spans $r+2$ vertices and so two of those vertices must lie in the same vertex class of $K_{r+1}^{(r)}(m)$. But then this pair of vertices cannot be contained in an edge $K_{r+1}^{(r)}(m)$ whereas $\mathcal{K}_4^{(r)}$ covers pairs.

Suppose now that H is a $\mathcal{K}_4^{(r)}$-free r-graph that covers pairs. By (2.8) it suffices to prove that $\lambda(H) \leq 1/(r+1)^{(r-1)}$. Suppose for contradiction that $\lambda(H) > 1/(r+1)^{(r-1)}$ then by Theorem 2.1 we can find $e, f \in E(H)$ such that $|e \cap f| = r-2$. Let $e = \{x_1, \ldots, x_r\}$ and $f = \{y_1, y_2, x_3, \ldots, x_r\}$ where $x_i \neq y_j$ for $i, j \in \{1, 2\}$. H covers pairs so for $i, j \in \{1, 2\}$ we may find $z_{ij1}, \ldots, z_{ij(r-2)} \in V(H)$ such that $\{x_i, y_j, z_{ij1}, \ldots, z_{ij(r-2)}\} \in E(H)$. It follows that H contains a homomorphic copy of $\mathcal{K}_4^{(r)}$, a contradiction. $\square$

2.5 $(r-2)$-avoiding r-graphs for large r.

In this section we prove Theorem 2.3 determining the order of the maximum Lagrangian attained by an $(r-2)$-avoiding r-graph (as a function of r). Theorem 2.3 shows that for large r, the complete graph $K_{r+1}^{(r)}$ is exponentially far from being optimal.

First we need to define the notion of a Sidon set.

Definition 2.14. *Let n be a positive integer. We say that $A \subseteq \mathbb{Z}_n$ is a* Sidon set *if all the ordered sums $x + y$, where $x, y \in A$ are distinct.*

A simple counting argument shows that a Sidon subset $A \subseteq \mathbb{Z}_n$ can have cardinality at most $\sqrt{2n}$ (indeed we have $\binom{|A|}{2} + |A|$ ordered sums and they must all be distinct so that $\binom{|A|}{2} + |A| \leq n$). A construction of Singer [81] shows that there exist Sidon subsets of $\mathbb{Z}_n$ with the same order of magnitude as this upper bound:

Proposition 2.15. *There exist Sidon subsets of $\mathbb{Z}_n$ of cardinality*

$$(1 - o(1))\sqrt{n}\,.$$

Proof of Theorem 2.3. First we show that the Lagrangian of any $(r-2)$-avoiding r-graph must be $O(\frac{1}{r^4 r!})$. This follows easily from the proof of Theorem 2.1: As in that proof, if H is an $(r-2)$-avoiding r-graph ($r \geq 4$) then there exists a $0 < s \leq 1/(r-3)$ such that

$$r!\lambda(H) \leq 24s^3(\lfloor\mu\rfloor + (\mu - \lfloor\mu\rfloor)^3)\prod_{i=1}^{r-4}(1 - is) \tag{2.9}$$

where $\mu = (1-(r-3)s)/4s$. Using the inequality $\lfloor\mu\rfloor + (\mu - \lfloor\mu\rfloor)^3 \leq \mu$ in (2.9) yields

$$r!\lambda(H) \leq 6s^2\prod_{i=1}^{r-3}(1 - is) \leq 6s^2 \exp\left\{-s\binom{r-2}{2}\right\} \tag{2.10}$$

where for the last inequality we use that $1 - x \leq e^{-x}$ for $x \in \mathbb{R}$. Considering the right hand side of (2.10) as a function of $s \geq 0$ we see that it is maximised when $s = 2/\binom{r-2}{2}$ and so

$$\lambda(H) \leq \frac{24}{e^2 r!\binom{r-2}{2}^2}.$$

For the lower bound we construct an $(r-2)$-avoiding r-graph whose Lagrangian matches the upper bound up to a constant factor. Fix a positive integer n to be determined later. Let $A \subseteq \mathbb{Z}_n$ be a Sidon set and for each $k \in [n]$ define the hypergraph $H_k = (A, E_k)$ where

$$E_k = \left\{e \in \binom{A}{r} : \sum_{v \in e} v \equiv k \pmod{n}\right\}.$$

Note that since A is a Sidon set H_k is $(r-2)$-avoiding for each k. Moreover the sets $E_1, \ldots, E_n$ form a partition of $\binom{A}{r}$ and thus by the pigeonhole principle we must have

$$|E_j| \geq \binom{|A|}{r} \Big/ n$$

for some $j \in [n]$. Let $m := |A|$ and let $H := H_j$. Recalling the definition of the Lagrangian we have

$$\lambda(H) \geq p_H\left(\frac{1}{m}, \ldots, \frac{1}{m}\right) \geq \frac{1}{nm^r}\binom{m}{r} = \frac{1}{nr!}\prod_{i=1}^{r-1}\left(1 - \frac{i}{m}\right). \tag{2.11}$$

By Proposition 2.15 we may choose $m = (1+o(1))\sqrt{n}$. Since we have not yet specified n we may now do so implicitly by setting $m = (r-1)^2$. Making these substitutions into (2.11) yields

$$\begin{aligned} \lambda(H) &\geq (1+o(1))\frac{1}{r^4 r!}\prod_{i=1}^{r-1}\left(1 - \frac{i}{(r-1)^2}\right) \\ &\geq (1+o(1))\frac{1}{r^4 r!}\left(1 - \frac{1}{(r-1)}\right)^{(r-1)} \\ &= (1+o(1))\frac{1}{er^4 r!} \end{aligned}$$

where we have used the fact that $(1-1/r)^r \to e^{-1}$ as $r \to \infty$. □

We end this chapter with some suggestions of possible avenues for further research.

It would be interesting to investigate at which point the complete graph $K_{r+1}^{(r)}$ ceases to maximise the Lagrangian over $(r-2)$-avoiding r-graphs. As mentioned in Section 2.1, $K_r^{(r)}$ (i.e. a single hyperedge) maximises the Lagrangian over all $(r-1)$-avoiding r-graphs for $r = 2, 3, 4$ after which more interesting extremal structures begin to appear. Frankl and Füredi [42] show that for $r = 5, 6$ the maximum Lagrangian over all $(r-1)$-avoiding r-graphs is attained by the Steiner systems $S(11, 5, 4)$ and $S(12, 6, 5)$ respectively.

The construction based on Sidon sets in the proof of Theorem 2.3 is an example of a *partial* Steiner $(n, r, r-2)$ system (i.e. an r-graph H on n vertices in which every element of $\binom{V(H)}{r-2}$ is contained in *at most* one hyperedge). Note that being a partial Steiner $(n, r, r-2)$ system is a stronger condition than being $(r-2)$-avoiding. Theorem 2.3 suggests that for $r > 7$, a Steiner system $S(n, r, r-2)$ would be a good candidate for maximising the Lagrangian over $(r-2)$-avoiding r-graphs if n is relatively small compared to r. However, to the author's knowledge no such Steiner systems are known to exist. Of course for r fixed and n large, Steiner $(n, r, r-2)$ systems are known to exist due to the breakthrough work of Keevash [59] on the existence of designs.

A natural generalisation of a Sidon set is the notion of a B_h-set. For an integer $h \geq 2$, a subset A of an abelian group is called a B_h-set if all unordered sums $a_1 + \ldots + a_h$, where $a_i \in A$ are distinct. It is known [11] that for fixed h, there exist B_h-sets in $\mathbb{Z}_n$ with size at least $(1 + o(1))n^{1/h}$. Mimicking the construction in the proof of Theorem 2.3, but with the Sidon set replaced with such a B_h-set (for some $h < r$), yields a a partial Steiner $(n, r, r-h)$ system whose Lagrangian is $\Omega(\frac{1}{r^{2h} r!})$. It would be interesting to investigate whether this is within a constant of the maximum Lagrangian attained by an $(r-h)$-avoiding r graph for fixed h and large r. In [42] it is shown that this is the case for $h = 1$ and Theorem 2.3 shows that it also holds for $h = 2$.

3

Exact Ramsey Numbers of Odd Cycles via Nonlinear Optimisation

This chapter is based on joint work with my supervisor Jozef Skokan.

3.1 Introduction

Recall that for graphs $G_1, G_2, \ldots, G_k$, the *Ramsey number* $R(G_1, \ldots, G_k)$ is the least integer N such that any colouring of the edges of the complete graph K_N on N vertices with k colours contains a monochromatic copy of G_i in the i-th colour for some i, $1 \leq i \leq k$. In the case where $G_1, \ldots, G_k$ are all isomorphic to the graph G, we call $R(G_1, \ldots, G_k)$ the *k-colour Ramsey number* of G and denote it by $R_k(G)$.

In Section 1.3, we surveyed some classical results in Ramsey theory focusing on the Ramsey numbers of complete graphs. Recall that $R(t)$ denotes the Ramsey number $R(K_t, K_t)$ and recall the bounds $2^{t/2} \leq R(t) \leq 4^t$. Despite considerable effort over the past 80 years, the bases in the exponent in both of these bounds has not been improved.

This inertia has motivated the study of Ramsey numbers of graphs with a 'simpler' structure, where the problem may be more tractable. In this spirit, there has been a large body of research dedicated to the study of Ramsey numbers of graphs that are *sparse* in some sense (e.g. they have bounded maximum degree). The path on n vertices P_n and the cycle on n vertices C_n are particularly simple examples and were some of the earliest subjects in the study of Ramsey numbers of sparse graphs. Formally, P_n is the graph on vertex set $[n]$ and edge set $\{\{i,j\} : j - i = 1\}$ and C_n is the graph on vertex set $[n]$ and edge set $\{\{i,j\} : j - i \equiv 1 \pmod{n}\}$.

An early success in the Ramsey theory of sparse graphs was a result of Gerencsér and Gyárfás [47] from 1967 who showed that for all $n \geq m \geq 2$

$$R(P_n, P_m) = n + \left\lfloor \frac{m}{2} \right\rfloor - 1 \,.$$

We highlight the fact that here the Ramsey number is determined exactly in stark contrast to the results of Section 1.3. The behaviour of the Ramsey number $R(C_n, C_m)$ has been studied by several authors, including Bondy and Erdős [10], Faudree and Schelp [34], and Rosta [75], and is now fully determined. For example it is known that

$$R(C_n, C_n) = \begin{cases} 2n-1, & \text{if } n \geq 5 \text{ is odd,} \\ \frac{3n}{2} - 1, & \text{if } n \geq 6 \text{ is even.} \end{cases} \tag{3.1}$$

Results such as these that exactly determine $R(G_1, G_2)$ for a pair of graphs G_1, G_2 are by now fairly plentiful. See Radziszowski [58] for an excellent survey of such results. However, in the case where more than two colours are involved such results are still rather rare. Again, cycles and paths serve as natural starting points. Gyárfás, Ruszinkó, Sárközy and Szemerédi [50] showed that for n sufficiently large

$$R(P_n, P_n, P_n) = \begin{cases} 2n-1, & \text{if } n \text{ is odd,} \\ 2n-2, & \text{if } n \text{ is even.} \end{cases} \tag{3.2}$$

Benevides and Skokan [5] and Kohayakawa, Simonovits and Skokan [61] showed that for n sufficiently large

$$R(C_n, C_n, C_n) = \begin{cases} 4n-3, & \text{if } n \text{ is odd,} \\ 2n, & \text{if } n \text{ is even.} \end{cases} \tag{3.3}$$

Both (3.2) and (3.3) were established by the regularity method pioneered by Łuczak which we will return to shortly. The only non-trivial class of graphs for which the k-colour Ramsey number is exactly determined for arbitrary k is that of matchings. Letting mP_2 denote a matching of m edges, Cockayne and Lorimer [18] showed that for $m_1 \geq \ldots \geq m_\ell$ we have

$$R(m_1 P_2, \ldots, m_\ell P_2) = m_1 + 1 + \sum_{i=1}^{\ell} m_i.$$

In this chapter we address the following conjecture attributed to Bondy and Erdős [10].

Conjecture 3.1. *If $k \geq 2$ and $n > 3$ is odd then*

$$R_k(C_n) = 2^{k-1}(n-1) + 1.$$

Note that the conjecture deals specifically with the case where n is odd. Odd and even cycles behave rather differently in this context due to the fact that an even cycle is bipartite whereas an odd cycle is not (note the dichotomy in (3.1) and (3.3)). Erdős and Graham [33] proved the bounds

$$2^{k-1}(n-1)+1 \leq R_k(C_n) \leq (k+2)!n, \tag{3.4}$$

for all $k \geq 2$ and all odd $n > 3$. In this chapter we show that for fixed k and n large, the lower bound is correct.

Theorem 3.2. *For any fixed $k \geq 2$ and odd n sufficiently large,*

$$R_k(C_n) = 2^{k-1}(n-1)+1.$$

We therefore resolve Conjecture 3.1 for large n. We will in fact prove a stability-type strengthening of this result (see Theorem 3.4 below). Recently Day and Johnson [26] showed that in the opposite regime, where we fix an odd n and let k be sufficiently large, one in fact has $R_k(C_n) > (n-1)(2+\varepsilon)^{k-1}$ for some $\varepsilon = \varepsilon(n) > 0$, and so Conjecture 3.1 is false when n is small with respect to k. The qualification that n is sufficiently large in Theorem 3.2 is therefore necessary, however due to the use of compactness arguments in the proof, we obtain no effective bound on how large n must be with respect to k.

In view of Theorem 3.2, let us call a k-colouring of the complete graph on $2^{k-1}(n-1)$ vertices which does not contain a monochromatic copy of C_n an *extremal k-colouring.* The lower bound in (3.4) was established by observing that one can naturally construct extremal k-colourings by induction. Indeed if there exists a k-colouring of the edges of the complete graph K_m with no monochromatic C_n, then by joining two such copies of K_m by edges of colour $k+1$, one obtains a $(k+1)$-colouring of K_{2m} with no monochromatic C_n (here we use that C_n is non-bipartite). The base construction, for $k = 1$, is simply a monochromatic clique of size $n-1$. It was believed that all extremal k-colourings come from such a doubling argument. We show that this is not the case, providing a classification of extremal k-colourings which exposes a surprising correspondence between extremal k-colourings and perfect matchings in the k-dimensional hypercube Q_k.

The first breakthrough towards Conjecture 3.4 was made by Łuczak [65] who used the regularity method to show that the $k = 3$ case holds asymptotically i.e. that for n odd,

$$R(C_n, C_n, C_n) = 4n + o(n) \text{ as } n \to \infty.$$

Łuczak's method of applying regularity in this setting has proven extremely fruitful (see e.g. [38, 50, 61, 65, 66]) and has since become a standard tool. We will come to describe the method in more detail as we make crucial use of it in this chapter.

Building on Łuczak's ideas, Kohayakawa, Simonovits and Skokan [61] paired the regularity method with stability arguments to resolve Conjecture 3.1 for $k = 3$ and n large. The case where $k \geq 4$ remained open. Progress was made by Łuczak, Simonovits and Skokan [66] who showed that for $k \geq 4$ and odd n,

$$R_k(C_n) \leq k2^k n + o(n) \text{ as } n \to \infty.$$

To conclude this section we give a broad overview of the proof method of Theorem 3.2. Let $\mathcal{G}_n$ denote the (finite) set of all k-coloured cliques with no monochromatic copy of C_n. Determining $R_k(C_n)$ is then equivalent to determining the maximum number of vertices an element of $\mathcal{G}_n$ can have. Using the regularity method, we relate this problem to finding the maximum ℓ_1-norm of an element in a certain compact subset $\mathcal{S}$ of $\mathbb{R}^{3^k}$. This allows us to import analytic tools in support of our proof. The relation is such that maximal elements of $\mathcal{S}$ correspond to maximal elements of $\mathcal{G}_n$ i.e. extremal k-colourings for Theorem 3.2. Moreover by classifying the extremal points of $\mathcal{S}$ we can classify the extremal k-colourings and prove a stability type strengthening of Theorem 3.2, generalising the main result from [61]. We show that each perfect matching of the hypercube Q_k gives rise to a class of extremal k-colourings. On the other hand, any extremal k-colouring must be 'close' to one such construction. We defer precise statements to Section 3.2. The number of essentially different classes of extremal k-colourings is equal to the number of equivalence classes of perfect matchings in Q_k with respect to its automorphism group and this number is doubly exponential in k. Such a plethora of extremal constructions is usually forbidding when trying to establish stability type results, we believe that the fact we can overcome this obstacle is largely down to our analytic perspective.

Before continuing, let us collect some notation and terminology that we use throughout this chapter.

Let $W = w_0w_1 \dots w_\ell$ be a walk in G (that is a sequence of vertices $w_0, \dots, w_\ell$ such that w_iw_{i+1} is an edge of G for all $i < \ell$). If all of the w_i are distinct then we call W a *path of length* ℓ (so that P_n is a path of length $n-1$). We may also refer to W as a w_0w_ℓ-path to distinguish its endpoints. If all the w_i are distinct except $w_0 = w_\ell$ then we call W a *cycle of length* ℓ. We will also concatenate walks in the natural way. For example if $U = u_0 \dots u_m$ is a walk in G such that $u_m = w_0$, we let UW denote the walk $u_0 \dots u_mw_1 \dots w_\ell$. If x is a vertex such that $w_\ell x$ is an edge of G then we let Wx denote the walk $w_0 \dots w_\ell x$.

A *k-coloured graph* is a graph $G = (V, E)$ equipped with some function $\varphi : E \to [k]$. Furthermore, for $i \in [k]$, we let G_i denote the subgraph $(V, \varphi^{-1}\{i\})$ of G. We call G_i the ith *colour class* of G.

A *digraph* $D = (V, A)$ consists of a set of vertices V and a set $A \subseteq V^2$ i.e. a set of ordered pairs from V which we call *directed edges*. For $v \in V$ we let $d^+(v)$ denote the size of the set $\{u : (v, u) \in A\}$ and call $d^+(v)$ the *outdegree* of v. Similarly we define the *indegree* of v as $d^-(v) = |\{u : (u, v) \in A\}|$.

For $x \in \mathbb{R}^d$ we let $\|x\|$ denote the ℓ_1-norm of x i.e. $\|x\| = \sum_{i=1}^d |x_i|$. Furthermore, given $\varepsilon > 0$, we let $B_\varepsilon(x) := \{z \in \mathbb{R}^d : \|z - x\| < \varepsilon\}$, the open ball of radius ε centred at x. We let *supp*(x) denote the support of x i.e. the set $\{i \in [d] : x_i \neq 0\}$.

In the statements of theorems and lemmas it will be useful to use the notation $\alpha \ll \beta$ to mean that there is an increasing function $\alpha(x)$ so that the statement is valid for $0 < \alpha < \alpha(\beta)$. When we need to refer to this function at a later stage, we include the number of the lemma (or theorem) the function appears in as a subscript. For example, $\delta_{3.46}(x)$ denotes the implied function $\delta(x)$ from Lemma 3.46.

3.2 A Graph Decomposition, Extremal Colourings and Stability

In this section we describe the extremal colourings and give precise statements of the stability results discussed in Section 3.1. We also introduce some key concepts and results that will be used throughout the chapter and give a more detailed overview of our proof methods. We begin by introducing a way of decomposing an arbitrary k-coloured graph. This decomposition will play a central role for us and is similar to a decomposition introduced in [66].

3.2.1 A Graph Decomposition

Let G be a k-coloured graph. For each $i \in [k]$, we write $G_i = G_i' \cup G_i''$, where G_i' is the union of the bipartite components of G_i and G_i'' is the union of the non-bipartite components of G_i. For each $i \in [k]$, write $V(G_i') = V_0^i \cup V_1^i$ where V_0^i and V_1^i are the vertex classes of a bipartition of G_i' and set $V_*^i = V(G_i'')$. For $\tau \in \{0, 1, *\}^k$, let $V_\tau = \bigcap_{j=1}^k V_{\tau_j}^j$ and note that

$$V(G) = \bigcup_{\tau \in \{0,1,*\}^k} V_\tau, \quad \text{a disjoint union.}$$

We call $(V_\tau : \tau \in \{0, 1, *\}^k)$ a *profile partition* of G and we call the corresponding vector $(|V_\tau| : \tau \in \{0, 1, *\}^k)$ a *profile* of G. We will often denote a profile of G by $x(G)$. Note that G may admit multiple profile partitions since we made an arbitrary choice in choosing the bipartition $V(G_i') = V_0^i \cup V_1^i$ for each $i \in [k]$.

3.2.2 Extremal Colourings and the Hypercube

For $k \in \mathbb{N}$, we let Q_k denote the k-dimensional hypercube i.e. the graph on vertex set $\{0, 1\}^k$ and edge set consisting of pairs differing in exactly one coordinate. It will be useful to think of an element $\tau \in \{0, 1, *\}^k$ as a subcube of the k-dimensional hypercube Q_k via the correspondence

$$\tau \leftrightarrow Q(\tau) := \{c \in \{0, 1\}^k : c_j = \tau_j \text{ if } \tau_j \in \{0, 1\}\}.$$

In other words we think of a coordinate j where $\tau_j = *$ as a 'missing bit' and let $Q(\tau)$ be the set of all possible ways of filling in these bits. For example, if $k = 3$ and $\tau = (0, *, *)$ then

$$Q(\tau) = \{(0,0,0), (0,0,1), (0,1,0), (0,1,1)\}.$$

We define the *weight* of τ to be the size of the set $\{i \in [k] : \tau_i = *\}$ (i.e. the number of missing bits) and denote it by $\omega(\tau)$. Note that $|Q(\tau)| = 2^{\omega(\tau)}$. In the language of the hypercube, $\omega(\tau)$ is the dimension of the subcube $Q(\tau)$. In particular if $\omega(\tau) = 1$, then we think of $Q(\tau)$ as an *edge* of Q_k.

We can now describe a class of extremal k-colourings in terms of perfect matchings in Q_k. Let $\mathcal{M}$ be a perfect matching of Q_k. We express each edge of $\mathcal{M}$ as an element (of weight 1) of $\{0,1,*\}^k$. Let $G = K_N$ where $N = 2^{k-1}(n-1)$ and let $V(G) = \bigcup_{\tau \in \mathcal{M}} V_\tau$ be a partition of $V(G)$ where $|V_\tau| = n - 1$ for all $\tau \in \mathcal{M}$. For each $\tau \in \mathcal{M}$, colour all edges in $G[V_\tau]$ with the colour i, where i is the coordinate for which $\tau_i = *$. For $\tau, \sigma \in \mathcal{M}$, arbitrarily colour the edges between V_τ and V_σ with any colour j for which $\{\sigma_j, \tau_j\} = \{0, 1\}$ (i.e. edges τ, σ lie in opposite subcubes of Q_k of codimension 1 separated by the jth coordinate). It follows that each colour class of such a colouring is the disjoint union of cliques of size $n - 1$ and a bipartite graph and therefore contains no monochromatic copy of C_n. We call such a colouring a *hypercube colouring with clique size* $n-1$. See Figure 3.1 for an illustrated example.

If we inductively construct a perfect matching on Q_k by taking two perfect matchings on a disjoint pair of subcubes of codimension 1 and consider the associated hypercube colouring, we recover the inductive colourings of Erdős and Graham [33] described in Section 3.1. However, for $k \geq 4$, not all perfect matchings of Q_k decompose as the union of two matchings on a pair of codimension 1 subcubes, and so we obtain some genuinely new colourings (an example of which is depicted in Figure 3.1). In particular, a novel feature that appears for $k \geq 4$ colours is that there exist extremal k-colourings that contain monochromatic cliques of size $n-1$ in all k possible colours.

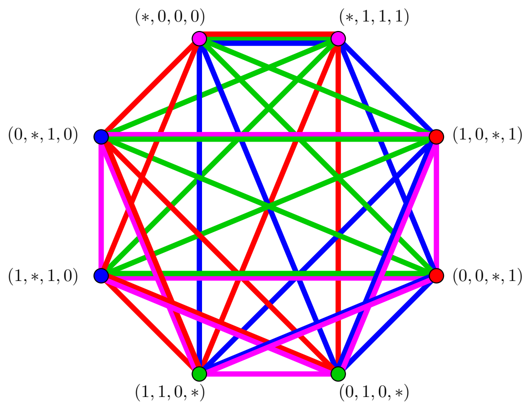


Figure 3.1 An extremal colouring for Theorem 3.2 in the case $k = 4$. Each node represents a clique of size $n - 1$ and is labelled by an edge of Q_4 where a '$*$' corresponds to the coordinate which changes across the edge. Each coordinate is associated with a colour and the position of the '$*$' determines the colour of each clique. Edges between cliques labelled τ, σ are coloured arbitrarily with colours j for which $\{\tau_j, \sigma_j\} = \{0, 1\}$.

3.2.3 Stability

In this subsection we state a theorem to the effect that the hypercube colourings considered in the previous subsection are the only extremal k-colourings for our problem. Moreover we assert that almost extremal colourings are in some sense 'close' to a hypercube colouring. Let us make this more precise.

Definition 3.3. *Let G and H be k-coloured graphs with $V(H) \subseteq V(G)$. Let $\varepsilon > 0$, then we say that G is* ε-close *to H if $|G_i \triangle H_i| \leq \varepsilon v(G)^2$ for all $i \in [k]$.*

Informally we may say that G and H as above are close in *edit distance*. We may now state the main result of this chapter.

Theorem 3.4. *Let $k \geq 2$, let $\frac{1}{n} \ll \eta \ll \varepsilon \ll 1$, where n is odd, and let $N > (2^{k-1} - \eta)n$. Then if $G = K_N$ is k-coloured with no monochromatic copy of C_n, then $N \leq 2^{k-1}(n - 1)$ and there exists a hypercube colouring H such that G is ε-close to H.*

Note that Theorem 3.2 follows as an immediate corollary. The $k = 3$ case of Theorem 3.4 was proved in [61] where the two classes of colourings the authors consider can be viewed as the colourings that arise from the two isomorphism classes of perfect matchings in Q_3. An interesting feature of Theorem 3.4 is that it deals with a wide variety of extremal colourings. Indeed if $\mathcal{M}_1$ and $\mathcal{M}_2$ are perfect matchings of Q_k that lie in distinct equivalence classes under the action of the automorphism group of Q_k, then there are hypercube colourings associated to $\mathcal{M}_1$ that are not isomorphic to any hypercube colouring associated to $\mathcal{M}_2$. It is also interesting to note that even though we can prove a stability statement around hypercube colourings, the structure of these colourings is not well understood. This is simply due to the fact that the structure of perfect matchings in the hypercube is not well understood. Indeed, even enumerating the perfect matchings (or their equivalence classes) in Q_k is a well-studied and difficult problem. Let $f(k)$ be the number of equivalence classes of perfect matchings in Q_k. It is clear that $f(3) = 2$ and so we obtain two essentially different extremal 3-colourings as in [61]. Graham and Harary [48] showed that $f(4) = 8$ and recently Östergård and Pettersson [69] determined (with a large amount of computer time) $f(5)$, $f(6)$ and $f(7)$. The function $f(k)$ grows rather rapidly; it is amusing to note that already $f(7) = 6071580464951208868206 21$ and so we have this many essentially different classes of extremal 7-colourings. It was shown in [17] that the number of perfect matchings in Q_k is $[(1+o(1))k/e]^{2^{k-1}}$ (although this result in fact follows from a theorem in [64, p.312]). Since the automorphism group of Q_k has size $k!2^k$ it follows that $f(k) = [(1+o(1))k/e]^{2^{k-1}}$ also.

3.2.4 Proof of Theorem 3.4: An Overview

Łuczak's application of the regularity method, discussed briefly in the introduction of this chapter, plays a central role in our proof. We include an informal discussion of the method here, deferring details until later. Those unfamiliar with the regularity method may skip this subsection and wait for the formal discussion in Section 3.7. We start with a definition.

Definition 3.5. *Let F be a connected graph whose largest matching saturates m vertices, then we call F a* connected matching of order m. *We distinguish a particular matching of largest size M_F in F and refer to an* *edge of M_F as a* matching edge *of F. If in addition F is non-bipartite, we call F an* odd *connected matching of order m.*

The idea behind Łuczak's application of the regularity method is as follows. Suppose that G is a k-coloured complete graph on N vertices. Let $G_1, \ldots, G_k$ be its colour classes. We apply the multicolour version of the Regularity Lemma [83] and obtain a regular partition of the vertex set $V(G)$ into $t+1$ classes $V(G) = V_0 \cup \ldots \cup V_t$. We construct an auxiliary graph R with vertex set $1, ..., t$ and the edge set formed by pairs $\{i, j\}$ for which (V_i, V_j) is regular with respect to $G_1, \ldots, G_k$. We colour each edge $\{i, j\}$ in R by the majority colour in the pair (V_i, V_j). The crucial point is that if R contains a monochromatic odd connected matching of order greater than m, then G contains a monochromatic cycle C_ℓ where ℓ can take essentially any odd value smaller than mN/t. It follows that if G contains no monochromatic copy of C_n, then R cannot contain a monochromatic odd connected matching of order larger than nt/N. The advantage of this perspective is that forbidding a large connected matching is far more restrictive than forbidding a cycle of a given length. Indeed a cycle is itself an example of a connected matching, and so if a graph contains no connected matching of order greater than m then it contains no cycle of length greater than m. The following theorem of Erdős and Gallai [32] shows that this is a very strict condition.

Theorem 3.6. *Let $m \geq 3$. If G is a graph which contains no cycle of length greater than m, then $e(G) \leq m(v(G)-1)/2$.*

The price one pays is that R is not a complete graph, however it can be chosen to be as dense as one likes. We are now able to state a theorem that is a major stepping stone toward the proof of Theorem 3.4. (Recall that $\|\cdot\|$ denotes the ℓ_1-norm).

Theorem 3.7. *Let $k \geq 2$ and let $\frac{1}{n} \ll \delta \ll \varepsilon \ll 1$, where n is odd. If G is a k-coloured graph with $v(G) = 2^{k-1}n$ and $e(G) \geq (1-\delta)\binom{v(G)}{2}$ containing no monochromatic odd connected matching of order $\geq (1+\delta)n$, then for any choice of profile $x(G)$ of G, there exists a hypercube colouring H with profile*

$x(H)$ satisfying

$$\|x(G) - x(H)\| \leq \varepsilon n.$$

The proof of Theorem 3.7 occupies the majority of this chapter. In the final section we show how Theorem 3.4 follows from Theorem 3.7 via combinatorial stability arguments and the regularity method. The outline of the proof of Theorem 3.7 is as follows. Let G be as in the statement of Theorem 3.7 and let $x(G)$ denote a profile of G. Our starting point is to translate the combinatorial constraint of containing no large monochromatic odd connected matching into an analytic constraint on $x(G)$ of the form

$$F(x(G)) \leq 0, \tag{3.5}$$

where F is a quadratic form which we derive in the next section. We then view (3.5) as a constraint in an optimisation problem where we wish to maximise the objective function $\|x(G)\|$. Recalling that $\|x(G)\| = v(G)$, we get a corresponding upper bound on the order of G. It turns out that optimal solutions to this optimisation problem correspond to the profiles of hypercube colourings. Solving the optimisation problem is the subject of Sections 3.4 and 3.5. In Section 3.6 we use compactness arguments to show that almost optimal solutions must be close in ℓ_1-norm to the profile of a hypercube colouring. We then translate this analytic stability into the more combinatorial stability statement of Theorem 3.7. Note that Theorem 3.7 will be applied to a reduced graph like the one described above. The focus of the final section is to show that if the profile of this reduced graph is close in ℓ_1-norm to the profile of a hypercube colouring, then the original graph is close in edit distance to a hypercube colouring.

3.3 Deriving the Analytic Constraints

Given a k-coloured graph G, we will show how to translate the combinatorial constraint of containing no large monochromatic odd connected matching into an analytic constraint on the profile of G.

From here on, throughout the chapter, we let $k \geq 2$ be a fixed integer. Let G be a k-coloured graph. First we distinguish between two types of edges of G.

If $e \in E(G)$ is coloured with the colour j and lies in a bipartite component of G_j then we call e a *bipartite* edge. We call e *non-bipartite* otherwise. Let $(V_\tau : \tau \in \{0,1,*\}^k)$ be a profile partition of G. We make two simple observations regarding the profile partition of a k-coloured graph.

Observation 3.8. *If $e \in E(G)$ is a bipartite edge of colour j then it must have endpoints in parts V_τ, V_σ for some $\tau, \sigma \in \{0,1,*\}^k$ such that $\tau_j = 0$ and $\sigma_j = 1$.*

Observation 3.9. *If $e \in E(G)$ is a non-bipartite edge of colour j then it must have endpoints in parts V_τ, V_σ for some (not necessarily distinct) $\tau, \sigma \in \{0,1,*\}^k$ such that $\tau_j = \sigma_j = *$.*

This motivates the following definitions.

Definition 3.10. *We say that $\sigma, \tau \in \{0,1,*\}^k$ are* distinguishable *if $\{\sigma_j, \tau_j\} = \{0,1\}$ for some $j \in [k]$. We say that σ and τ are* indistinguishable *otherwise.*

Definition 3.11. *If $\sigma, \tau \in \{0,1,*\}^k$ are such that either (i) σ, τ are distinguishable or (ii) $\sigma_j = \tau_j = *$ for some $j \in [k]$, then we say that σ and τ are* compatible. *We say that σ, τ are* incompatible *otherwise.*

Viewing elements of $\{0,1,*\}^k$ as subcubes of Q_k, we may reinterpret these definitions as follows.

Lemma 3.12. *Let $\sigma, \tau \in \{0,1,*\}^k$. Then σ, τ are distinguishable if and only if $Q(\tau) \cap Q(\sigma) = \emptyset$. Furthermore, σ, τ are incompatible if and only if $|Q(\tau) \cap Q(\sigma)| = 1$.*

Proof. By the definition of the sets $Q(\tau), Q(\sigma)$ we have

$$Q(\tau) \cap Q(\sigma) = \{c \in \{0,1\}^k : c_j = \tau_j \text{ if } \tau_j \in \{0,1\} \text{ and } c_j = \sigma_j \text{ if } \sigma_j \in \{0,1\}\}.$$

This is empty if and only if there exists a $j \in [k]$ such that $\sigma_j, \tau_j \in \{0,1\}$ and $\sigma_j \neq \tau_j$ i.e. if and only if σ, τ are distinguishable. Let $T = \{i \in [k] : \sigma_i = \tau_i = *\}$. If σ, τ are indistinguishable then we see that $|Q(\tau) \cap Q(\sigma)| = 2^{|T|}$. Therefore, $|Q(\tau) \cap Q(\sigma)| = 1$ if and only if σ, τ are indistinguishable and $T = \emptyset$ i.e. σ, τ are incompatible. $\square$

From now on, we let

$$\Delta = \left\{ \{\sigma, \tau\} \in \binom{\{0,1,*\}^k}{2} : \ \sigma, \tau \text{ are distinguishable} \right\}.$$

It will also be convenient to make the following definition.

Definition 3.13. *Let $\alpha > 0$ and let G be a graph such that $e(G) \geq \alpha\binom{v(G)}{2}$. Then we say that G is* α-dense*.*

This next proposition provides the link between our combinatorial problem and a problem in nonlinear optimisation. (Recall the definition of a connected matching, Definition 3.5).

Proposition 3.14. *Let $C > 1$, $0 < \delta < 1$ and let $n > 1/\delta$. Suppose that G is a $(1-\delta)$-dense, k-coloured graph with $v(G) = Cn$, containing no monochromatic odd connected matching of order $\geq (1+\delta)n$. Let x be a profile of G and let $v = x/n$. Then the following hold:*

1. $$\left(\sum_{\tau \in \{0,1,*\}^k} v_\tau \right)^2 - 2 \sum_{\{\sigma,\tau\} \in \Delta} v_\sigma v_\tau - \sum_{\tau \in \{0,1,*\}^k} \omega(\tau) v_\tau \leq \delta k C^2.$$

2. *$v_\tau \leq 1 + 2\sqrt{\delta} C$ whenever $\omega(\tau) = 1$.*

3. *$v_\tau v_\sigma \leq 2\delta C^2$ whenever σ and τ are incompatible.*

Proof. Let us first remind ourselves of the graph decomposition discussed in Subsection 3.2.1. For each $i \in [k]$, we write $G_i = G_i' \cup G_i''$, where G_i' is the union of the bipartite components of G_i and G_i'' is the union of the non-bipartite components of G_i. For each $i \in [k]$, write $V(G_i') = V_0^i \cup V_1^i$ where V_0^i and V_1^i are the vertex classes of a bipartition of G_i' and set $V_*^i = V(G_i'')$. For $\tau \in \{0,1,*\}^k$, set $V_\tau = \bigcap_{j=1}^k V_{\tau_j}^j$. Let $x = (|V_\tau| : \tau \in \{0,1,*\}^k)$ be the profile corresponding to this partition. Let $N = v(G)$ and note that

$$N = \sum_{\tau \in \{0,1,*\}^k} x_\tau. \tag{3.6}$$

It follows from Observation 3.8 that the number of bipartite edges in G is at most $\sum_{\{\sigma,\tau\}\in\Delta} x_\sigma x_\tau$. Letting e_0 denote the number of non-bipartite edges in G we therefore have that

$$e_0 \geq e(G) - \sum_{\{\sigma,\tau\}\in\Delta} x_\sigma x_\tau. \tag{3.7}$$

Since $N \geq 1/\delta$, we have

$$e(G) \geq (1-\delta)\binom{N}{2} \geq (1-2\delta)\frac{N^2}{2}. \tag{3.8}$$

Combining (3.6), (3.7) and (3.8) gives

$$e_0 \geq \frac{1}{2}\left(\sum_{\tau\in\{0,1,*\}^k} x_\tau\right)^2 - \sum_{\{\sigma,\tau\}\in\Delta} x_\sigma x_\tau - \delta N^2. \tag{3.9}$$

We now find a corresponding upper bound for e_0. Recall that for $\tau \in \{0,1,*\}^k$, the weight $\omega(\tau)$ of τ is defined to be the size of the set $\{i \in [k] : \tau_i = *\}$.

By assumption, for each $i \in [k]$, every connected component of G_i'' has no matching on $(1+\delta)n$ vertices and so in particular G_i'' has no cycle of length greater than $(1+\delta)n$. The Erdős-Gallai Theorem, Theorem 3.6, therefore implies that

$$e(G_i'') \leq (1+\delta)\frac{n}{2}|V_*^i|. \tag{3.10}$$

Observe that

$$|V_*^i| = \sum_{\{\tau\in\{0,1,*\}^k:\tau_i=*\}} x_\tau. \tag{3.11}$$

Since each non-bipartite edge of G belongs to $E(G_i'')$ for some i, (3.10) and (3.11) provide the upper bound

$$e_0 \leq \sum_{i=1}^k e(G_i'') \leq (1+\delta)\frac{n}{2}\sum_{i=1}^k |V_*^i| = (1+\delta)\frac{n}{2}\sum_{\tau\in\{0,1,*\}^k}\omega(\tau)x_\tau. \tag{3.12}$$

Since $\omega(\tau) \leq k$ for all $\tau \in \{0,1,*\}^k$ by definition, (3.6) and (3.12) imply the bound

$$e_0 \leq \frac{1}{2}\delta nkN + \frac{n}{2}\sum_{\tau\in\{0,1,*\}^k}\omega(\tau)x_\tau. \tag{3.13}$$

Recall that $v = x/n$. Comparing the bounds (3.9) and (3.13) and scaling the resulting inequality by $2/n^2$ yields

$$\left(\sum_{\tau\in\{0,1,*\}^k} v_\tau\right)^2 - 2\sum_{\{\sigma,\tau\}\in\Delta} v_\sigma v_\tau - 2\delta C^2 \le \sum_{\tau\in\{0,1,*\}^k} \omega(\tau)v_\tau + \delta kC.$$

This establishes (1). Notice that if $\tau \in \{0,1,*\}^k$ is such that $\omega(\tau) = 1$, then $G[V_\tau]$ is monochromatic with all edges non-bipartite by Observations 3.8 and 3.9. $G[V_\tau]$ therefore contains no cycle of length greater than $(1+\delta)n$ and so by Theorem 3.6 and the fact that G has at most $\delta\binom{N}{2}$ edges missing

$$\binom{x_\tau}{2} - \delta\binom{N}{2} \le e(G[V_\tau]) \le (1+\delta)\frac{n}{2}x_\tau.$$

It follows that

$$v_\tau^2 \le (1+2\delta)v_\tau + \delta C^2,$$

from which (2) follows. Finally, let us note that by Observations 3.8 and 3.9, if σ, τ are incompatible, then there can be no edges lying between V_σ and V_τ. Since G has at most $\delta\binom{N}{2}$ edges missing we must then have

$$x_\tau x_\sigma \le 2\delta N^2$$

(note that this inequality also accounts for the case where $\sigma = \tau$) and so (3) follows.

$\square$

Given a graph G its profile lies in the space $\mathbb{R}^{\{0,1,*\}^k}$ which we will denote by $\mathbb{R}^*$. In view of Proposition 3.14 we define the function $F : \mathbb{R}^* \to \mathbb{R}$ by

$$F(x) = \left(\sum_{\tau\in\{0,1,*\}^k} x_\tau\right)^2 - 2\sum_{\{\sigma,\tau\}\in\Delta} x_\sigma x_\tau - \sum_{\tau\in\{0,1,*\}^k} \omega(\tau)x_\tau.$$

Let us also define the following subsets of $\mathbb{R}^*$.

$\boldsymbol{X(\gamma)}$: For $\gamma \ge 0$, let $X(\gamma)$ denote the set of elements $x \in \mathbb{R}^*$ satisfying:

(X1) $F(x) \le \gamma$

(X2) $x_\tau \le 1+\gamma$ whenever $w(\tau)=1$.

(X3) $x_\tau x_\sigma \le \gamma$ whenever σ and τ are incompatible.

(X4) $x_\tau \ge 0$ for all τ.

Now let G be as in the statement of Theorem 3.7 and let x be a profile of G. By the above proposition we have $x/n \in X(\sqrt{\delta}k2^{2k})$ whereas we also have $\|x\| = 2^{k-1}n$. We will show that for δ small, this means that x/n is an element of almost maximal norm in $X(\sqrt{\delta}k2^{2k})$. We will also show that elements of large norm in $X(\sqrt{\delta}k2^{2k})$ have a very specific structure (in fact they resemble the profile of a hypercube colouring) and so this imposes a lot of structure on x. For now we focus our attention on the set $X(0)$ which we denote simply by X. Later on, we use compactness arguments to relate properties of X and $X(\gamma)$ for γ small.

Our next goal is to classify elements of maximal ℓ_1-norm in X. To describe these elements we need a definition.

Definition 3.15. *Call a set $\mathcal{A} \subseteq \{0,1,*\}^k$* distinguishable *if every pair of distinct elements of $\mathcal{A}$ are distinguishable and also $\omega(\tau) \ge 1$ for all $\tau \in \mathcal{A}$.*

The requirement that elements have weight at least 1 is for notational convenience later in the chapter. Viewing elements of $\{0,1,*\}^k$ as subcubes of Q_k, a distinguishable set is simply a collection of disjoint subcubes of Q_k (of dimension at least 1). If this collection covers the whole cube we give it a special name.

Definition 3.16. *Call a distinguishable set $\mathcal{A} \subseteq \{0,1,*\}^k$ a* decomposition *if $\bigcup_{\tau\in\mathcal{A}} Q(\tau) = \{0,1\}^k$.*

Let us quickly record a simple result concerning distinguishable sets which will become useful later.

Lemma 3.17. *Let $\mathcal{D} \subset \{0,1,*\}^k$ be a distinguishable set. Then*

$$\sum_{\tau\in\mathcal{D}} 2^{\omega(\tau)} \le 2^k,$$

with equality if and only if $\mathcal{D}$ is a decomposition.

Proof. This is simply the observation that a distinguishable set $\mathcal{D}$ is a collection of disjoint subcubes of Q_k and so the sum of their sizes $\sum_{\tau\in\mathcal{D}} 2^{\omega(\tau)}$ is bounded by the size of Q_k. Moreover we have equality if and only if these subcubes cover all of Q_k i.e. $\mathcal{D}$ is a decomposition. $\square$

We define the following subset of $\mathbb{R}^*$.

O: Let O denote the set of elements $x \in \mathbb{R}^*$ satisfying:

(O1) $supp(x)$ is a decomposition where $\omega(\tau) = 1$ or 2 for all $\tau \in supp(x)$.

(O2) For all $\tau \in supp(x)$, if $\omega(\tau) = 1$ then $x_\tau = 1$ and if $\omega(\tau) = 2$ then $x_\tau = 2$.

It is easy to check that $O \subseteq X$. The next proposition asserts that O is the set of elements of maximal ℓ_1-norm in X.

Proposition 3.18. *If $x \in X$, then $\|x\| \le 2^{k-1}$ with equality if and only if $x \in O$.*

We note that the 'if' statement in the above proposition is immediate. Indeed if $x \in O$ then $supp(x)$ is a decomposition so that $\sum_{\tau\in supp(x)} 2^{\omega(\tau)} = 2^k$ by Lemma 3.17. Moreover $2^{\omega(\tau)} = 2x_\tau$ for all $\tau \in supp(x)$ by (O2).

Definition 3.19. *If $x \in X$ is such that $\|x\| = \sup_{z\in X}\|z\|$ then we say that x is an* optimal point *of X.*

We note that since X is compact, optimal points of X exist. The proof of Proposition 3.18 is split over the next two sections.

3.4 Compressions and a Spherical Constraint

In this section we make the first steps towards a proof of Proposition 3.18. Broadly speaking we apply the combinatorial technique of 'shifting' or 'compression' to transform the complicated nonlinear constraint in the definition of $X = X(0)$ into a spherical constraint which is much more amenable to analysis. In Section 3.5 we apply optimisation tools to this transformed problem. We begin with a simple lemma concerning elements of $\{0, 1, *\}^k$.

Lemma 3.20. *If $\sigma, \tau \subseteq \{0,1,*\}^k$ are indistinguishable and compatible and $\omega(\tau) = 1$, then $Q(\tau) \subseteq Q(\sigma)$. In particular if $\omega(\sigma) = 1$ also, then $\sigma = \tau$.*

Proof. Since σ, τ are indistinguishable and compatible we have $|Q(\tau) \cap Q(\sigma)| \geq 2$ by Lemma 3.12. However, $|Q(\tau)| = 2$ and so it follows that $Q(\tau) \subseteq Q(\sigma)$. If $\omega(\sigma) = 1$ also, then clearly $Q(\sigma) = Q(\tau)$ i.e. $\sigma = \tau$. □

Definition 3.21. *Let $x \in \mathbb{R}^*$. If all pairs of (not necessarily distinct) elements of supp(x) are compatible, then we say that x has* compatible support.

Let us note that condition (X3) (with $\gamma = 0$) in the definition of the set X is simply the condition that elements of X have compatible support. In particular, if $x \in X$ and $\tau \in \{0,1,*\}^k$ has weight 0, then $x_\tau = 0$ since τ is not compatible with itself.

The following lemma establishes an important property of optimal points. For $\tau \in \{0,1,*\}^k$ we let $e_\tau \in \mathbb{R}^*$ denote the standard unit vector whose entries are all 0 except the entry labelled τ which is 1.

Lemma 3.22. *Let $x \in X$ be an optimal point then $F(x) = 0$.*

Proof. Suppose for contradiction that $F(x) < 0$. Assume first that there exists $\tau \in supp(x)$ with $\omega(\tau) \geq 2$. By the continuity of F we may choose $\alpha > 0$ small enough so that $F(x + \alpha e_\tau) < 0$. Let $x' = x + \alpha e_\tau$. Since $supp(x') = supp(x)$ it is clear that $x' \in X$. However, $\|x'\| = \|x\| + \alpha > \|x\|$ contradicting the fact that x is optimal.

We may assume then that $supp(x)$ consists only of elements of weight 1 and therefore is a distinguishable set by Lemma 3.20 and the fact that x has compatible support. It follows from the definition of F that

$$F(x) = \sum_{\tau \in supp(x)} (x_\tau^2 - x_\tau) < 0,$$

and so $x_\tau < 1$ for some $\tau \in supp(x)$. As before there exists some $\alpha > 0$ sufficiently small so that $F(x + \alpha e_\tau) < 0$. Let $x' = x + \alpha e_\tau$. If we pick α small enough so that $x'_\tau = x_\tau + \alpha \leq 1$ also, then again we have $x' \in X$ with $\|x'\| > \|x\|$ contradicting the optimality of x. □

We now describe the transformations alluded to at the beginning of this section. They will be of great use in simplifying our analysis of optimal points of X.

Definition 3.23. *Let $x \in \mathbb{R}^*$ and let $\pi, \rho \in \{0, 1, *\}^k$ be distinct. We define the* (π, ρ)-compression *of x, denoted $x(\pi, \rho)$, as follows:*

- *If $\omega(\rho) \geq 2$, or if $\omega(\rho) \leq 1$ and $x_\pi + x_\rho < 1$, then let $x(\pi, \rho)$ be the vector x' with coordinates: $x'_\pi = 0$, $x'_\rho = x_\pi + x_\rho$ and $x'_\tau = x_\tau$ for all $\tau \in \{0, 1, *\}^k \backslash \{\pi, \rho\}$.*
- *If $\omega(\rho) \leq 1$ and $x_\pi + x_\rho \geq 1$ then let $x(\pi, \rho)$ be the vector x' with coordinates: $x'_\pi = x_\pi + x_\rho - 1$, $x'_\rho = 1$ and $x'_\tau = x_\tau$ for all $\tau \in \{0, 1, *\}^k \backslash \{\pi, \rho\}$.*

If $x(\pi, \rho) = x$ then we say that x is (π, ρ)-compressed.

Let $x \in X$ be an optimal point, we will be interested in instances where $x(\pi, \rho)$ is also an optimal point of X. We observe that if $x \in \mathbb{R}^*$ and $\pi, \rho \in \{0, 1, *\}^k$ are distinct then $\|x(\pi, \rho)\| = \|x\|$. However, if $x \in X$ then it does not follow in general that $x(\pi, \rho) \in X$.

For reasons that will become clear, we only consider (π, ρ)-compressions in the case where π and ρ are indistinguishable. It will therefore be useful to associate to each point $x \in X$, the digraph $D(x) = (V(x), E(x))$ where $V(x) = supp(x)$ and

$$E(x) = \{(\pi, \rho) : \pi, \rho \text{ are distinct, indistinguishable and } x(\pi, \rho) \in X\}.$$

In particular if $x \in X$ is (π, ρ)-compressed, where π and ρ are distinct and indistinguishable, then $(\pi, \rho) \in E(x)$. We draw attention to the fact that edges of $D(x)$ only occur between indistinguishable pairs. Conversely, the following lemma shows that, when $x \in X$ is optimal, at least one edge occurs between any indistinguishable pair in $D(x)$.

Lemma 3.24. *Let $x \in X$ be optimal and suppose that $\pi, \rho \in V(x)$, are indistinguishable and distinct. Then one of the following holds:*

(i) x is (π, ρ)-compressed, $x_\rho = 1$, $\omega(\rho) = 1$, $\omega(\pi) \geq 2$ and $(\rho, \pi) \notin E(x)$,

(ii) x *is* (ρ, π)*-compressed,* $x_\pi = 1$, $\omega(\pi) = 1$, $\omega(\rho) \geq 2$ *and* $(\pi, \rho) \notin E(x)$,

(iii) (ρ, π) *and* (π, ρ) *both lie in* $E(x)$.

Proof. Recall that

$$F(x) = \left(\sum_{\tau \in \{0,1,*\}^k} x_\tau \right)^2 - 2 \sum_{\{\sigma,\tau\} \in \Delta} x_\sigma x_\tau - \sum_{\tau \in \{0,1,*\}^k} \omega(\tau) x_\tau,$$

where Δ is the set of unordered distinguishable pairs from $\{0, 1, *\}^k$. Since π, ρ are indistinguishable, the sum $\sum_{\{\sigma,\tau\}\in\Delta} x_\sigma x_\tau$ does not contain the term $x_\rho x_\pi$. We may therefore express $F(x)$ in the form

$$F(x) = \left(\sum_{\tau \in \{0,1,*\}^k} x_\tau \right)^2 - Ax_\rho - Bx_\pi - C, \tag{3.14}$$

where A, B and C do not depend on x_ρ or x_π and $A, B \geq 0$. Suppose that $A \geq B$ and let $x' = x(\pi, \rho)$. Let us show that $(\pi, \rho) \in E(x)$ i.e. $x' \in X$. By (3.14) we have $F(x') \leq F(x)$ and so x' satisfies $(X1)$ in the definition of X. By the definition of (π, ρ)-compression it is clear that x' also satisfies $(X2)$ and $(X4)$. Since $\rho \in supp(x)$, we also have $supp(x') \subseteq supp(x)$. Since x has compatible support the same is true for x' i.e. x' satisfies (X3) and so $x' \in X$. Note that since compressions preserve the ℓ_1-norm, x' is also an optimal point of X.

In the case $A = B$, an identical argument shows that $(\rho, \pi) \in E(x)$ also, and so (iii) holds.

Suppose then that $A > B$. In this case, looking again at (3.14), we see that if x is not (π, ρ)-compressed then we in fact have $F(x') < F(x) = 0$, contradicting Lemma 3.22. We conclude that x is (π, ρ)-compressed. Suppose $\omega(\rho) \geq 2$, then by the definition of (π, ρ)-compression we have $x_\pi = x'_\pi = 0$ contradicting the fact that $\pi \in supp(x)$ and so $\omega(\rho) = 1$. Since π and ρ are compatible, indistinguishable and distinct, it follows from Lemma 3.20 that $\omega(\pi) \geq 2$. Let $x'' = x(\rho, \pi)$. It follows that $x''_\rho = 0$ and $x''_\pi = x_\rho + x_\pi$ and so by (3.14), $F(x'') > F(x) = 0$. We conclude that $x'' \notin X$ i.e. $(\rho, \pi) \notin E(x)$. The fact that $x_\rho = 1$ follows from the fact that $\omega(\rho) = 1$

and x is (π, ρ)-compressed. Thus, (i) holds, and similarly if $A < B$ then (ii) holds. $\square$

We obtain the following immediate corollary.

Corollary 3.25. *Let $x \in X$ be an optimal point and suppose that I is an independent set in $D(x)$. Then I is a distinguishable set.*

Definition 3.26. *We call an optimal point $x \in X$* compressed *if it is (π, ρ)-compressed for all $(\pi, \rho) \in E(x)$.*

We now show that compressed optimal points of X exist. In fact we show that given any optimal point of $x \in X$ we may obtain a compressed optimal point by applying a finite number of compressions to x. The simpler structure of compressed optimal points will make it easier to bound their ℓ_1-norm which is the goal of Proposition 3.18.

Lemma 3.27. *Compressed optimal points of X exist.*

Proof. Let x be an arbitrary optimal point of X and define a sequence $x_0, x_1, x_2, \ldots$ of elements of X recursively as follows: Set $x_0 = x$. Having chosen $x_0, \ldots, x_t$, if x_t is compressed then stop the sequence at x_t. If not, then there exists $(\pi, \rho) \in E(x_t)$ such that x_t is not (π, ρ)-compressed. By Lemma 3.24, we must therefore have that $x(\pi, \rho)$ and $x(\rho, \pi)$ are both optimal points of X. Note that by the definition of $D(x_t)$, ρ and π are indistinguishable and $\{\pi, \rho\} \subseteq supp(x_t)$. Since x_τ has compatible support, it follows from Lemma 3.20 that either $\omega(\pi) \geq 2$ or $\omega(\rho) \geq 2$. If $\omega(\rho) \geq 2$ then set $x_{t+1} = x(\pi, \rho)$, if not (so that $\omega(\pi) \geq 2$) set $x_{t+1} = x(\pi, \rho)$. In either case x_{t+1} is an optimal point of X satisfying $|V(x_{t+1})| = |V(x_t)| - 1$. Since $0 \leq |V(x)| \leq 3^k$ for all $x \in X$ it follows that the sequence must terminate in at most 3^k steps. $\square$

Having discovered compressed optimal points, we now explore some of their properties. First we need a definition.

Definition 3.28. *A* star *is a digraph with vertex set $\{\rho, \pi_1, \ldots, \pi_m\}$ (for some $m \geq 0$) and edge set $\{(\rho, \pi_1), \ldots, (\rho, \pi_m)\}$. We refer to ρ as the* root *of the star and we call* $\pi_1, \ldots, \pi_m$ leaves*. Note that we have included the possibility of a star with no leaves.*

Lemma 3.29. *Let $x \in X$ be a compressed optimal point, then $D(x)$ is a disjoint union of stars. Moreover if ρ is a root of positive outdegree then $\omega(\rho) \geq 2$ and if π is a leaf then $\omega(\pi) = 1$ and $x_\pi = 1$.*

Proof. It suffices to prove the following:

1. If $(\rho, \pi) \in E(x)$ then $\omega(\rho) \geq 2$, $\omega(\pi) = 1$, $x_\pi = 1$ and $(\pi, \rho) \notin E(x)$.
2. If $(\rho_1, \pi), (\rho_2, \pi) \in E(x)$ then $\rho_1 = \rho_2$.

Suppose that $(\rho, \pi) \in E(x)$, in particular ρ and π are indistinguishable. If $(\pi, \rho) \in E(x)$ also, then since x is compressed we have by definition that $x(\rho, \pi) = x = x(\pi, \rho)$. However, from the definition of compression we see that the only way we can have $x(\rho, \pi) = x(\pi, \rho)$ is if $\omega(\pi) = \omega(\rho) = 1$. But then by Lemma 3.20, $\pi = \rho$, a contradiction. We conclude that $(\pi, \rho) \notin E(x)$ and so (1) follows from Lemma 3.24.

Suppose now that $(\rho_1, \pi), (\rho_2, \pi) \in E(x)$. By (1) we know that $\omega(\pi) = 1$ and $\omega(\rho_i) \geq 2$ for $i = 1, 2$. By Lemma 3.20 it follows that $Q(\pi) \subseteq Q(\rho_1) \cap Q(\rho_2)$ and so ρ_1, ρ_2 are indistinguishable by Lemma 3.12. If $\rho_1 \neq \rho_2$ then by Lemma 3.24 we have that either $(\rho_1, \rho_2) \in E(x)$ or $(\rho_2, \rho_1) \in E(x)$, but this contradicts (1). $\square$

Given a compressed optimal point $x \in X$ let

$$L(x) = \{\tau \in V(x) : d^-(\tau) > 0\}$$

and

$$R(x) = V(x) \backslash L(x).$$

By Lemma 3.29, $L(x)$ and $R(x)$ are the set of leaves and the set of roots of $D(x)$ respectively.

Lemma 3.30. *Let $x \in X$ be a compressed optimal point. Then*

$$F(x) = \sum_{\tau \in R(x)} \left(x_\tau^2 + (2d^+(\tau) - \omega(\tau)) x_\tau \right).$$

Proof. Lemma 3.29 shows that for any indistinguishable pair $\pi, \rho \in V(x)$, where $\pi \neq \rho$, exactly one of (π, ρ) and (ρ, π) is in $E(x)$ and so

$$F(x) = \sum_{\tau \in V(x)} x_\tau^2 + 2 \sum_{(\sigma,\tau)\in E(x)} x_\sigma x_\tau - \sum_{\tau \in V(x)} \omega(\tau) x_\tau.$$

By Lemma 3.29 we may write

$$\sum_{(\sigma,\tau)\in E(x)} x_\sigma x_\tau = \sum_{\rho \in R(x)} x_\rho \left(\sum_{\pi:(\rho,\pi)\in E(x)} x_\pi \right) = \sum_{\rho\in R(x)} d^+(\rho) x_\rho.$$

Moreover by Lemma 3.29 we have $\sum_{\tau \in L(x)} (x_\tau^2 - \omega(\tau)x_\tau) = 0$. The result follows. □

The key feature here is that for a compressed optimal point x, the constraint equation $F(x) = 0$ is spherical. This allows us to more easily apply standard optimisation techniques and this will be the concern of the next section. For now it will be useful for us to establish some degree conditions on the vertices of $D(x)$ for a compressed optimal point $x \in X$.

Lemma 3.31. *Let $x \in X$ be a compressed optimal point, then $d^+(\sigma) \leq 2^{\omega(\sigma)-1}$ for all $\sigma \in V(x)$.*

Proof. Suppose that $\rho \in V(x)$ is such that $d^+(\rho) > 0$. By Lemma 3.29, ρ is the root of a star in $D(x)$. Let L be the set of leaves of this star (so in particular $|L| = d^+(\rho)$ and $\omega(\pi) = 1$ for all $\pi \in L$). Note that L is an independent set in $D(x)$ and therefore it is a distinguishable set by Corollary 3.25. Note further that for each $\pi \in L$, ρ and π are indistinguishable and compatible and hence $Q(\pi) \subseteq Q(\rho)$ by Lemma 3.20. It follows that

$$2d^+(\rho) = \sum_{\pi\in L} |Q(\pi)| \leq |Q(\rho)| = 2^{\omega(\rho)}.$$

□

We can now bootstrap, using the previous two lemmas to establish a much stronger degree condition. The idea behind the proof of the following lemma is readily explained however it is notationally laborious. The idea is that if $x \in X$ is a compressed optimal point and a star in $D(x)$ with root ρ has

$> \omega(\rho)$ leaves, then one can contradict the optimality of x by replacing this star with a collection of stars whose roots have less weight. First let us generalise an earlier notation.

Definition 3.32. *Let $\sigma \in \{0, 1, *\}^k$ and let $W = \{i \in [k] : \sigma_i = *\}$. Then for $S \subseteq W$ define*

$$Q(\sigma; S) = \{\tau \in \{0, 1, *\}^k : \tau_i \in \{0, 1\} \text{ for } i \in W \backslash S \text{ and } \tau_i = \sigma_i \text{ otherwise}\}.$$

Note that elements of $Q(\sigma; S)$ are pairwise distinguishable and that $Q(\sigma; \emptyset)$ is simply the set $Q(\sigma)$. We may think of $Q(\sigma; S)$ as a decomposition of $Q(\sigma)$ into 'parallel' subcubes of dimension $|S|$.

Lemma 3.33. *Let $x \in X$ be a compressed optimal point, then $d^+(\sigma) \leq \omega(\sigma)$ for all $\sigma \in V(x)$.*

Proof. By Lemma 3.29 we may write $V(x) = S_1 \cup \ldots \cup S_q$, a disjoint union where each S_i is the vertex set of a star in $D(x)$. Suppose that there exists $\sigma \in V(x)$ such that $d^+(\sigma) > \omega(\sigma)$. Without loss of generality assume σ is the root of S_1. By Lemma 3.30 we then have that $\omega(\sigma) < d^+(\sigma) \leq 2^{\omega(\sigma)-1}$ and so $\omega(\sigma) \geq 3$. Without loss of generality assume that $\sigma_1 = \sigma_2 = *$. By Lemma 3.29 we have $x_\tau = 1$ for all $\tau \in L(x)$ and so

$$\|x\| = |L(x)| + \sum_{\tau \in R(x)} x_\tau. \tag{3.15}$$

We proceed by modifying x, being careful to stay within the set X. Take $\pi \in Q(\sigma; \{1, 2\})$ and note that $\omega(\pi) = 2$. Consider now the element $x' \in \mathbb{R}^*$ defined as follows. Let $x'_\pi = x_\sigma$, $x'_\tau = 1$ for all $\tau \in Q(\sigma; \{1\})$, $x'_\tau = x_\tau$ for all $\tau \in S_2 \cup \ldots \cup S_q$ and $x'_\tau = 0$ otherwise. We now check that $x' \in X$. Clearly $x'_\tau \leq 1$ whenever $\omega(\tau) = 1$. Note that $supp(x') = \{\pi\} \cup Q(\sigma; \{1\}) \cup S_2 \cup \ldots \cup S_q$. Now, if $\tau \in \{\pi\} \cup Q(\sigma; \{1\})$ we have $Q(\tau) \subseteq Q(\sigma)$ and since $\sigma \in S_1$, we know that σ, and hence also τ, is distinguishable from each element of $S_2 \cup \ldots \cup S_k$. Since $\tau_1 = *$ for each $\tau \in \{\pi\} \cup Q(\sigma; \{1\})$ we see that $\{\pi\} \cup Q(\sigma; \{1\})$ contains no incompatible pairs. It follows that x' has compatible support. Finally, note that by a calculation similar to that in the proof of Lemma 3.30 we have

$$\begin{aligned} F(x') &= x_\sigma^2 + 2x_\sigma + \sum_{\tau \in R(x)\setminus\{\sigma\}} \left(x_\tau^2 + \left(2d^+(\tau) - \omega(\tau)\right) x_\tau\right) \\ &= F(x) - (2d^+(\sigma) - \omega(\sigma) - 2)x_\sigma. \end{aligned} \tag{3.16}$$

Recalling that $d^+(\sigma) > \omega(\sigma)$ we have $2d^+(\sigma) - \omega(\sigma) > \omega(\sigma) \geq 3$. Since $\sigma \in$ *supp*(x), we also have $x_\sigma > 0$ and so (3.16) implies that $F(x') < F(x) = 0$. Thus, we do indeed have $x' \in X$. Note that

$$\|x'\| = |L(x)| - d^+(\sigma) + |Q(\sigma;\{1\})| + \sum_{\tau \in R(x)} x_\tau, \tag{3.17}$$

and observe that $d^+(\sigma) \leq 2^{\omega(\sigma)-1} = |Q(\sigma;\{1\})|$ by Lemma 3.30. It now follows from (3.15) and (3.17) that $\|x'\| \geq \|x\|$, and so x' is an optimal point of X. However, we have shown that $F(x') < 0$ contradicting Lemma 3.22.

□

Gathering all the information we have obtained on compressed optimal points, we show that a proof of the following proposition is almost enough to deduce Proposition 3.18. Let us remind ourselves that in the definition of a distinguishable set (Definition 3.15), we require all elements of the set to have weight at least 1.

Proposition 3.34. *Let $\mathcal{D} \subseteq \{0,1,*\}^k$ be a distinguishable set and let $\Omega = \{d_\tau : \tau \in \mathcal{D}\}$ be a set of integers satisfying $0 \leq d_\tau \leq \omega(\tau)$ for all $\tau \in \mathcal{D}$, and $d_\tau = 0$ whenever $\omega(\tau) = 1$. Suppose that $x \in \mathbb{R}^*$ is a vector with supp$(x) = \mathcal{D}$ satisfying*

1. $$\sum_{\tau \in \mathcal{D}} \left(x_\tau^2 + (2d_\tau - \omega(\tau))x_\tau\right) = 0,$$

2. *$x_\tau \leq 1$ whenever $\omega(\tau) = 1$.*

Then

$$\sum_{\tau \in \mathcal{D}} x_\tau \leq 2^{k-1} - \sum_{\tau \in \mathcal{D}} d_\tau.$$

Furthermore we have equality only if $x \in O$ and $\Omega = \{0\}$.

A proof of Proposition 3.34 will be the focus of the next section, for now we note that it has the following corollary

Corollary 3.35. *O is the set of compressed optimal points of X. In particular, if $x \in X$ then $\|x\| \leq 2^{k-1}$.*

Proof. Let $x \in X$ be a compressed optimal point. Note that $R(x)$ is an independent set in the digraph $D(x)$ and hence by Corollary 3.25, $R(x)$ is a distinguishable set. Set $\Omega = \{d^+(\tau) : \tau \in R(x)\}$ and let x' be the element of $\mathbb{R}^*$ supported on $R(x)$ such that $x'_\tau = x_\tau$ for all $\tau \in R(x)$. Note that by Lemmas 3.22, 3.30, 3.33 and by the definition of the set X, we have that Ω and x' satisfy the conditions in the statement of Proposition 3.34. Assuming Proposition 3.34, it therefore follows that

$$\sum_{\tau \in R(x)} x_\tau \leq 2^{k-1} - \sum_{\tau \in R(x)} d^+(\tau), \tag{3.18}$$

with equality only if $x' \in O$ and $d^+(\tau) = 0$ for all $\tau \in R(x)$. The latter condition implies that $x' = x$ and so we have equality in (3.18) only if $x \in O$. By Lemma 3.29,

$$\|x\| = \sum_{\tau \in R(x)} x_\tau + \sum_{\tau \in R(x)} d^+(\tau),$$

and so it follows that $\|x\| \leq 2^{k-1}$ with equality only if $x \in O$. The result follows by noting that for all $z \in O$, $\|z\| = 2^{k-1}$ and z is compressed.

□

It is now clear that Proposition 3.18 would follow if we could also prove the following.

Proposition 3.36. *If $x \in X$ is an optimal point, then x is compressed.*

We prove Proposition 3.36 in Section 3.6.

3.5 Constrained Optimisation and a Proof of Proposition 3.34

In this section we prove Proposition 3.34 thus finalising the main stepping stone toward a proof of Proposition 3.18. We exploit the convexity of the spherical constraint found in the previous section by using the Karush-Kuhn-Tucker framework with Slater's condition (Theorem 1.5). This will lead us to consider the possible distributions of weights in distinguishable sets which we optimise over in a separate argument.

In view of the statement of Proposition 3.34 it is natural to apply Theorem 1.5 to establish the following.

Lemma 3.37. *Let $\alpha_1, \ldots, \alpha_m$ be integers where $\alpha_i = 1$ for $i = 1, \ldots, \ell$, $(0 \le \ell \le m)$ and consider the following optimisation problem for $x \in \mathbb{R}^m$.*

$$\begin{aligned} &\textit{maximise} && \sum_{i=1}^{m} x_i \\ &\textit{subject to} && \sum_{i=1}^{m} (x_i^2 - \alpha_i x_i) \le 0\,, \\ & && x_i \le 1, \quad i = 1, \ldots, \ell\,. \end{aligned} \tag{3.19}$$

If $\sum_{i=1}^{m} \alpha_i^2 > m$, then we have the unique optimal point $x^* = (x_1, \ldots, x_m)$ where $x_i = 1$ for $i \le \ell$ and $x_i = \frac{1}{2}\left(\alpha_i + \sqrt{\frac{1}{(m-\ell)} \sum_{i=\ell+1}^{m} \alpha_i^2}\right)$ for $i > \ell$.

If instead $\sum_{i=1}^{m} \alpha_i^2 \le m$, then we have the unique optimal point x^* where $x_i = \frac{1}{2}\left(\alpha_i + \sqrt{\frac{1}{m} \sum_{i=1}^{m} \alpha_i^2}\right)$ for all i.

Proof. Note first that if $\alpha_i = 0$ for all i (so in particular $\ell = 0$) then constraint (3.19) implies that $x_i = 0$ for all i in which case there's nothing to prove. Suppose then that this is not the case and define functions $f, g_1, \ldots, g_{\ell+1} : \mathbb{R}^m \to \mathbb{R}$ as follows. Let $f(x) = \sum_{i=1}^{m} x_i$, $g_i(x) = x_i - 1$ for $i = 1, \ldots, \ell$ and $g_{\ell+1}(x) = \sum_{i=1}^{m}(x_i^2 - \alpha_i x_i)$. Note that the functions just defined are all convex and differentiable. Let $x_0 = (\alpha_1/2, \ldots, \alpha_m/2)$, the centre of the spherical region described by (3.19) and observe that $g_i(x_0) < 0$ for $i = 1, \ldots, \ell + 1$. Let $S = \{x \in \mathbb{R}^m : g_i(x) \le 0 \text{ for } i = 1, \ldots, \ell + 1\}$. Since S

is compact and f is continuous there exists an optimal point x^* of our optimisation problem. Let $x^* = (x_1, \ldots, x_m)$. By Theorem 1.5 (with Slater's condition), there exist real numbers $\lambda_1, \ldots, \lambda_\ell$ and Λ such that the following hold (for notational convenience we define $\lambda_j = 0$ for $j > \ell$):

(1) $\Lambda(2x_i - \alpha_i) + \lambda_i = 1$ for all i,

(2) $\Lambda \geq 0$ and $\lambda_i \geq 0$ for all i,

(3) $\Lambda \left(\sum_{i=1}^m (x_i^2 - \alpha_i x_i)\right) = 0$ and $\lambda_i(x_i - 1) = 0$ for all i.

We consider three cases depending on the value of Λ. First let us suppose that $\Lambda = 0$. In this case, by (1) we must have $\lambda_i = 1$ for all i. Recalling that $\lambda_j = 0$ for $j > \ell$ by definition, we must also have $\ell = m$ and so $\alpha_i = 1$ for all i. Moreover, it follows from (3) that $x_i = 1$ for all i and so we're done.

By (2), we may now assume that $\Lambda > 0$ and so we may rewrite (1) as

$$x_i = \frac{1}{2}\left(\frac{1 - \lambda_i}{\Lambda} + \alpha_i\right) \text{ for all } i. \tag{3.20}$$

Moreover, $\sum_{i=1}^m (x_i^2 - \alpha_i x_i) = 0$ by (3) which by (3.20) gives

$$\frac{1}{\Lambda^2} \sum_{i=1}^m (1 - \lambda_i)^2 = \sum_{i=1}^m \alpha_i^2. \tag{3.21}$$

Now, note that for $i \leq \ell$ we have $\alpha_i = 1$ and $x_i \leq 1$ and so by (3.20) we have

$$1 - \Lambda \leq \lambda_i \text{ for } i \leq \ell. \tag{3.22}$$

If $\Lambda < 1$ then by (3.22) we have $\lambda_i > 0$ for $i \leq \ell$ and so by (3), $x_i = 1$ for $i \leq \ell$ and so in fact by (3.20)

$$1 - \Lambda = \lambda_i \text{ for } i \leq \ell.$$

Recalling that $\alpha_i = 1$ for $i \leq \ell$ and $\lambda_i = 0$ for $i > \ell$ by definition, (3.21) then gives

$$\frac{1}{\Lambda} = \sqrt{\frac{1}{m - \ell} \sum_{i=\ell+1}^m \alpha_i^2}. \tag{3.23}$$

From (3.20) it now follows that

$$x_i = \frac{1}{2}\left(\alpha_i + \sqrt{\frac{1}{m-\ell}\sum_{i=\ell+1}^{m}\alpha_i^2}\right) \text{ for } i > \ell,$$

Recalling that $\Lambda < 1$, it follows from (3.23) that $\sum_{i=1}^{m}\alpha_i^2 > m$.

It remains to consider the case where $\Lambda \geq 1$. Recall that if $\lambda_i > 0$ for some i then $x_i = 1$ by (3) and so $\Lambda = 1 - \lambda_i$ by (3.20). However, this contradicts the assumption that $\Lambda \geq 1$ and so we conclude that $\lambda_i = 0$ for all i. It follows from (3.21) that

$$\frac{1}{\Lambda} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\alpha_i^2}, \tag{3.24}$$

so that by (3.20),

$$x_i = \frac{1}{2}\left(\alpha_i + \sqrt{\frac{1}{m}\sum_{i=1}^{m}\alpha_i^2}\right) \text{ for all } i.$$

The result follows, noting that by (3.24) we have $\sum_{i=1}^{m}\alpha_i^2 \leq m$ in this case. □

We are almost ready to prove Proposition 3.34, but first we need the following inequality.

Lemma 3.38. *Let $\alpha_1, \ldots, \alpha_m$ be integers ≥ 2 then*

$$\sum_{i=1}^{m}\alpha_i + \sqrt{m\sum_{i=1}^{m}\alpha_i^2} \leq \sum_{i=1}^{m}2^{\alpha_i},$$

and equality holds if only if $\alpha_i = 2$ for all i.

Proof. Let $\alpha = (\alpha_1, \ldots, \alpha_m)$. We induct on the value of $S_\alpha := \sum_{i=1}^{m}(2^{\alpha_i - 2} - 1)$. If $S_\alpha = 0$, then $\alpha_i = 2$ for all i, so that

$$\sum_{i=1}^{m}\alpha_i + \sqrt{m\sum_{i=1}^{m}\alpha_i^2} = 4m = \sum_{i=1}^{m}2^{\alpha_i}.$$

Suppose then that $S_\alpha > 0$ so that $\alpha_j \geq 3$ for some $j \in [m]$. Without loss of generality assume that $j = 1$. Define a new sequence of integers

$\alpha' = (\alpha'_1, \ldots, \alpha'_{m+1})$, as follows: Let $\alpha'_1 = \alpha'_2 = \alpha_1 - 1$ and $\alpha'_i = \alpha_{i-1}$ for $i = 3, 4, \ldots, m+1$. Note that $\alpha'_i \geq 2$ for all i and $S_{\alpha'} = S_\alpha - 1$ and so by the inductive hypothesis

$$\sum_{i=1}^{m+1} \alpha'_i + \sqrt{(m+1)\sum_{i=1}^{m+1} \alpha'^2_i} \leq \sum_{i=1}^{m+1} 2^{\alpha'_i}, \tag{3.25}$$

Note that

$$\sum_{i=1}^{m+1} 2^{\alpha'_i} = \sum_{i=1}^{m} 2^{\alpha_i} \text{ and } \sum_{i=1}^{m+1} \alpha'_i - \sum_{i=1}^{m} \alpha_i = \alpha_1 - 2 > 0, \tag{3.26}$$

and also

$$\begin{aligned} (m+1)\sum_{i=1}^{m+1} \alpha'^2_i - m\sum_{i=1}^{m} \alpha_i^2 &= \sum_{i=1}^{m} \alpha_i^2 + (m+1)(\alpha_1^2 - 4\alpha_1 + 2) \\ &\geq \sum_{i=1}^{m} \alpha_i^2 - (m+1), \end{aligned} \tag{3.27}$$

where in the last inequality we used the fact that $\alpha^2 - 4\alpha + 2 \geq -1$ for $\alpha \geq 3$. Note that since $\alpha_i \geq 2$ for all i, we certainly have that $\sum_{i=1}^m \alpha_i^2 > m+1$. It follows then from (3.27) that

$$(m+1)\sum_{i=1}^{m+1} \alpha'^2_i > m\sum_{i=1}^{m} \alpha_i^2. \tag{3.28}$$

Combining (3.25), (3.26), and (3.28) we have

$$\sum_{i=1}^{m} \alpha_i + \sqrt{m\sum_{i=1}^{m} \alpha_i^2} < \sum_{i=1}^{m+1} \alpha'_i + \sqrt{(m+1)\sum_{i=1}^{m+1} \alpha'^2_i} \leq \sum_{i=1}^{m+1} 2^{\alpha'_i} = \sum_{i=1}^{m} 2^{\alpha_i}$$

as required. Note the strict inequality, and so we only have equality in the case where $\alpha_i = 2$ for all i. □

Proof of Proposition 3.34. Consider first the case where $\sum_{\tau\in\mathcal{D}}(\omega(\tau) - 2d_\tau)^2 > |\mathcal{D}|$. Suppose that ℓ elements of $\mathcal{D}$ have weight 1 and let $\mathcal{D}' = \{\tau \in \mathcal{D} : \omega(\tau) \geq 2\}$. Note that by Lemma 3.17 we have $\sum_{\tau\in\mathcal{D}} 2^{\omega(\tau)} \leq 2^k$ and hence

$$\sum_{\tau\in\mathcal{D}'} 2^{\omega(\tau)} \leq 2^k - 2\ell. \tag{3.29}$$

Applying Lemma 3.37, recalling that $0 \leq d_\tau \leq w(\tau)$ for all $\tau \in \mathcal{D}$, that $d_\tau = 0$ whenever $\omega(\tau) = 1$ and using (3.29) and Lemma 3.38 we have

$$\sum_{\tau\in\mathcal{D}} x_\tau \leq \ell + \frac{1}{2}\left(\sum_{\tau\in\mathcal{D}'}(\omega(\tau) - 2d_\tau) + \sqrt{|\mathcal{D}'|\sum_{\tau\in\mathcal{D}'}(\omega(\tau) - 2d_\tau)^2}\right) \tag{3.30}$$

$$\leq \ell + \frac{1}{2}\left(\sum_{\tau\in\mathcal{D}'}\omega(\tau) + \sqrt{|\mathcal{D}'|\sum_{\tau\in\mathcal{D}'}\omega(\tau)^2}\right) - \sum_{\tau\in\mathcal{D}} d_\tau \tag{3.31}$$

$$\leq \ell + \sum_{\tau\in\mathcal{D}'} 2^{\omega(\tau)-1} - \sum_{\tau\in\mathcal{D}} d_\tau \tag{3.32}$$

$$\leq 2^{k-1} - \sum_{\tau\in\mathcal{D}} d_\tau. \tag{3.33}$$

We analyse the conditions for equality to hold. For equality to hold in (3.31) it must be the case that for all $\tau \in \mathcal{D}$, either $d_\tau = 0$ or $d_\tau = \omega(\tau)$. By Lemma 3.38, for equality to hold in (3.32) it must be the case that $\omega(\tau) = 2$ for all $\tau \in \mathcal{D}'$. It now follows from Lemma 3.37 that for equality to also hold in (3.30), we must have $x_\tau = 1$ whenever $\omega(\tau) = 1$, $x_\tau = 2$ for all $\tau \in \mathcal{D}'$ such that $d_\tau = 0$ and $x_\tau = 0$ for all $\tau \in \mathcal{D}'$ such that $d_\tau = \omega(\tau)$. However, since each x_τ is non-zero by assumption we conclude that $d_\tau = 0$ for all $\tau \in \mathcal{D}$ i.e. $\Omega = \{0\}$. Finally, for equality to hold in (3.33) we must have equality in (3.29) and so $\mathcal{D}$ is a decomposition by Lemma 3.17. It follows that $x \in O$.

It remains to consider the case where $\sum_{\tau\in\mathcal{D}}(\omega(\tau) - 2d_\tau)^2 \leq |\mathcal{D}|$. By Lemma 3.37 and Lemma 3.17 we then have

$$\sum_{\tau\in\mathcal{D}} x_\tau \leq \frac{1}{2}\left(|\mathcal{D}| + \sum_{\tau\in\mathcal{D}}\omega(\tau)\right) - \sum_{\tau\in\mathcal{D}} d_\tau$$

$$\leq \frac{1}{2}\left(|\mathcal{D}| + \sum_{\tau\in\mathcal{D}} 2^{\omega(\tau)-1}\right) - \sum_{\tau\in\mathcal{D}} d_\tau$$

$$\leq 2^{k-1} - \sum_{\tau\in\mathcal{D}} d_\tau. \tag{3.34}$$

For equality to hold in (3.34), we must have that $|\mathcal{D}| = 2^{k-1}$ and so $\mathcal{D}$ is a decomposition consisting only of elements of weight 1. It follows that $d_\tau = 0$ and $x_\tau \leq 1$ for all $\tau \in \mathcal{D}$. If equality holds throughout the above, we then have that $x_\tau = 1$ for all $\tau \in \mathcal{D}$ and so $x \in O$. $\square$

3.6 Towards an Exact Result: Analytic and Combinatorial Stability

In this section we prove Proposition 3.36 thus concluding our proof of Proposition 3.18. Note that Proposition 3.18 classifies the optimal points of X. We use compactness arguments to prove a result to the effect that 'almost optimal' points of X must be close (in ℓ_1 norm) to a genuine optimal point of X. Furthermore, compactness allows us to derive similar properties for $X(\gamma)$ when γ is small. We then investigate what implications this has in our original combinatorial setting and complete the proof of Theorem 3.7.

Proof of Proposition 3.36. Let H be the matrix with rows and columns indexed by $\{0,1,*\}^k$ where

$$H_{\sigma\tau} = \begin{cases} 1 & \text{if } \sigma,\tau \text{ are indistinguishable,} \\ 0 & \text{if } \sigma,\tau \text{ are distinguishable.} \end{cases}$$

Note that in particular, all diagonal entries of H are equal to 1. Let $w = (-\omega(\tau) : \tau \in \{0,1,*\}^k) \in \mathbb{R}^*$, then for $x \in \mathbb{R}^*$ we may write

$$F(x) = w^T x + x^T H x.$$

Suppose now that $x \in X$ is an optimal point. By the proof of Lemma 3.27, there is a finite sequence $x = x_0, x_1, \ldots, x_m$ of distinct optimal points of X where x_m is compressed, and for $i = 0, \ldots, m-1$, $x_{i+1} = x_i(\pi_i, \rho_i)$ for some indistinguishable pair $\pi_i, \rho_i \in supp(x_i)$. Moreover we know that $\omega(\rho_i) \geq 2$ and that $\pi_i \notin supp(x_{i+1})$ for all i.

Suppose that x is not compressed so that $m \geq 1$. Let $y = x_{m-1}$, $z = x_m$ and let $\pi = \pi_{m-1}$, $\rho = \rho_{m-1}$. Since $z = y(\pi,\rho)$, it follows from the definition of compression that $z = y + \alpha(e_\rho - e_\pi)$ for some $\alpha > 0$. Let $p = e_\pi - e_\rho$. It follows, by the Taylor expansion of F, that

$$F(y) = F(z + \alpha p) = F(z) + \alpha p^T \nabla F(z) + \alpha^2 p^T H p. \tag{3.35}$$

Recall that $F(y) = F(z) = 0$ by Lemma 3.22. Furthermore by direct calculation we also have $p^T H p = 0$. It follows from (3.35) that $p^T \nabla F(z) = 0$

i.e.

$$\frac{\partial F}{\partial x_\pi}(z) = \frac{\partial F}{\partial x_\rho}(z). \tag{3.36}$$

Let I_ρ be the set of elements of $\{0, 1, *\}^k$ that are indistinguishable from ρ excluding ρ itself. Define I_π similarly. From the definition of F we have

$$\frac{\partial F}{\partial x_\rho}(z) = 2z_\rho + 2\sum_{\tau \in I_\rho} z_\tau - \omega(\rho). \tag{3.37}$$

Since z is a compressed optimal point we have $z \in O$ by Corollary 3.35. Since $\omega(\rho) \geq 2$ and $\rho \in supp(z)$ we conclude that in fact $\omega(\rho) = 2$ and so $z_\rho = 2$. Moreover since $supp(z)$ is a distinguishable set we conclude that $z_\tau = 0$ for all $\tau \in I_\rho$. It follows from (3.37) that $\frac{\partial F}{\partial x_\rho}(z) = 2$ and hence from (3.36) that

$$\frac{\partial F}{\partial x_\pi}(z) = 2z_\pi + 2\sum_{\tau \in I_\pi} z_\tau - \omega(\pi) = 2. \tag{3.38}$$

Since $z \in O$ we know that for all $\tau \in supp(z)$, $\omega(\tau) = 1$ or 2 and $z_\tau = \omega(\tau)$. Let w_1, w_2 be the number of elements of $I_\pi \cap supp(z)$ with weights 1, 2 respectively. Since $\pi \notin supp(z)$, we can then infer from (3.38) that

$$2w_1 + 4w_2 - \omega(\pi) = 2. \tag{3.39}$$

We also know that $supp(z)$ is a decomposition and so

$$\begin{aligned} Q(\pi) &= \bigcup_{\tau \in supp(z)} (Q(\tau) \cap Q(\pi)) \\ &= \bigcup_{\tau \in I_\pi \cap supp(z)} (Q(\tau) \cap Q(\pi)) \\ &\subseteq \bigcup_{\tau \in I_\pi \cap supp(z)} Q(\tau). \end{aligned} \tag{3.40}$$

The second equality comes from the fact that $Q(\tau) \cap Q(\pi) = \emptyset$ whenever τ and π are distinguishable. Comparing the cardinality of the sets in (3.40) yields

$$2^{\omega(\pi)} \leq \sum_{\tau \in I_\pi \cap supp(z)} 2^{\omega(\tau)} = 2w_1 + 4w_2. \tag{3.41}$$

Note also that $\rho \in I_\pi \cap supp(z)$ and $\omega(\rho) = 2$ so that $w_2 \geq 1$. Using (3.39), this last observation implies that $\omega(\pi) \geq 2$ whereas combining (3.39) and (3.41) we have

$$2^{\omega(\pi)} - \omega(\pi) \leq 2w_1 + 4w_2 - \omega(\pi) = 2 \tag{3.42}$$

We deduce that $\omega(\pi) = 2$ and so we have equality throughout (3.42), in particular we have equality in (3.41) and so also in (3.40). Note that $|Q(\pi)| = |Q(\rho)|$ since $\omega(\pi) = \omega(\rho) = 2$. Recalling that $\rho \in I_\pi \cap supp(z)$ equality in (3.40) would therefore imply that $Q(\pi) = Q(\rho)$ i.e. $\pi = \rho$. This is a contradiction and so x must be compressed. $\square$

Proposition 3.18 has the following corollary that says an almost optimal point of X must be close in norm to an actual optimal point of X.

Proposition 3.39. *Let $\eta \ll \varepsilon$. If $x \in X$ satisfies $\|x\| > 2^{k-1} - \eta$, then there exists an $x^* \in O$ such that $\|x - x^*\| < \varepsilon$.*

Proof. Consider the set

$$\tilde{X} := X \setminus \bigcup_{x^* \in O} B_\varepsilon(x^*).$$

$\tilde{X}$ is compact and so $\sup_{z \in \tilde{X}} \|z\| = \|\tilde{x}\|$ for some $\tilde{x} \in \tilde{X}$. By the definition of $\tilde{X}$, $\tilde{x} \notin O$ and so by Proposition 3.18, $\|\tilde{x}\| = 2^{k-1} - \eta$ for some $\eta > 0$. It follows that if $x \in X$ satisfies $\|x\| > 2^{k-1} - \eta$ then $x \notin \tilde{X}$ and so $x \in B_\varepsilon(x^*)$ for some $x^* \in O$. $\square$

The following lemma allows us to relate properties of X and $X(\gamma)$ for γ small.

Lemma 3.40. *Let $\gamma \ll \eta$. If $x \in X(\gamma)$, then there exists $x_0 \in X$ for which $\|x - x_0\| < \eta$.*

Proof. Let $(\gamma_i)_{i \in \mathbb{N}}$ be a strictly decreasing sequence tending to 0, and let $X_i = X(\gamma_i)$ for $i \in \mathbb{N}$. Then $X_1, X_2, \ldots$ is a decreasing sequence of compact sets i.e. $X_{i+1} \subseteq X_i$ for $i \in \mathbb{N}$. Consider the set

$$U = \bigcup_{z \in X} B_\eta(z),$$

an open set containing X. Note that $(X_i \setminus U)_{i \in \mathbb{N}}$ is also a decreasing sequence of compact sets and that

$$\bigcap_{i=1}^{\infty} (X_i \setminus U) = \left(\bigcap_{i=1}^{\infty} X_i \right) \Big\backslash U = X \setminus U = \emptyset.$$

By Cantor's Intersection Theorem (see [76, Theorem 2.36, p.38]) it follows that $X_m \backslash U = \emptyset$ for some $m \in \mathbb{N}$. In other words, if $x \in X(\gamma_m)$ then $x \in U$ so that $x \in B_\eta(x_0)$ for some $x_0 \in X$. The result follows by taking $\gamma \leq \gamma_m$. $\square$

Corollary 3.41. *Let $\gamma \ll \varepsilon$. If $x \in X(\gamma)$ satisfies $\|x\| = 2^{k-1}$, then there exists an $x^* \in O$ such that $\|x - x^*\| < \varepsilon$.*

Proof. Given $\varepsilon > 0$, let $\eta = \min\{\eta_{3.39}(\varepsilon/2), \varepsilon/2\}$ and suppose that $\gamma \leq \gamma_{3.40}(\eta)$. Suppose that $x \in X(\gamma)$ satisfies $\|x\| = 2^{k-1}$. By Lemma 3.40 there exists an $x_0 \in X$ such that $\|x_0 - x\| < \eta$ and so $\|x_0\| > \|x\| - \eta = 2^{k-1} - \eta$. It follows from Proposition 3.39 that there exists an $x^* \in O$ such that $\|x_0 - x^*\| < \varepsilon/2$ and so

$$\|x - x^*\| \leq \|x - x_0\| + \|x_0 - x^*\| < \eta + \varepsilon/2 \leq \varepsilon.$$

$\square$

Let

$$O^* = \{x \in O : \omega(\tau) = 1 \text{ for all } \tau \in supp(x)\}.$$

In words, O^* is the set of all elements $x \in \mathbb{R}^*$ such that x is supported on a perfect matching of Q_k and all non-zero entries of x are equal to 1. We can also view O^* as the set of profiles of hypercube colourings normalised by clique size. Our aim is to use the stability-type statement of Corollary 3.41 to prove Theorem 3.7 in the following form.

Theorem 3.42. *Let $\frac{1}{n} \ll \delta \ll \varepsilon \ll 1$. If G is a $(1-\delta)$-dense, k-coloured graph with $v(G) = 2^{k-1}n$, containing no monochromatic odd connected matching of order $\geq (1+\delta)n$, then for any choice of profile $x(G)$ of G, there exists some $x^* \in O^*$ such that*

$$\|x(G)/n - x^*\| < \varepsilon.$$

First we need the following two colour Ramsey type result which is a direct consequence of the more general Theorem 1.8 in [4].

Lemma 3.43. *Let $\frac{1}{n} \ll \delta \ll \varepsilon$. If H is a $(1-\delta)$-dense, 2-coloured graph with $v(H) \geq (\frac{3}{2}+\varepsilon)n$, then H contains a monochromatic connected matching of order $\geq (1+\delta)n$.* $\square$

Proof of Theorem 3.42. Given $\varepsilon > 0$, let $\gamma = \gamma_{3.41}(\varepsilon)$ and $\delta' = \delta_{3.43}(\varepsilon)$. Suppose that $\delta < \min\{\gamma^2 k^{-2} 2^{-4k}, \delta' 2^{-2k}\}$ and that $n \geq \max\{n_{3.43}(\delta'), \delta^{-1}\}$. Let G be a k-coloured graph as in the statement of Theorem 3.42. Let $x(G)$ be any choice of profile for G and let the corresponding profile partition be $(V_\tau : \tau \in \{0,1,*\}^k)$. Note that $\|x(G)/n\| = 2^{k-1}$ and by Proposition 3.14, we have that $x(G)/n \in X(\sqrt{\delta} k 2^{2k}) \subseteq X(\gamma)$. By Corollary 3.41 there exists an element $x^* \in O$ such that

$$\|x(G)/n - x^*\| < \varepsilon. \tag{3.43}$$

Suppose that $x^* \in O \backslash O^*$, then $x_\tau = 2$ for some $\tau \in \{0,1,*\}^k$ such that $\omega(\tau) = 2$. It follows from (3.43) that $x(G)_\tau = |V_\tau| > (2 - \varepsilon)n \geq (3/2 + \varepsilon)n$. Let $H = G[V_\tau]$. By the definition of V_τ, H is a 2-coloured graph. Moreover since G has at most $\delta\binom{v(G)}{2} \leq \delta'\binom{v(H)}{2}$ edges missing, the same is true for H. It follows by Lemma 3.43 that H contains a monochromatic connected matching of order $\geq (1+\delta')n > (1+\delta)n$. However, by the definition of $V_\tau = V(H)$, any monochromatic component of H is contained in a non-bipartite monochromatic component of G. Thus, G contains a monochromatic odd connected matching of order $> (1+\delta)n$ contrary to assumption. We conclude that $x^* \in O^*$. $\square$

3.7 The Regularity Method

In this section we discuss the tools and results we need from the regularity method. Our starting point is Szemerédi's Regularity Lemma [83] which we discuss briefly now.

Let G be a graph and let A, B be disjoint subsets of $V(G)$. We call

$$d_G(A,B) := \frac{e_G(A,B)}{|A||B|}$$

the *density* of the pair (A, B). For $\delta > 0$, we say that the pair (A, B) is δ-regular with respect to G if, for every $A' \subseteq A$ and $B' \subseteq B$ satisfying $|A'| \geq \delta|A|$ and $|B'| \geq \delta|B|$, we have

$$|d_G(A',B') - d_G(A,B)| < \delta.$$

If, for $d > 0$, we also have that $|N_G(a) \cap B| \geq d|B|$ for all $a \in A$ and $|N_G(b) \cap A| \geq d|A|$ for all $b \in B$, then we say that (A, B) is (δ, d)-super-regular with respect to G. We may omit the subscripts from the above notation if the graph G is clear from the context. The following is a version of Szemerédi's Regularity Lemma that appears as Theorem 1.18 in [62].

Theorem 3.44 (Multicolour Regularity Lemma)**.** *For all $\delta > 0$ and $k, \ell \in \mathbb{N}$ there exists $L = L(\delta, k, \ell)$ and $M = M(\delta, k, \ell)$ such that the following holds. For all k-coloured graphs G on at least M vertices, $V(G)$ may be partitioned into sets $V_0, V_1 \dots, V_t$ such that*

- *$\ell \leq t \leq L$;*
- *$|V_0| < \delta v(G)$ and $|V_1| = |V_2| = \dots = |V_t|$;*
- *apart from at most $\delta\binom{t}{2}$ exceptional pairs, the pairs (V_i, V_j), $1 \leq i < j \leq t$, are δ-regular with respect to G_s for $s = 1, \dots, k$.*

We now state some technical lemmas related to Łuczak's method of connected matchings. First we need a definition. (It might be useful at this point to recall Definition 3.5.)

Definition 3.45. *Let $\delta, d \in [0, 1]$ and $q, m \geq 1$ be integers.*

- *Let F be a graph on vertex set $[q]$ and let $U_1, \dots, U_q$ be disjoint sets of size m. We call a graph H on vertex set $\bigcup_{i\in[q]} U_i$ a* (δ, m)-regular blow-up *of F if whenever $\{i, j\} \in E(F)$, we have that (U_i, U_j) is a δ-regular pair.*
- *If in addition to the above, $d(U_i, U_j) \geq d$ for each edge $\{i, j\}$ of F then we say that H has* minimum density d*.*
- *Suppose that F is a connected matching and H is a (δ, m)-regular blow-up of F with minimum density d. If for each matching edge $\{i, j\}$ of F, the pair (U_i, U_j) is in fact (δ, d)-super-regular in H, then we say that H is a* (δ, d, m)-super-regular blow-up *of F.*

Versions of the following two lemmas abound in the literature (e.g. [61], [65]), but here we give statements tailored to our needs. However, since they are not new, we defer their proofs to the Appendix (Section B).

Lemma 3.46. *Let $q \geq 4$ and suppose that $\frac{1}{m} \ll \delta \ll d$. Let F be a connected matching of order q such that every vertex of F is incident to a matching edge and let H be a (δ, d, m)-super-regular blow-up of F. Then the following holds:*

If $i, j \in V(F)$ and there is an ij-path of length r in F, then for every pair of vertices $u \in U_i$, $w \in U_j$, there exists a uw-path of length ℓ in H for each $3q \leq \ell \leq (1-6\delta)qm$ such that $\ell \equiv r \pmod 2$.

Lemma 3.47. *Let $q \geq 4$ and let $\frac{1}{m} \ll \delta \ll d$. Let F be an odd connected matching of order q and suppose that H is a (δ, m)-regular blow-up of F with minimum density d. Then H contains a cycle of length ℓ for each odd $3q \leq \ell \leq (1-6\delta)qm$.*

We borrow the following fact.

Fact 3.48. *([50, Lemma 9]). Let H be a $(1-\delta)$-dense graph on t vertices. Then H has a subgraph H' such that $v(H') \geq (1-\sqrt{\delta})t$ and $\delta(H') \geq (1 - 2\sqrt{\delta})t$.*

We will also need the following two standard facts whose proofs we omit it here.

Fact 3.49. *Let $0 < \delta \leq 1/2$ and let (A, B) be a δ-regular pair with density d. Suppose that $A' \subseteq A$, $B' \subseteq B$ such that $|A'| \geq (1-\delta)|A|$, $|B'| \geq (1-\delta)|B|$. Then (A', B') is 2δ-regular with density $d' > d - \delta$. Moreover, if (A, B) is in fact (δ, β)-super-regular for some $\beta > 0$, then (A', B') is $(2\delta, \beta - \delta)$-super-regular.* $\square$

Fact 3.50. *Let $0 < \delta \leq 1/2$ and let (A, B) be a δ-regular pair with density d. Then there exist $A' \subseteq A$, $B' \subseteq B$ such that $|A'| = (1-\delta)|A|$, $|B'| = (1-\delta)|B|$ and (A', B') is $(2\delta, d - 2\delta)$-super-regular.* $\square$

3.8 Proof of Main Theorem

In this final section we prove our main result Theorem 3.4, and therefore also Theorem 3.2. The idea is to invoke Theorem 3.42 to show that the profile of a certain reduced graph is close in ℓ_1-norm to the profile of a hypercube colouring. We then translate this information to show that the original graph is close in edit distance to a hypercube colouring.

The stability-type methods of this section require some care due to the plethora of extremal constructions. Luckily hypercube colourings share enough common features for these methods to be viable and surprisingly we require no case analysis.

First let us state a result which is a corollary of a classical theorem of Bondy [9].

Theorem 3.51. *Let G be a graph on at least 3 vertices with minimum degree $> v(G)/2$, then G is pancyclic i.e. G contains cycles of all lengths $3 \leq \ell \leq v(G)$.* □

Proof of Theorem 3.4. Let $0 < \varepsilon < 2^{-4k}$, let

$$\eta < \delta \leq \min \left\{ \frac{1}{9}\delta_{3.42}^2(\varepsilon), \delta_{3.46}\left(\frac{1}{k}\right), \delta_{3.47}\left(\frac{1}{k}\right) \right\}, \tag{3.44}$$

let $n_0 \geq \max\{n_{3.42}(\delta), \delta^{-1/2}\}$ and let $L = L_{3.44}(\delta, k, 2^k n_0)$. Let n be odd with

$$n \geq \max\{Lm_{3.46}(\delta), Lm_{3.47}(\delta), M_{3.44}(\delta, k, 2^k n_0)\}. \tag{3.45}$$

Finally let G be a k-coloured copy of K_N where $N > (2^{k-1} - \eta)n$ and assume that

$$G \text{ contains no monochromatic copy of } C_n. \tag{†}$$

Applying Theorem 3.44 to G we obtain a partition of $V(G)$ into sets $V_0, \ldots, V_{t_0}$ such that

(i) $2^k n_0 \leq t_0 \leq L$;

(ii) $|V_0| < \delta N$ and $|V_1| = |V_2| = \ldots = |V_{t_0}|$;

(iii) apart from at most $\delta\binom{t_0}{2}$ exceptional pairs, the pairs (V_i, V_j), $1 \leq i < j \leq t_0$, are δ-regular with respect to G_s for $s = 1, \ldots, k$.

It follows that for $i \in [t_0]$,

$$m := |V_i| \geq \frac{(1-\delta)N}{t_0}. \tag{3.46}$$

We construct a *reduced graph* R_0 with vertex set $\{1, ..., t_0\}$ and edge set formed by pairs $\{u, w\}$ for which (V_u, V_w) is δ-regular with respect to G_i for $i = 1, \ldots, k$. It follows from (iii) of the above that R_0 is $(1-\delta)$-dense. Fact 3.48 allows us to find a subgraph $R \subseteq R_0$ satisfying $v(R) \geq (1-\sqrt{\delta})t_0$ and $\delta(R) \geq (1-2\sqrt{\delta})t_0$. Let $t = v(R)$ and assume without loss of generality that $V(R) = \{1, \ldots, t\}$. We k-colour R by colouring an edge $\{u, w\}$ with the least colour i for which

$$d_{G_i}(V_u, V_w) \geq \frac{1}{k}. \tag{3.47}$$

Let $t' = t/2^{k-1}$ and note that by (3.46) and the definition of t,

$$mt' \geq (1-2\sqrt{\delta})n. \tag{3.48}$$

Suppose that R contains a monochromatic odd connected matching F of order $q \geq (1+3\sqrt{\delta})t'$. Then G contains a monochromatic (δ, m)-regular blow-up of F with minimum density d for some $d \geq 1/k$ by (3.47). Note that since $t_0 \leq L$ we have $m \geq m_{3.47}(\delta)$ by (3.46) and (3.45). It follows from Lemma 3.47 that G contains a monochromatic copy of C_n since n is odd and

$$3q \leq 3L \leq n \leq (1-6\delta)(1+3\sqrt{\delta})mt' \leq (1-6\delta)qm,$$

contradicting (†). We conclude that R contains no such odd connected matching. Let $(W_\tau : \tau \in \{0,1,*\}^k)$ be a profile partition of R and let $x(R) = (|W_\tau| : \tau \in \{0,1,*\}^k)$ be the corresponding profile. It follows by Theorem 3.42 that there exists $x^* \in O^*$ such that

$$\|x(R)/t' - x^*\| < \varepsilon. \tag{3.49}$$

This tells us a lot about the structure of R, indeed it is 'close to' a hypercube colouring. The aim is to use this fact to eventually say the same for G. By the definition of O^* we have that

$$supp(x^*) = \mathcal{M} \subseteq \{0,1,*\}^k,$$

for some perfect matching $\mathcal{M}$ of the hypercube Q_k and $x^*_\tau = 1$ for all $\tau \in \mathcal{M}$. Let

$$W = R \setminus \bigcup_{\tau \in \mathcal{M}} W_\tau.$$

We will treat W as a 'leftover set' of vertices of R and study only the structure of $R\backslash W$. Note that by (3.49) we have

$$(1-\varepsilon)t' < |W_\tau| < (1+\varepsilon)t' \text{ for all } \tau \in \mathcal{M}, \tag{3.50}$$

and so by removing at most $2\varepsilon t'$ vertices from each part W_τ, where $\tau \in \mathcal{M}$, and absorbing these removed vertices into W, we may assume that these parts W_τ all have the same size $> (1-\varepsilon)t'$. Note that even after this absorption we have

$$|W| = t - \sum_{\tau \in \mathcal{M}} |W_\tau| < t - 2^{k-1}(1-\varepsilon)t' = \varepsilon t.$$

We make a couple of observations regarding the colouring of R with respect to these vertex classes. For $j \in [k]$ we let

$$I_j = \{\tau \in \mathcal{M} : \tau_j = *\}.$$

Lemma 3.52. *Let $\tau \in I_j$. Then $R[W_\tau]$ is monochromatic in the colour j and has minimum degree at least $(1 - 2^{k+1}\sqrt{\delta})|W_\tau|$.*

Proof. By the definition of the profile partition, for each colour $i \neq j$, each pair $v, w \in W_\tau$ must lie in the same vertex class in an induced bipartite subgraph of R_i. It follows that if $\{v, w\} \in E(R)$ then it cannot receive the colour i and hence must receive colour j. Since $\delta(R) \geq (1 - 2\sqrt{\delta})t$ we have

$$\delta(R[W_\tau]) \geq |W_\tau| - 1 - 2\sqrt{\delta}t \geq (1 - 2^{k+1}\sqrt{\delta})|W_\tau|$$

where for the last inequality we used (3.50). □

Definition 3.53. *Let $\sigma, \tau \in \{0, 1, *\}^k$. We denote the set $\{i \in [k] : \{\sigma_i, \tau_i\} = \{0, 1\}\}$ by $\Delta(\sigma, \tau)$. We call $|\Delta(\sigma, \tau)|$ the* distance *between σ and τ and denote it by $d(\sigma, \tau)$.*

Lemma 3.54. *Let $\sigma, \tau \in \mathcal{M}$ be distinct, then*

(i) Each edge of $R[W_\sigma, W_\tau]$ receives a colour from the set $\Delta(\sigma, \tau)$;

(ii) $R[W_\sigma, W_\tau]$ has minimum degree $\geq (1 - 2^{k+1}\sqrt{\delta})|W_\sigma|$, in particular $R[W_\sigma, W_\tau]$ is connected and contains a perfect matching.

Proof. Let $\sigma \in I_j$, $\tau \in I_\ell$. Suppose that $j \neq \ell$. By the definition of the profile partition, for each colour $i \notin \Delta(\sigma, \tau)$, each pair $v \in W_\sigma$, $w \in W_\tau$ must lie in either the same vertex class in an induced bipartite subgraph of R_i or they lie in different connected components of R_i. It follows that if $\{v, w\} \in E(R)$ then it must receive a colour from $\Delta(\sigma, \tau)$. Similarly, if $j = \ell$ then each edge of $R[W_\sigma, W_\tau]$ must receive a colour from $\Delta(\sigma, \tau) \cup \{j\}$. However, by Lemma 3.52, $R[W_\tau]$ and $R[W_\sigma]$ are both monochromatic in the colour j and both have minimum degree at least $(1-2^{k+1}\sqrt{\delta})|W_\sigma| > |W_\sigma|/2$ (recall that $|W_\sigma| = |W_\tau|$). It follows, by Theorem 3.51 for example, that $R[W_\tau]$ and $R[W_\sigma]$ are both Hamiltonian and non-bipartite. Using (3.50), we deduce that if an edge of $R[W_\sigma, W_\tau]$ receives the colour j, then R contains a monochromatic odd connected matching in the colour j of order at least

$$|W_\sigma| + |W_\tau| - 2 \geq 2(1 - \varepsilon)t' - 2 \geq (1 + 3\sqrt{\delta})t',$$

which we showed previously was not the case. Part (i) of the lemma follows. If $v \in W_\tau$ then, since $\delta(R) \geq (1 - 2\sqrt{\delta})t$, we have

$$|N(v) \cap W_\sigma| \geq |W_\sigma| - 2\sqrt{\delta}t \geq (1 - 2^{k+1}\sqrt{\delta})|W_\sigma|.$$

Similarly if $w \in W_\sigma$, then $|N(w) \cap W_\tau| \geq (1 - 2^{k+1}\sqrt{\delta})|W_\sigma|$. Since $1 - 2^{k+1}\sqrt{\delta} > 1/2$, it follows that $R[W_\sigma, W_\tau]$ is connected and (e.g. by Hall's theorem) contains a perfect matching.

□

Let Γ denote the k-coloured multigraph on vertex set $\mathcal{M}$ where we have an edge between σ and τ in each colour j for which $R[W_\sigma, W_\tau]$ contains an edge of colour j. Note that since $\delta(R) > (1 - 2\sqrt{\delta})t$ and $|W_\sigma| = |W_\tau| > (1 - \varepsilon)t' > 2\sqrt{\delta}t$, $R[W_\sigma, W_\tau]$ always contains an edge. Let Γ^* denote the subgraph of Γ where we keep only those edges that occur as the unique edge between a given pair of vertices in Γ. Recall that for $j \in [k]$, Γ_j, Γ^*_j denote the jth colour class of Γ, Γ^* respectively.

Lemma 3.55. *For each $j \in [k]$, the vertices of Γ^*_j can be covered by a matching $T_j \subseteq \Gamma^*_j$ and the set I_j. Moreover I_j is a set of isolated vertices in Γ_j.*

Proof. Fix $j \in [k]$. If $\sigma \in I_j$ then σ is an isolated vertex in Γ_j by Lemma 3.54(i). If $\sigma \notin I_j$ then we may assume without loss of generality that $\sigma_j = 0$. Let σ' be the element of $\{0, 1, *\}^k$ such that $\sigma'_j = 1$ and $\sigma'_i = \sigma_i$ for all $i \neq j$. Let H be the graph on $\mathcal{M}$ with edge set $\{\{\rho, \pi\} : \Delta(\rho, \pi) = \{j\}\}$. By Lemma 3.54(i) we have $H \subseteq \Gamma^*_j$. The neighbours of σ in H are precisely those elements of $\mathcal{M}$ that are indistinguishable from σ' i.e. those elements of $\mathcal{M}$ (viewed as edges of Q_k) that intersect $Q(\sigma')$. Since $\mathcal{M}$ is a perfect matching of Q_k, and $|Q(\sigma')| = 2$, there are either 1 or 2 such elements of $\mathcal{M}$. It follows that H is the disjoint union of cycles (where we consider an edge a cycle) and the independent set I_j. Since H is bipartite with bipartition $\{\tau \in \mathcal{M} : \tau_j = 0 \text{ or } *\} \cup \{\tau \in \mathcal{M} : \tau_j = 1\}$, the cycles in H are all even. The result follows. □

Let $j \in [k]$, then for each $\{\sigma, \tau\} \in T_j$ (T_j as in the statement of Lemma 3.55), we may fix a monochromatic perfect matching $M^j_{\sigma\tau}$ in the colour j in $R[W_\sigma, W_\tau]$ by Lemmas 3.54(ii) and 3.55. Let

$$\mathcal{T}_j = \bigcup_{\{\sigma,\tau\} \in T_j} M^j_{\sigma\tau}.$$

and note that $\mathcal{T}_j$ is a matching in R, monochromatic in the colour j, which covers the vertex set $\bigcup_{\tau \notin I_j} W_\tau$. The following corollary hints at an important common feature of all hypercube colourings.

Corollary 3.56. *Given $j \in [k]$ and $\rho \in \mathcal{M} \backslash I_j$, there exists $\pi \in \mathcal{M}$ such that $R[W_\rho, W_\pi]$ contains a monochromatic connected perfect matching in the colour j whose matching edges are edges of $\mathcal{T}_j$.*

Proof. Since $\rho \notin I_j$, by Lemma 3.55 there must exist $\pi \in \mathcal{M}$ such that $\{\rho, \pi\}$ is an edge of $T_j \subseteq \Gamma^*_j$. By the definition of Γ^*, we have that $R[W_\rho, W_\pi]$ is monochromatic in the colour j. The result follows from the definition of $\mathcal{T}_j$ and Lemma 3.54(ii). □

It will be useful to prune the sets V_i for $i \in R$ in such a way that if $\{x, y\}$ is an edge of the matching $\mathcal{T}_j$ then $G_j[V_x, V_y]$ is super-regular.

Lemma 3.57. *For each $i \in R$ there exists $V'_i \subseteq V_i$ such that* *(i)* $|V_i'| = (1-2^k\delta)m$ *for all* $i \in R$,

(ii) $G_j[V_x', V_y']$ *is* $2^{k+1}\delta$*-regular with density* $\geq \frac{1}{k+1}$ *for all* $j \in [k], \{x,y\} \in R_j$,

(iii) $G_j[V_x', V_y']$ *is* $(2^{k+1}\delta, \frac{1}{k+1})$*-super-regular for all* $j \in [k], \{x,y\} \in \mathcal{T}_j$.

Proof. For $i \in R$ we define a sequence of subsets $V_i^0, \ldots, V_i^k$ of V_i recursively. Let $V_i^0 := V_i$ for all $i \in R$. Suppose that for all $i \in R$ we have found $V_i^\ell \subseteq V_i$ with the following properties.

(a) $|V_i^\ell| \geq (1-2^\ell\delta)m$ for all $i \in R$,

(b) $G_j[V_x^\ell, V_y^\ell]$ is $2^\ell\delta$-regular with density $\geq 1/k - 2^\ell\delta$ for all $j \in [k], \{x,y\} \in R_j$,

(c) $G_j[V_x^\ell, V_y^\ell]$ is $(2^\ell\delta, 1/k - 2^{\ell+1}\delta)$-super-regular for all $j \in [\ell], \{x,y\} \in \mathcal{T}_j$.

By Fact 3.50, for each edge $\{u,w\}$ in the matching $\mathcal{T}_{\ell+1}$ (so in particular $\{u,w\} \in R_{\ell+1}$) there exists $V_u^{\ell+1} \subseteq V_u^\ell$ and $V_w^{\ell+1} \subseteq V_w^\ell$ such that $|V_u^{\ell+1}| = (1-2^\ell\delta)|V_u^\ell|$, $|V_w^{\ell+1}| = (1-2^\ell\delta)|V_w^\ell|$ and $G_{\ell+1}[V_u^{\ell+1}, V_w^{\ell+1}]$ is $(2^{\ell+1}\delta, 1/k - 2^{\ell+2}\delta)$-super-regular. If i is not incident to any edge of $\mathcal{T}_{\ell+1}$ then simply set $V_i^{\ell+1} = V_i^\ell$. Note that by (a), for all $i \in R$,

$$|V_i^{\ell+1}| \geq (1-2^\ell\delta)|V_i^\ell| \geq (1-2^\ell\delta)^2 m \geq (1-2^{\ell+1}\delta)m.$$

For $j \in [k]$ and $\{x,y\} \in R_j$, by (b) and Fact 3.49 we have that $G_j[V_x^{\ell+1}, V_y^{\ell+1}]$ is $2^{\ell+1}\delta$-regular with density $\geq 1/k - 2^{\ell+1}\delta$. Using (c) and Fact 3.49, it also follows that $G_j[V_x^{\ell+1}, V_y^{\ell+1}]$ is $(2^{\ell+1}\delta, 1/k - 2^{\ell+2}\delta)$-super-regular for all $j \in [\ell], \{x,y\} \in \mathcal{T}_j$. We have shown that the sets $V_i^{\ell+1}, i \in R$, satisfy (a)-(c) (with ℓ replaced by $\ell+1$). The result follows by letting V_i' be any subset of V_i^k of size $(1-2^k\delta)m$ for all $i \in R$ and appealing to Fact 3.49, noting that $1/(k+1) \leq 1/k - 2^{k+2}\delta$.

□

Given $\sigma \in \mathcal{M}$, let

$$\widetilde{W}_\sigma = \bigcup_{i \in W_\sigma} V_i' \subseteq V(G),$$

and let

$$\widetilde{W} = V(G) \setminus \bigcup_{\tau \in \mathcal{M}} \widetilde{W}_\tau.$$

As with W, we think of $\widetilde{W}$ as a small leftover set of vertices. Let $m' := (1 - 2^k\delta)m$ and note that by (3.48), $m't' \geq (1 - 3\sqrt{\delta})n$ and so by (3.50)

$$|\widetilde{W}_\tau| \geq (1-\varepsilon)m't' \geq (1-2\varepsilon)n \text{ for all } \tau \in \mathcal{M}. \tag{3.51}$$

We also have

$$|\widetilde{W}| \leq N - 2^{k-1}(1-\varepsilon)m't' = N - (1-\varepsilon)(1-2^k\delta)mt \leq 2\varepsilon N. \tag{3.52}$$

Where for the last inequality we recalled (3.46).

We can now establish our first piece of structure on the graph G. We show that almost all of $V(G)$ can be covered by 2^{k-1} monochromatic cliques of equal size. First we make a quick definition.

Definition 3.58. *If $\tau \in \{0, 1, *\}^k$ has weight 1, we let $c(\tau)$ denote the unique element of $i \in [k]$ for which $\tau_i = *$.*

Lemma 3.59. *For all $\sigma \in \mathcal{M}$, $G[\widetilde{W}_\sigma]$ is monochromatic in the colour $c(\sigma)$.*

Proof. Suppose that $G[\widetilde{W}_\sigma]$ contains an edge $\{x, y\}$ of colour $j \neq c(\sigma)$ (so that $\sigma \notin I_j$). By Corollary 3.56 there exists $\tau \in \mathcal{M}$ such that $R_j[W_\sigma, W_\tau]$ contains a connected perfect matching, F, whose matching edges are edges of $\mathcal{T}_j$. Let $q := v(F)$, then by (3.50) we have $2(1-\varepsilon)t' \leq q \leq 2(1+\varepsilon)t'$. By Lemma 3.57 we see that $G_j[\widetilde{W}_\sigma, \widetilde{W}_\tau]$ contains a spanning $(2^{k+1}\delta, 1/(k+1), m')$-super-regular blow-up of F (with the V_i' playing the role of the U_i in Definition 3.45). Suppose that $x \in V_a'$ and $y \in V_b'$, then a and b lie on the same side of the bipartition of the connected graph F and so F contains an ab-path of even length. Note that $m' \geq n_{3.46}(\delta)$, by (3.46) and (3.45). We may therefore apply Lemma 3.46 to deduce that $G_j[\widetilde{W}_\sigma, \widetilde{W}_\tau]$ contains a path of length $n-1$ joining x and y since $n - 1 \equiv 0 \pmod 2$ and

$$(1 - 6 \cdot 2^k\delta)qm' \geq 2(1 - 6 \cdot 2^k\delta)(1-\varepsilon)m't' \geq n - 1 \geq 3L \geq 3q, \tag{3.53}$$

where we used (3.44), (3.45) and (3.51). Together with the edge $\{x, y\}$ this creates a monochromatic copy of C_n in G, contrary to assumption (†). $\square$

Our aim now is to say something about the edges of G lying between $\widetilde{W}$ and the rest of the graph (see Lemma 3.62 below). With the multigraph Γ in mind, we make the following definition.

Definition 3.60. *Let $\mathcal{M}'$ be a perfect matching of Q_k and let $\varphi : \mathcal{M} \to \mathcal{M}'$ be a bijection such that $c(\varphi(\tau)) = c(\tau)$ for all $\tau \in \mathcal{M}$. Suppose that Ψ is a k-coloured multigraph on vertex set $\mathcal{M}$. We call φ an* admissible labelling *of Ψ if for all $\sigma, \tau \in \mathcal{M}$, the edges between σ, τ in Ψ only take colours from the set $\Delta(\varphi(\sigma), \varphi(\tau))$.*

Note that by Lemma 3.54(i) the identity map $\iota : \mathcal{M} \to \mathcal{M}$ is an admissible labelling of Γ. The following Lemma gives a useful way of generating new admissible labellings of Γ. For $\tau \in \{0, 1, *\}^k$, such that $\tau_j \in \{0, 1\}$, we let $\tau^j := (\tau_1, \dots, \tau_{j-1}, 1 - \tau_j, \tau_{j+1}, \dots, \tau_k)$ i.e. τ^j denotes τ with the jth coordinate flipped.

Lemma 3.61. *Let φ be an admissible labelling of Γ. Let $j \in [k]$ and let C be the vertex set of a component of Γ_j such that $\tau_j \neq *$ for all $\tau \in C$. Let φ' be the function on $\mathcal{M}$ given by $\varphi'(\tau) = \varphi(\tau)^j$ for all $\tau \in C$, $\varphi'(\tau) = \varphi(\tau)$ otherwise. Then φ' is an admissible labelling of Γ.*

Proof. First note that by the definition of φ' and the fact that φ is admissible, each element of $\varphi'(\mathcal{M})$ has weight 1 and $c(\varphi'(\tau)) = c(\varphi(\tau)) = c(\tau)$ for all $\tau \in \mathcal{M}$. Let us check that the image of φ' is a perfect matching of Q_k (i.e. a distinguishable set of size 2^{k-1}). It suffices to show that if $\sigma, \tau \in \mathcal{M}$ are distinct, then $\varphi'(\sigma), \varphi'(\tau)$ are distinguishable (i.e. $\Delta(\varphi'(\sigma), \varphi'(\tau)) \neq \emptyset$). We do this by considering an edge between σ and τ in Γ and showing that if it has the colour i then $i \in \Delta(\varphi'(\sigma), \varphi'(\tau))$. Note that this in fact suffices to show that φ' is admissible.

Suppose then that there is an edge between σ, τ in Γ in the colour i. Since φ is admissible we have $i \in \Delta(\varphi(\sigma), \varphi(\tau))$. Suppose that $i \neq j$ then by the definition of φ', $\varphi'(\tau)_i = \varphi(\tau)_i$ and $\varphi'(\sigma)_i = \varphi(\sigma)_i$ and so $i \in \Delta(\varphi'(\sigma), \varphi'(\tau))$ also. Suppose then that $i = j$, so that either $\sigma, \tau \in C$ or $\sigma, \tau \in \mathcal{M} \backslash C$. If $\sigma, \tau \in C$, then $\varphi'(\tau)_i = 1 - \varphi(\tau)_i$, $\varphi'(\sigma)_i = 1 - \varphi(\sigma)_i$ and if $\sigma, \tau \in \mathcal{M} \backslash C$, then $\varphi'(\tau)_i = \varphi(\tau)_i$, $\varphi'(\sigma)_i = \varphi(\sigma)_i$. In either case $i \in \Delta(\varphi'(\sigma), \varphi'(\tau))$. $\square$

The following lemma allows us to associate each vertex in $\widetilde{W}$ to some class $\widetilde{W}_\sigma$ in G.

Lemma 3.62. *Let $v \in \widetilde{W}$. Then there exists $\sigma \in \mathcal{M}$ such that $G[v, \widetilde{W}_\sigma]$ is monochromatic in the colour $c(\sigma)$.*

Proof. Suppose otherwise, then for each $\sigma \in \mathcal{M}$ there exists a $u \in \widetilde{W}_\sigma$ such that the edge $\{v, u\}$ receives a colour $j_\sigma \neq c(\sigma)$. We augment the multigraph Γ in the following way. We add the vertex v to Γ and for each $\sigma \in \mathcal{M}$ we add an edge between v and σ in the colour j_σ. Let us call this augmented multigraph Γ^+.

Claim 3.63. Γ^+ *contains a monochromatic odd cycle.*

Proof of Claim. Suppose otherwise and choose an admissible labelling φ of Γ that minimises the function

$$S(\varphi) = \sum_{i \in [k]} |\{\tau \in \mathcal{M} : j_\tau = i \text{ and } \varphi(\tau)_i = 1\}|.$$

Suppose that $S(\varphi) > 0$, then there exists a colour $j \in [k]$ and an element $\sigma \in \mathcal{M}$ for which $j_\sigma = j$ and $\varphi(\sigma)_j = 1$. Let C denote the component of Γ_j containing the vertex σ and note that by the definition of admissibility C is bipartite with parts $\{\tau \in C : \varphi(\tau)_j = 0\}$ and $\{\tau \in C : \varphi(\tau)_j = 1\}$. Note that since C is connected in Γ_j this is the unique bipartition of C. Since Γ_j^+ is bipartite by assumption we must therefore have that $\varphi(\tau)_j = 1$ for all $\tau \in C$ such that $j_\tau = j$. Let φ' denote the function on $\mathcal{M}$ given by $\varphi'(\tau) = \varphi(\tau)^j$ for all $\tau \in C$, $\varphi'(\tau) = \varphi(\tau)$ otherwise. By Lemma 3.61, φ' is an admissible labelling of Γ, however $S(\varphi') < S(\varphi)$ contradicting the minimality of φ. We conclude that $S(\varphi) = 0$ i.e.

$$\text{For all } i \in [k],\ \tau \in \mathcal{M}, \text{ if } j_\tau = i \text{ then } \varphi(\tau)_i = 0. \tag{3.54}$$

Since $\varphi(\mathcal{M})$ is a perfect matching of Q_k, there must exist $\rho \in \mathcal{M}$ such that the edge $\varphi(\rho)$ is incident to the vertex $(1, 1, \ldots, 1)$ (formally $Q(\varphi(\rho))$ contains $(1, 1, \ldots, 1)$). Without loss of generality suppose $\varphi(\rho) = (*, 1, \ldots, 1)$. However, whatever value j_ρ takes, we contradict (3.54). This concludes the proof of the claim. $\square$

Suppose that Γ^+ contains a monochromatic odd cycle in the colour j. Since Γ_j is bipartite and Γ_j^+ is not, there must exist $\sigma, \tau \in \mathcal{M}$ such that σ, τ lie in opposite parts of the bipartition of a connected component in Γ_j and the edges $\{v,\sigma\}, \{v,\tau\}$ both have colour j in Γ^+. By the definition of Γ^+, there exist vertices $u \in \widetilde{W}_\sigma$, $w \in \widetilde{W}_\tau$ such that $\{v,u\}$ and $\{v,w\}$ both have colour j in G and $j \neq c(\sigma)$ or $c(\tau)$ i.e. $\sigma, \tau \notin I_j$. Suppose that $u \in V_a'$ and $w \in V_b'$ then by Lemmas 3.54(ii) and 3.55 and the definition of $\mathcal{T}_j$, a and b lie in opposite parts of a bipartite connected matching, F, in R_j whose matching edges span F and are edges of $\mathcal{T}_j$ (in particular there is an ab-path of odd length in F). Moreover we may assume that F spans the vertex sets W_σ, W_τ in R_j and so by (3.50), $2(1-\varepsilon)t' \leq v(F) \leq v(R) \leq L$. By Lemma 3.57, we have a $(2^{k+1}\delta, 1/(k+1), m')$-super-regular blow-up of F in G_j. By Lemma 3.46 (using inequalities as in (3.53) and noting that $n-2$ is odd) there exists a path of length $n-2$ joining u and w in G_j. This together with the edges $\{v,u\}$, $\{v,w\}$ forms a monochromatic copy of C_n in G contrary to assumption (†). This concludes the proof of Lemma 3.62. $\square$

Using Lemma 3.62 we may define a function $f : \widetilde{W} \to \mathcal{M}$ where $f(v)$ is an element of $\mathcal{M}$ such that $G[v, \widetilde{W}_{f(v)}]$ is monochromatic in the colour $c(f(v))$. For each $\tau \in \mathcal{M}$, let $U_\tau = \widetilde{W}_\tau \cup f^{-1}(\{\tau\})$. By (3.51), (3.52), (3.46) and Lemma 3.59 we have that

$$\delta(G_{c(\tau)}[U_\tau]) \geq (1 - 2^{k+1}\varepsilon)|U_\tau| \text{ for all } \tau \in \mathcal{M}. \tag{3.55}$$

Note that the sets U_τ, $\tau \in \mathcal{M}$, partition the vertex set of G and so if $N \geq 2^{k-1}(n-1)+1$ then by the pigeonhole principle there exists $\sigma \in \mathcal{M}$ such that $|U_\sigma| \geq n$. However, by (3.55) and Theorem 3.51, it follows that U_σ contains a monochromatic copy of C_n in the colour $c(\sigma)$, contrary to assumption (†). We therefore have that $N \leq 2^{k-1}(n-1)$. Note that at this point we have done enough to prove Theorem 3.2.

It remains to show that G is close in edit distance to a hypercube colouring. Recall that $|\widetilde{W}| \leq 2\varepsilon N$ and so there are at most $2\varepsilon N^2$ edges of G incident to $\widetilde{W}$. We now aim to show that $G \backslash \widetilde{W}$ is close to a hypercube colouring. Recall that we have partitioned the vertex set of $G \backslash \widetilde{W}$ into the monochromatic, equally sized cliques $\{\widetilde{W}_\tau : \tau \in \mathcal{M}\}$. For $\sigma \in \mathcal{M}$, we showed that $|\widetilde{W}_\sigma| \geq$

$(1-2\varepsilon)n$ and $\widetilde{W}_\sigma$ is monochromatic in the colour $c(\sigma)$. First note that at most $2\varepsilon n$ vertices of $G\backslash\widetilde{W}_\sigma$ have more than $2\varepsilon n$ neighbours in $\widetilde{W}_\sigma$ in the colour $c(\sigma)$ else we immediately find a monochromatic C_n in the colour $c(\sigma)$ in G. It follows that there are at most $2\varepsilon nN$ edges leaving the clique $\widetilde{W}_\sigma$ in the colour $c(\sigma)$. Over all $\tau \in \mathcal{M}$, there are therefore at most $2^k\varepsilon nN < 3\varepsilon N^2$ edges in total leaving a clique $\widetilde{W}_\tau$ in the colour $c(\tau)$.

Let Φ now be the multigraph on vertex set $\mathcal{M}$ where we have an edge between σ and τ in the colour j for each $j \notin \{c(\sigma), c(\tau)\}$ for which $G[\widetilde{W}_\sigma, \widetilde{W}_\tau]$ contains a matching of two edges in the colour j. First we observe that to complete the proof it suffices to show that there exists an admissible labelling φ of Φ (recall Definition 3.60). Indeed suppose that this is the case, then since φ is admissible, for each pair of distinct $\sigma, \tau \in \mathcal{M}$ and each $j \notin \Delta(\varphi(\sigma), \varphi(\tau)) \cup \{c(\sigma), c(\tau)\}$, we have that $G_j[\widetilde{W}_\sigma, \widetilde{W}_\tau]$ contains no matching of two edges and hence contains at most $|\widetilde{W}_\sigma| < n$ edges in total. It follows that there is a hypercube colouring H associated to the perfect matching $\varphi(\mathcal{M})$ of Q_k, where H has vertex set $V(G)\backslash\widetilde{W}$, such that for each $i \in [k]$,

$$|G_i \triangle H_i| \leq 2\varepsilon N^2 + |(G\backslash\widetilde{W})_i \triangle H_i| \leq 2\varepsilon N^2 + 3\varepsilon N^2 + n\binom{2^{k-1}}{2} \leq 6\varepsilon N^2.$$

The $2\varepsilon N^2$ term accounts for edges of G_i incident to $\widetilde{W}$, the $3\varepsilon N^2$ term accounts for edges of G_i leaving a clique $\widetilde{W}_\tau$ where $c(\tau) = i$, and the $n\binom{2^{k-1}}{2}$ term accounts for edges of G_i lying between pairs $\widetilde{W}_\tau, \widetilde{W}_\sigma$ for which $i \notin \Delta(\varphi(\sigma), \varphi(\tau)) \cup \{c(\sigma), c(\tau)\}$. We have thus shown that G is 6ε-close to H. It remains to show that we have the desired labelling of Φ.

Claim 3.64. *Φ contains no monochromatic odd cycle.*

Proof of Claim. Suppose otherwise and let $\sigma_1 \ldots \sigma_\ell$ be an odd cycle in Φ in the colour j. This allows us to fix a matching of size two in graphs $G_j[\widetilde{W}_{\sigma_i}, \widetilde{W}_{\sigma_{i+1}}]$ for $i = 1, \ldots, \ell$ (where $\sigma_{\ell+1} := \sigma_1$). Let S be the subset of vertices of G saturated by these matchings and note that $|S| < 2^{k+1}$. We first aim to build a short even path in G_j with endpoints in $\widetilde{W}_{\sigma_1}$ and $\widetilde{W}_{\sigma_\ell}$.

Let $x \in S \cap \widetilde{W}_{\sigma_1}$ and suppose that for some $2 \leq r < \ell$ there exists $y \in \widetilde{W}_{\sigma_r}$ such that G_j contains an xy-path P_r of length $r - 1 + 2L(r-2)$ where

$|P_r \cap S \cap \widetilde{W}_{\sigma_r}| = 1$ and $P_r \cap S \cap \widetilde{W}_{\sigma_s} = \emptyset$ for $r < s \leq \ell$ (note that this does indeed hold for $r = 2$). We may then pick $w \in \widetilde{W}_{\sigma_r} \cap S$ and $z \in \widetilde{W}_{\sigma_{r+1}} \cap S$ such that $\{w, z\}$ is an edge of $G_j[\widetilde{W}_{\sigma_r}, \widetilde{W}_{\sigma_{r+1}}]$ and $w \neq y$ (here we are using that we have a matching of size two available to us by the definition of Φ). By the definition of Φ, σ_r is not in I_j and so by Corollary 3.56 there exists $\pi \in \mathcal{M}$ such that $R_j[W_{\sigma_r}, W_\pi]$ contains a connected perfect matching, F, whose matching edges are edges of $\mathcal{T}_j$. By Lemma 3.57 we see that $G_j[\widetilde{W}_{\sigma_r}, \widetilde{W}_\pi]$ contains a spanning $(2^{k+1}\delta, 1/(k+1), m')$-super-regular blow-up of F where $2(1-\varepsilon)t' \leq v(F) \leq 2(1+\varepsilon)t'$ by (3.50). Moreover w and y lie in the same part in the bipartition of this blow-up. By calculations similar to those made previously, we may apply Lemma 3.46 to deduce that $G_j[\widetilde{W}_{\sigma_r}, \widetilde{W}_\pi]$ contains an yw-path Q of length $2L$. Moreover, since $|P_r \cup S| \leq 2^k L$ and using Fact 3.49 it is easy to ensure that Q only intersects $P_r \cup S$ at its endpoints. It follows that $P_{r+1} := P_r Q z$ is an xz-path of length $r + 2L(r-1)$ where $|P_{r+1} \cap S \cap \widetilde{W}_{\sigma_{r+1}}| = 1$ and $P_{r+1} \cap S \cap \widetilde{W}_{\sigma_s} = \emptyset$ for $r+1 < s \leq \ell$. It follows by recursion that there exists $u \in \widetilde{W}_{\sigma_\ell}$ and an xu-path P_ℓ of length $p := \ell - 1 + 2L(\ell-2)$ and $|P_\ell \cap S \cap \widetilde{W}_{\sigma_\ell}| = 1$. Note that the length of P_ℓ is even.

Finally, let $\{v, t\} \subseteq S$ be an edge in $G_j[\widetilde{W}_{\sigma_\ell}, \widetilde{W}_{\sigma_1}]$ where $v \in \widetilde{W}_{\sigma_\ell}$ and $v \neq u$. If $x = t$ then applying Lemma 3.46 as above we find a uv-path Q_0 in the colour j of length $n-p-1$ intersecting $P_\ell \cup S$ only at its endpoints. It follows that $P_\ell Q_0 x$ is a monochromatic copy of C_n contradicting (†). Similarly, if $x \neq t$, we find a uv-path Q_1 of length $2L$ and a tx-path Q_2 of length $n - p - 2L - 1$ both in the colour j so that $P_\ell Q_1 t Q_2$ is a monochromatic copy of C_n contradicting (†). □

We now construct an admissible labelling of Φ recursively. Suppose that $\mathcal{M}'$ is a perfect matching of Q_k and that $\psi : \mathcal{M} \to \mathcal{M}'$ is some bijection. Let $\sigma, \tau \in \mathcal{M}$ and suppose there is an edge f in Φ between σ and τ with colour j not in $\Delta(\psi(\sigma), \psi(\tau))$. We will call such an edge 'bad' (with respect to ψ).

Let $\{f_1, \ldots, f_t\}$ be the set of edges of Φ that are bad with respect to the identity map $\iota : \mathcal{M} \to \mathcal{M}$ and note that ι is an admissible labelling of $\Phi \backslash \{f_1, \ldots, f_t\}$. Suppose now that φ_i is an admissible labelling of $\Phi^i :=$

$\Phi \backslash \{f_1, \ldots, f_i\}$ for some $1 \leq i \leq t$. Suppose that f_i is bad with respect to φ_i and that f_i has colour j and lies between $\sigma, \tau \in \mathcal{M}$. Note that $j \notin \{c(\sigma), c(\tau)\}$ by the definition of Φ. Moreover by the admissibility of φ_i we have $c(\sigma) = c(\varphi_i(\sigma))$ and $c(\tau) = c(\varphi_i(\tau))$. Since f_i is bad it follows that we must have $\varphi_i(\sigma)_j = \varphi_i(\tau)_j \in \{0, 1\}$. Let us show that σ, τ lie in separate components of Φ^i_j (the jth colour class of Φ^i). Suppose otherwise and take a path in Φ^i_j joining σ and τ. Since φ_i is admissible for Φ^i and $\varphi_i(\sigma)_j = \varphi_i(\tau)_j$ this path must have even length. It follows that f_i completes this path to a monochromatic odd cycle in Φ contradicting Claim 3.64. Let C then denote the component of Φ^i_j containing τ (so that $\sigma \notin C$). Let φ_{i-1} be the function on $\mathcal{M}$ given by $\varphi_{i-1}(\tau) = \varphi_i(\tau)^j$ for all $\tau \in C$, $\varphi_{i-1}(\tau) = \varphi_i(\tau)$ otherwise. By Lemma 3.61, φ_{i-1} is an admissible labelling of Φ^i. Since

$$j \in \Delta(\varphi_i(\sigma), \varphi_i(\tau)^j) = \Delta(\varphi_{i-1}(\sigma), \varphi_{i-1}(\tau)),$$

we also have that φ_{i-1} is an admissible labelling of Φ^{i-1}. If f_i is not bad with respect to φ_i we simply let $\varphi_{i-1} = \varphi_i$. Running this recursion to the end we obtain an admissible labelling φ_0 of Φ as required.

□

This completes the proof of Theorem 3.4. We end this chapter with a few remarks about the off-diagonal case and a related problem. A simple adaptation of the proof method in this chapter proves the following generalisation of Theorem 3.2.

Theorem 3.65. *For all $k \geq 3$ there exists N_k such that the following holds. If $N_k \leq n_1 \leq n_2 \ldots \leq n_k$ are all odd then*

$$R(C_{n_1}, \ldots, C_{n_k}) = 2^{k-1}(n_k - 1) + 1.$$

The off-diagonal case has been well-studied. Erdős *et al.* [28] determined the value of $R(C_n, C_{\ell_1}, C_{\ell_2})$ and $R(C_n, C_{\ell_1}, C_{\ell_2}, C_{\ell_3})$ for ℓ_i fixed and n sufficiently large. In a similar vein, as a corollary to a more general result in the study of Ramsey goodness, Allen, Brightwell and Skokan [2] determined the value of $R(C_n, C_{\ell_1}, \ldots, C_{\ell_k})$ for ℓ_i fixed and odd satisfying $\ell_i > 2^i$ for $1 \leq i \leq k$ and n sufficiently large. In [39], Figaj and Łuczak asymptotically determine the Ramsey number of a triple of large cycles with any fixed combination of parities for the cycle lengths. In the case where not all of the cycles have the same parity, Ferguson [35, 36, 37] strengthened the asymptotic results of [39] to exact results. It would be interesting to extend the methods of the present chapter to such a mixed parity setting. More generally, we would like to investigate whether the analytic approach presented here has wider applications in Ramsey theory. As a starting point, we believe that the methods presented in this chapter would be useful for approaching the following conjecture of Benevides, Łuczak, Scott, Skokan and White [4] (at least for large n).

Conjecture 3.66 ([4], Conjecture 8.1)**.** *Let $n \geq 3$ and let k be an integer. Let G be a k-coloured graph on n vertices with $\delta(G) \geq (1 - 2^{-k})n$, then either:*

- *For all $\ell \in [\min\{2^k, 3\}, \lceil n/2^{k-1} \rceil]$, G contains a monochromatic copy of C_ℓ, or;*
- *G is a complete 2^k-partite graph with vertex classes of equal size where each colour class is bipartite.*

4

Independent Sets and the Hard-Core Model

This chapter is based on joint work with Ewan Davies, Will Perkins and Barnaby Roberts published in [24] and [25].

4.1 Introduction

4.1.1 Independent Sets in Regular Graphs

In this chapter we explore extremal problems related to independent sets in regular graphs and triangle-free graphs of bounded degree. For a graph G let $\mathcal{I}(G)$ be the set of independent sets in G. Recall from Section 1.4 that the hard-core model with fugacity λ on G is a random independent set I drawn according to the distribution

$$\mathbb{P}_{G,\lambda}[I] = \frac{\lambda^{|I|}}{P_G(\lambda)}, \quad \text{where} \quad P_G(\lambda) = \sum_{I \in \mathcal{I}(G)} \lambda^{|I|}.$$

The occupancy fraction $\alpha_G(\lambda)$ of G is both the expected fraction of vertices of G belonging to a random independent set I and the scaled logarithmic derivative of the partition function P_G:

$$\alpha_G(\lambda) = \frac{\sum_{I \in \mathcal{I}(G)} |I| \lambda^{|I|}}{v(G) P_G(\lambda)} = \frac{\lambda P_G'(\lambda)}{v(G) P_G(\lambda)} = \frac{\lambda}{v(G)} \left(\log P_G(\lambda)\right)'. \tag{4.1}$$

In this chapter we will prove:

Theorem 4.1. *For all d-regular graphs G and all $\lambda > 0$, we have*

$$\alpha_G(\lambda) \le \alpha_{K_{d,d}}(\lambda) = \frac{\lambda(1+\lambda)^{d-1}}{2(1+\lambda)^d - 1}.$$

Moreover, the maximum is achieved only by disjoint unions of $K_{d,d}$'s.

If $2d$ divides n, let $H_{d,n}$ denote a disjoint union of $K_{d,d}$'s on n vertices. Note that by linearity of expectation, we have $\alpha_{K_{d,d}}(\lambda) = \alpha_{H_{d,n}}(\lambda)$.

From (4.1), it follows that

$$\frac{1}{v(G)} \log P_G(\lambda) = \int_0^\lambda \frac{\alpha_G(t)}{t}\, dt. \tag{4.2}$$

Thus, an immediate corollary of Theorem 4.1 is the result of Zhao [86] stating that for $\lambda > 0$, $P_G(\lambda)^{1/v(G)}$ is maximised over d-regular graphs by $K_{d,d}$. Even more, it says that the ratio $P_{K_{d,d}}(\lambda)^{1/(2d)}/P_G(\lambda)^{1/v(G)}$ is strictly increasing in λ for any d-regular graph G that is not a disjoint union of $K_{d,d}$'s.

Broadly speaking, the proof of Theorem 4.1 proceeds by writing the occupancy fraction in terms of local probabilities related to our model and then adding consistency constraints on these probabilities that must hold for all regular graphs. We then view the occupancy fraction as an objective function and maximise under these constraints via linear programming. If we restrict our attention to triangle-free graphs, and *minimise* at this point, rather than maximise, we obtain a general lower bound on $\alpha_G(\lambda)$ for triangle-free graphs.

4.1.2 Independent Sets in Triangle-Free Graphs

The following theorem is written naturally in terms of the Lambert W function $W(z)$: for $z > 0$, $W(z)$ denotes the unique positive real satisfying the relation $W(z)e^{W(z)} = z$. It will be useful to note that for $z \geq e$ we have $W(z) \geq \log z - \log\log z$.

Theorem 4.2. *Let G be a triangle-free graph with maximum degree d. Then for any $\lambda > 0$,*

$$\alpha_G(\lambda) \geq \frac{\lambda}{1+\lambda}\frac{W(d\log(1+\lambda))}{d\log(1+\lambda)}. \tag{4.3}$$

Note that we only require our graphs to be of maximum degree d, rather than d-regular here. Theorem 4.2 yields the following corollary.

Corollary 4.3. *Let G be a triangle-free graph on n vertices with maximum degree d. Then*

$$\frac{1}{|\mathcal{I}(G)|}\sum_{I\in\mathcal{I}(G)}|I| \geq (1+o_d(1))\frac{\log d}{d}n.$$

In other words, the average size of the independent sets in any n-vertex, triangle-free graph of maximum degree d is at least $(1+o_d(1))\frac{\log d}{d}n$. This should be compared to Shearer's celebrated result [77] that in any n-vertex, triangle-free graph of *average* degree d, the *maximum* size of its independent sets is at least $(1+o_d(1))\frac{\log d}{d}n$. Both Theorem 4.3 and Shearer's result imply the best known upper bound on the Ramsey number $R(3,k)$. We give the short argument here.

Corollary 4.4 (Shearer [77])**.**

$$R(3,k) \le (1+o(1))\frac{k^2}{\log k}.$$

Proof. Let G be a triangle-free graph with no independent set of size k. Since G is triangle-free, all vertex neighbourhoods are independent sets and so G must have maximum degree less than k. Applying Corollary 4.3 we see that G contains an independent set of size at least $(1+o_k(1))\frac{\log k}{k}v(G)$ but less than k, and so $v(G) < (1+o_k(1))\frac{k^2}{\log k}$ as required. □

Independent work of Bohman and Keevash [7] and Fiz Pontiveros, Griffiths, and Morris [40] shows that $R(3,k) \ge (1/4+o(1))k^2/\log k$. Reducing the factor 4 gap between these bounds is a major open problem in Ramsey theory. The above proof of Corollary 4.4 simply uses the average size of an independent set as a lower bound for the maximum size. In Section 4.5 we consider whether the discrepancy between the maximum and average size can be exploited to improve the upper bound on $R(3,k)$.

It is interesting to note that the lower bound in Theorem 4.2 is not monotone in λ whereas for any graph G, $\alpha_G(\lambda)$ is monotone increasing (see Proposition 4.11). Simply substituting $\lambda = 1$ into Theorem 4.2, does not quite suffice to prove Corollary 4.3. Surprisingly it turns out to be better to use a smaller λ and then appeal to the monotonicity of $\alpha_G(\lambda)$. As an example, using $\lambda = 1/\log d$ in Theorem 4.2 is enough to prove Corollary 4.3. This shows that Theorem 4.3 holds even when we replace the average size of an independent set with a weighted average biased toward small sets. In fact we can afford to bias using any λ of the form $d^{-o(1)}$.

We can use equation (4.2) to turn our lower bound on occupancy fraction (Theorem 4.2) into a lower bound on partition function.

Theorem 4.5. *Let G be a triangle-free graph on n vertices with maximum degree d. Then for all $\lambda > 0$,*

$$P_G(\lambda) \geq \exp\left(\left[W(d\log(1+\lambda))^2 + 2W(d\log(1+\lambda))\right]\frac{n}{2d}\right).$$

Taking $\lambda = 1$ in this Theorem yields the following immediate corollary.

Corollary 4.6. *Let G be a triangle-free graph on n vertices with maximum degree d. Then*

$$|\mathcal{I}(G)| \geq e^{\left(\frac{1}{2}+o_d(1)\right)\frac{\log^2 d}{d}n}.$$

In comparison, Cooper, Dutta, and Mubayi [20] (improving on previous results of Cooper and Mubayi [21]) proved that any triangle-free graph of average degree d has at least $e^{\left(\frac{1}{4}+o_d(1)\right)\frac{\log^2 d}{d}n}$ independent sets.

As a further corollary we get the following lower bound without degree restrictions.

Corollary 4.7. *Let G be a triangle-free graph on n vertices. Then*

$$|\mathcal{I}(G)| \geq e^{\left(\frac{\sqrt{2\log 2}}{4}+o(1)\right)\sqrt{n}\log n}.$$

This improves on a result of Cooper, Dutta, and Mubayi [20] by a factor of $\sqrt{2}$ in the exponent. The authors of [20] also provide a construction based on the analysis of the triangle-free process in [7, 40] showing that the optimal constant is at most $1 + \log 2 \approx 1.693$ (compared to the constant $\frac{\sqrt{2\log 2}}{4} \approx .294$ in Corollary 4.7).

The layout of this chapter is as follows. In the next section we introduce our method in the simpler context of triangle-free graphs. We give a proof of Theorem 4.2 and show how the proof can be easily adapted to prove Theorem 4.1 in the case where we restrict ourselves to triangle-free graphs. In Section 4.3 we deduce Theorem 4.5. In Section 4.4 we prove Theorem 4.1 in full generality. Finally in Section 4.5 we end with some conjectures and suggest new strategies for improving the upper bound on the Ramsey number $R(3, k)$.

4.2 Occupancy Fraction in Triangle-Free Graphs

In this section we restrict our attention to triangle-free graphs and introduce the occupancy method. We will prove Theorem 4.2 and show how the proof can be easily adapted to prove the triangle-free case of Theorem 4.1. We also derive the various corollaries of Theorem 4.2.

There are two key steps to our proof of Theorem 4.2. First, we define a random variable that depends on two sources of randomness: the random independent set drawn from the hard-core model on G and a uniformly chosen random vertex $v \in V(G)$. We then express the occupancy fraction in terms of two different expectations involving this random variable. This gives a constraint on the distribution of the random variable. We then optimise over all random variables that satisfy the constraint, and deduce a bound on the occupancy fraction.

First let us introduce some useful terminology. Let I be an independent set of G. We say a vertex v is *occupied* by I if $v \in I$ and *unoccupied* by I otherwise. Furthermore we say v is *covered* by I if $N(v) \cap I \neq \emptyset$ and *uncovered* by I otherwise. If there is no ambiguity we will simply say that v is covered etc. without referring to the independent set I. Note that if v is covered, it must be unoccupied. Finally, for a subset $S \subseteq V(G)$, we define the *free vertices* of S to be the set

$$S^I := \{v \in S : v \text{ uncovered by } I \setminus S\}.$$

In other words, if we reveal I only on the vertices outside of S, the free vertices of S are those that could potentially be in I.

Before moving on to the proof of Theorem 4.2 let us establish a lemma that we will make repeated use of throughout the rest of this chapter. We establish what we will refer to as the 'Gibbs property' of the hard-core distribution. The following lemma is simply a consequence of the well-known fact that the hard-core distribution is a Gibbs distribution, however we include a proof for completeness.

Lemma 4.8 (Gibbs Property)**.** *Let G be a graph and let $U \subseteq S \subseteq V(G)$. Let I be an independent set of G drawn from the hard-core model at fugacity*

λ. *Then for* $T \in \mathcal{I}(G[U])$ *we have (provided that* $\mathbb{P}(S^I = U) > 0$*)*

$$\mathbb{P}[I \cap U = T | S^I = U] = \frac{\lambda^{|T|}}{P_{G[U]}(\lambda)}$$

i.e. conditioned on the event that $S^I = U$, $I \cap U$ *is distributed according to the hard-core model on* $G[U]$ *at fugacity* λ.

Proof. For any $J \in \mathcal{I}(G[U])$ let

$$\mathcal{I}_J := \{I \in \mathcal{I}(G) : S^I = U \text{ and } I \cap U = J\}.$$

Note that for $J \in \mathcal{I}(G[U])$ the function

$$\begin{aligned} f : \mathcal{I}_\emptyset &\longrightarrow \mathcal{I}_J \\ I &\longmapsto I \cup \{J\} \end{aligned}$$

is a bijection. It follows that $\sum_{L \in \mathcal{I}_J} \lambda^{|L|} = \lambda^{|J|} \sum_{L \in \mathcal{I}_\emptyset} \lambda^{|L|}$ and so

$$\mathbb{P}[I \cap U = T | S^I = U] = \frac{\sum_{L \in \mathcal{I}_T} \lambda^{|L|}}{\sum_{J \in \mathcal{I}(G[U])} \sum_{L \in \mathcal{I}_J} \lambda^{|L|}} = \frac{\lambda^{|T|}}{\sum_{J \in \mathcal{I}(G[U])} \lambda^{|J|}}$$

as required.

□

Proof of Theorem 4.2. Let G be a triangle-free graph on n vertices and let I be a random independent set drawn from G according to the hard-core model at fugacity λ. We begin by recording two simple consequences of Lemma 4.8.

Claim 4.9. $\mathbb{P}[v \in I] = \frac{\lambda}{1+\lambda} \mathbb{P}[v \text{ uncovered}]$

Proof. Note that if $\mathbb{P}[v \text{ uncovered}] > 0$ then $\mathbb{P}[v \in I | v \text{ uncovered}] = \lambda/(1 + \lambda)$ by Lemma 4.8 with $S = U = T = \{v\}$. The result follows by noting that if v is occupied then v is also uncovered. □

Claim 4.10. *If* $\mathbb{P}[|N(v)^I| = j] > 0$*, then*

$$\mathbb{P}[v \text{ uncovered} || N(v)^I| = j] = (1 + \lambda)^{-j}.$$

Proof. First note that the event that v is uncovered is the same as the event that the free vertices in $N(v)$ are all unoccupied. Let W be any subset of $N(v)$ of size j for which $\mathbb{P}[N(v)^I = W] > 0$. Note that since G is triangle free, $N(v)$ is an independent set and so $P_{G[W]}(\lambda) = (1+\lambda)^j$. The result follows by taking $S = N(v)$, $U = W$ and $T = \emptyset$ in Lemma 4.8. □

We now write the occupancy fraction as:

$$\begin{aligned}\alpha_G(\lambda) &= \frac{1}{n}\sum_{v\in G}\mathbb{P}[v\in I]\\ &= \frac{\lambda}{1+\lambda}\cdot\frac{1}{n}\sum_{v\in G}\mathbb{P}[v \text{ uncovered}] &(4.4)\\ &= \frac{\lambda}{1+\lambda}\cdot\frac{1}{n}\sum_{v\in G}\sum_{j=0}^{d}\mathbb{P}[|N(v)^I| = j]\cdot(1+\lambda)^{-j} &(4.5)\end{aligned}$$

where (4.4) follows from Claim 4.9 and (4.5) from Claim 4.10. We define the random variable $Z = |N(v)^I|$ for a uniformly chosen vertex v. Z has two layers of randomness, that of I drawn from the hard-core measure, and that of selecting v at random. Interpreting the RHS of (4.5) in terms of Z, we obtain

$$\alpha_G(\lambda) = \frac{\lambda}{1+\lambda}\mathbb{E}[(1+\lambda)^{-Z}]\,. \tag{4.6}$$

Note also that since G is triangle-free, Z is simply the number of uncovered neighbours of v (since $N(v)$ is an independent set, elements of $N(v)$ can only be covered by vertices outside of $N(v)$). It follows that

$$\mathbb{E}Z = \frac{1}{n}\sum_{v\in G}\sum_{u\sim v}\mathbb{P}[u \text{ uncovered}] = \frac{1+\lambda}{\lambda}\cdot\frac{1}{n}\sum_{v\in G}\sum_{u\sim v}\mathbb{P}[u\in I],$$

where for the last equality we used Claim 4.9. Observe that in the sum $\sum_{v\in G}\sum_{u\sim v}\mathbb{P}[u\in I]$ each vertex u appears $\deg(u)$ times. Since G has maximum degree d we can relate Z and $\alpha_G(\lambda)$ in a second way as follows:

$$\alpha_G(\lambda) = \frac{1}{n}\sum_{v\in G}\mathbb{P}[v\in I] \geq \frac{1}{dn}\sum_{v\in G}\sum_{u\sim v}\mathbb{P}[u\in I] = \frac{\lambda}{1+\lambda}\frac{\mathbb{E}Z}{d}. \tag{4.7}$$

We aim to minimise the occupancy fraction subject to the constraints on the distribution of Z given by (4.6) and (4.7). In fact we relax the optimisation problem to optimise over all distributions of random variables Z that satisfy these constraints, not only those that arise from the hard-core model on a graph.

By Jensen's inequality applied to (4.6) we have

$$\alpha_G(\lambda) \geq \frac{\lambda}{1+\lambda} \cdot (1+\lambda)^{-\mathbb{E}Z}.$$

Recalling also the lower bound (4.7) we have

$$\frac{1+\lambda}{\lambda} \cdot \alpha_G(\lambda) \geq \max\left\{\frac{\mathbb{E}Z}{d}, (1+\lambda)^{-\mathbb{E}Z}\right\} \geq \min_{x\in\mathbb{R}^+} \max\left\{\frac{x}{d}, (1+\lambda)^{-x}\right\}.$$

To compute the minimum observe that x/d is increasing in x, whereas $(1+\lambda)^{-x}$ is decreasing. Then the minimum occurs at the value of x which makes these quantities equal i.e. the x that satisfies

$$xe^{\log(1+\lambda)x} = d$$

and hence

$$\log(1+\lambda)x = W(d\log(1+\lambda)).$$

The result follows. □

Before we explore the consequences of Theorem 4.2, we show how essentially the same proof can be used to establish Theorem 4.1 in the case where G is triangle free. The following proof is not needed for the general proof of Theorem 4.1, however we believe it is worth giving a unified argument for two results which have classical results of Kahn [55] and Shearer [77] as corollaries.

Proof of Theorem 4.1 for triangle-free graphs. Note that since G is triangle free, equation (4.6) holds for G and since G is d-regular, (4.7) holds with equality throughout. That is we have

$$\alpha_G(\lambda) = \frac{\lambda}{1+\lambda}\frac{\mathbb{E}Z}{d} = \frac{\lambda}{1+\lambda}\mathbb{E}[(1+\lambda)^{-Z}], \tag{4.8}$$

where we recall that Z is a random variable bounded between 0 and d. Now instead of asking for the minimum value of $\alpha_G(\lambda)$ over all distributions of Z as we did in the proof of Theorem 4.2, we ask for the maximum. Note that since $0 \leq Z/d \leq 1$, convexity of the function $x \mapsto (1+\lambda)^{-x}$ implies that

$$(1+\lambda)^{-Z} \leq \frac{Z}{d}(1+\lambda)^{-d} + 1 - \frac{Z}{d}. \tag{4.9}$$

Substituting this into (4.8) and using linearity of expectation yields

$$\alpha_G(\lambda) = \frac{\lambda}{1+\lambda}\frac{\mathbb{E}Z}{d} \leq \frac{\lambda(1+\lambda)^{d-1}}{2(1+\lambda)^d - 1}, \tag{4.10}$$

where the right hand side is the occupancy fraction of $K_{d,d}$. For uniqueness, note that to have equality in (4.10) we must have had equality in (4.9) which is only possible if Z takes only the values 0 and d. This distribution of Z can only occur in a disjoint union of copies of $K_{d,d}$. To see this recall that Z is the number of uncovered neighbours of a randomly selected vertex v. The only way every vertex v can always have either 0 or d uncovered neighbours is for all the neighbours of v to have the same neighbourhood. For d-regular graphs this property holds only for disjoint unions of $K_{d,d}$. □

In Section 4.4 we establish Theorem 4.1 in full generality. For now let us return to Theorem 4.2. A curious feature of this result is that the lower bound $\frac{\lambda}{1+\lambda}\frac{W(d\log(1+\lambda))}{d\log(1+\lambda)}$ is not monotone increasing with λ, whereas for any graph G, the occupancy fraction $\alpha_G(\lambda)$ *is* monotone increasing.

Proposition 4.11. *For any graph G, $\alpha_G(\lambda)$ is monotone increasing in λ.*

Proof. We will show that the derivative of $\alpha_G(\lambda)$ with respect to λ is positive. By (4.1) (using P for $P_G(\lambda)$), we have

$$\begin{aligned} v(G)\cdot\alpha_G'(\lambda) &= \left(\frac{\lambda P'}{P}\right)' = \frac{P'}{P} + \frac{\lambda P P'' - \lambda(P')^2}{P^2} \\ &= \frac{P'}{P} + \frac{1}{\lambda}\left(\frac{\lambda^2 P''}{P} - \left(\frac{\lambda P'}{P}\right)^2\right) \\ &= \frac{\mathbb{E}(|I|) + \mathbb{E}(|I|^2) - \mathbb{E}(|I|) - (\mathbb{E}(|I|))^2}{\lambda} \\ &= \frac{\operatorname{var}(|I|)}{\lambda} \geq 0 \end{aligned}$$

where I is a random independent set drawn from the hard-core model at fugactiy λ. □

We now prove Corollary 4.3. Rather than substituting $\lambda = 1$ in Theorem 4.2, it turns out to be better to use a smaller λ and then appeal to monotonicity.

Proof of Corollary 4.3. Substituting $\lambda = 1/\log d$ in (4.3) and recalling the bound $W(z) \geq \log z - \log\log z$ for $z \geq e$ we obtain

$$\alpha_G(\lambda) \geq (1 + o(1))\frac{\log d}{d}.$$

By monotonicity (Proposition 4.11), we have $\alpha_G(1) \geq \alpha_G(\lambda)$ and the result follows. $\square$

4.3 Counting Independent Sets in Triangle-Free Graphs

In this section we prove Theorem 4.5 (and hence Corollary 4.6) by integrating the lower bound on the occupancy fraction of a triangle-free graph given in Theorem 4.2. We also prove Corollary 4.7.

Proof of Theorem 4.5. By (4.2) and Theorem 4.2 we have

$$\begin{aligned}\log P_G(\lambda) &\geq \frac{n}{d}\int_0^\lambda \frac{W(d\log(1+t))}{(1+t)\log(1+t)}\,dt \\ &= \frac{n}{d}\int_0^{W(d\log(1+\lambda))} (1+u)\,du \\ &= \frac{n}{2d}\left[W(d\log(1+\lambda))^2 + 2W(d\log(1+\lambda))\right], \qquad (4.11)\end{aligned}$$

where for the first equality we used the substitution $u = W(d\log(1+t))$. In particular when $\lambda = 1$, using the inequality $W(z) \geq \log z - \log\log z$ for $z \geq e$, we have

$$\log P_G(\lambda) \geq \left(\frac{1}{2} + o_d(1)\right)\frac{\log^2 d}{d}n. \qquad \square$$

Proof of Corollary 4.7. In a triangle-free graph the neighbourhood of any vertex forms an independent set. Let d be the largest degree of a vertex in G, then we have the bound

$$P_G(\lambda) \geq \max\left\{(1+\lambda)^d, \exp\left[\frac{n}{2d}W(d\log(1+\lambda))^2\right]\right\},$$

by considering the neighbourhood of a vertex of maximum degree and by inequality (4.11). The first expression is increasing in d while the second is decreasing, and at

$$d = \frac{1}{2}\sqrt{\frac{n}{2\log(1+\lambda)}}\log\left(\frac{n\log(1+\lambda)}{2}\right)$$

they are equal. It follows that for $\lambda > 0$,

$$P_G(\lambda) \geq \exp\left[\frac{1}{2}\sqrt{\frac{n\log(1+\lambda)}{2}}\log\left(\frac{n\log(1+\lambda)}{2}\right)\right]. \tag{4.12}$$

Take $\lambda = 1$ to complete the proof. □

Inequality (4.12) may be of independent interest, giving a general lower bound for the independence polynomial of a triangle-free graph on n vertices.

4.4 Proof of Theorem 4.1

In this section, we prove Theorem 4.1 in full generality. Let G be a d-regular graph on n vertices. For a vertex $v \in G$ and an independent set I, we define the *free neighbourhood* $F(v)$ of v to be the subgraph of G induced by the neighbours of v which are not adjacent to any vertex in $I \setminus N(v)$ (i.e. the graph induced by $N(v)^I$). Recall that in the triangle-free case, the free neighbourhood of a vertex v was simply an independent set of size j for some $0 \leq j \leq d$. Here, the free neighbourhood could be any graph on at most d vertices. The vertices in the free neighbourhood may be uncovered or covered by I, but if they are covered it must be from another vertex in the free neighbourhood. Note that if $v \in I$, then the free neighbourhood of v is necessarily empty.

Let C be the random free neighbourhood of v when we draw I according to the hard-core model and choose vertex v uniformly at random from G. For any graph F, let p_F be the probability that C is isomorphic to F. Note that

$$p_F = \frac{1}{n}\sum_{v \in G}\mathbb{P}[F(v) = F]. \tag{4.13}$$

Note that we write $F(v) = F$ to mean that $F(v)$ is isomorphic to F. Let $\mathcal{C}_d$ be the set of all graphs on at most d vertices, including the empty graph. A key observation is that the occupancy fraction $\alpha_G(\lambda)$ can be expressed in two distinct ways in terms of the random free neighbourhood C. Recall that $P_C(\lambda)$ denotes the independence polynomial of C at fugacity λ.

Claim 4.12.

$$\alpha_G(\lambda) = \frac{\lambda}{1+\lambda}\mathbb{E}\left[\frac{1}{P_C(\lambda)}\right] = \frac{\lambda}{d}\mathbb{E}\left[\frac{P_C'(\lambda)}{P_C(\lambda)}\right]$$

where the expectations are over the random free neighbourhood C.

Proof. We proceed in a similar way to the proof of Theorem 4.2. First note that for $v \in G$, v is uncovered if and only if all the vertices in the free neighbourhood of v are unoccupied and so for $F \in \mathcal{C}_d$

$$\mathbb{P}[v \text{ uncovered} | F(v) = F] = \frac{1}{P_F(\lambda)} \tag{4.14}$$

by Lemma 4.8, taking $S = N(v)$, $T = \emptyset$ and letting U be any subset of $N(v)$ for which $G[U]$ is isomorphic to F. Equation (4.14) is the analogue of Claim 4.10 in this more general setting. Now we may write

$$\begin{aligned}
\alpha_G(\lambda) &= \frac{1}{n}\sum_{v \in G} \mathbb{P}[v \in I] \\
&= \frac{\lambda}{1+\lambda} \cdot \frac{1}{n}\sum_{v \in G} \mathbb{P}[v \text{ uncovered}] && (4.15) \\
&= \frac{\lambda}{1+\lambda} \cdot \frac{1}{n}\sum_{v \in G}\sum_{F \in \mathcal{C}_d} \frac{1}{P_F(\lambda)}\mathbb{P}[F(v) = F] && (4.16) \\
&= \frac{\lambda}{1+\lambda} \sum_{F \in \mathcal{C}_d} \frac{1}{P_F(\lambda)} \left(\frac{1}{n}\sum_{v \in G} \mathbb{P}[F(v) = F]\right) \\
&= \frac{\lambda}{1+\lambda} \sum_{F \in \mathcal{C}_d} \frac{1}{P_F(\lambda)} p_F. && (4.17)
\end{aligned}$$

For (4.15) we used Claim 4.9, for (4.16) we used (4.14) and for (4.17) we used (4.13). This establishes the first equality .

Alternatively we may write

$$\alpha_G(\lambda) = \frac{1}{dn}\sum_{v\in G}\sum_{u\sim v}\mathbb{P}[u\in I] \tag{4.18}$$

$$= \frac{1}{dn}\sum_{v\in G}\sum_{F\in\mathcal{C}_d}\left(\sum_{u\sim v}\mathbb{P}[u\in I|F(v)=F]\right)\mathbb{P}[F(v)=F] \tag{4.19}$$

$$= \frac{1}{dn}\sum_{v\in G}\sum_{F\in\mathcal{C}_d}\frac{\lambda P_F'(\lambda)}{P_F(\lambda)}\mathbb{P}[F(v)=F] \tag{4.20}$$

$$= \frac{1}{d}\sum_{F\in\mathcal{C}_d}\frac{\lambda P_F'(\lambda)}{P_F(\lambda)}\left(\frac{1}{n}\sum_{v\in G}\mathbb{P}[F(v)=F]\right)$$

$$= \frac{1}{d}\sum_{F\in\mathcal{C}_d}\frac{\lambda P_F'(\lambda)}{P_F(\lambda)}p_F.$$

For (4.18), we used that G is d-regular. For (4.20) note that the inner bracket of (4.19) is the expected number of occupied neighbours of v, given that $F(v) = F$ which by Lemma 4.8 is $\frac{\lambda P_F'(\lambda)}{P_F(\lambda)}$ i.e. the expected size of an independent set drawn from the hard-core model on F. □

Now let

$$\alpha^* = \frac{\lambda}{1+\lambda}\cdot\sup\left\{\mathbb{E}\left[\frac{1}{P_C(\lambda)}\right] : \mathbb{E}\left[\frac{1}{P_C(\lambda)}\right] = \frac{1+\lambda}{d}\cdot\mathbb{E}\left[\frac{P_C'(\lambda)}{P_C(\lambda)}\right]\right\} \tag{4.21}$$

where the supremum is over all distributions of random free neighbourhoods C supported on graphs of at most d vertices. From Claim 4.12, the distribution obtained from G satisfies the constraint above and so $\alpha_G(\lambda) \le \alpha^*$.

To complete the proof of Theorem 4.1 we will show that $\alpha^* = \alpha_{K_{d,d}}(\lambda)$. Moreover we will show that any distribution attaining the supremum in (4.21) must be supported only on the empty graph and the graph consisting of d isolated vertices, $\overline{K_d}$. The theorem follows since a disjoint union of $K_{d,d}$'s is the only graph which gives rise to a distribution with such a support:

Claim 4.13. *Suppose that G is a d-regular graph and $p_F = 0$ for all $F \in \mathcal{C}_d \setminus \{\emptyset, \overline{K_d}\}$. Then G is a disjoint union of $K_{d,d}$'s.*

Proof. First note that if $F = G[N(v)]$ for some $v \in G$, then $p_F > 0$ since we could pick v and the empty independent set. It follows that all vertex neighbourhoods must induce a copy of $\overline{K_d}$ in G i.e. G is triangle-free.

Suppose that G has a component H not isomorphic to $K_{d,d}$ and pick $v \in H$. Since H is not isomorphic to $K_{d,d}$ we can pick $u, w \in N(v)$ such that $N(u) \neq N(w)$ i.e. we can choose z such that z is adjacent to u but not w. If we choose the independent set consisting only of the vertex z then w is in the free neighbourhood of v, but u is not. Thus, the free neighbourhood of v is neither $\emptyset$ not $\overline{K_d}$. □

To prove our claim regarding the distributions that achieve α^*, we use the language of linear programming introduced in Chapter 1.

4.4.1 The Linear Program

Recall that p_F is the probability of a given free neighbourhood F. Equation (4.21) leads us to consider the following linear program with the decision variables $\{p_F\}_{C \in \mathcal{C}_d}$.

$$\begin{aligned} \text{maximise } & \frac{\lambda}{1+\lambda} \sum_{F \in \mathcal{C}_d} p_F \cdot a_F \\ \text{subject to } & \sum_{F \in \mathcal{C}_d} p_F = 1 \\ & \sum_{F \in \mathcal{C}_d} p_F (a_F - b_F) = 0 \\ & p_F \geq 0 \ \ \forall F \in \mathcal{C}_d \end{aligned}$$

where $a_F = \frac{1}{P_F(\lambda)}$ and $b_F = \frac{(1+\lambda)P_F'(\lambda)}{dP_F(\lambda)}$. Note that an optimal solution of this linear program will have objective value α^*.

The dual linear program is

$$\begin{aligned} \text{minimise } & \Lambda_1 \\ \text{subject to } & \Lambda_1 + \Lambda_2(a_F - b_F) \geq \frac{\lambda}{1+\lambda} a_F \ \ \forall F \in \mathcal{C}_d \end{aligned}$$

where Λ_1, Λ_2 are the decision variables.

We will show that $\Lambda_1 = \alpha_{K_{d,d}}(\lambda) = \frac{\lambda(1+\lambda)^{d-1}}{2(1+\lambda)^d - 1}$ and $\Lambda_2 = \frac{\lambda}{1+\lambda} - \Lambda_1$ is a feasible solution to the dual program i.e. we will show that

$$\Lambda_1 + \Lambda_2(a_F - b_F) \geq \frac{\lambda}{1+\lambda} a_F \qquad \forall F \in \mathcal{C}_d. \tag{4.22}$$

We can calculate $a_\emptyset = 1$, $b_\emptyset = 0$, $a_{\overline{K}_d} = (1+\lambda)^{-d}$, $b_{\overline{K}_d} = 1$ and so (4.22) holds with equality for $F = \emptyset, \overline{K}_d$. We will in fact show that (4.22) holds with strict inequality for all $F \in \mathcal{C}_d \setminus \{\emptyset, \overline{K}_d\}$. Substituting our values of Λ_1, Λ_2, this inequality reduces to showing that

$$\frac{\lambda P_F'(\lambda)}{P_F(\lambda) - 1} < \frac{\lambda d(1+\lambda)^{d-1}}{(1+\lambda)^d - 1} \qquad \forall F \in \mathcal{C}_d \setminus \{\emptyset, \overline{K}_d\}. \tag{4.23}$$

The LHS of (4.23) is the expected size of the random independent set from the hard-core model on F conditioned on it being non-empty. The RHS is the same quantity for $\overline{K}_d$.

Inequality (4.23) follows directly from the observation that, over all $F \in \mathcal{C}_d$, the graph $\overline{K}_d$ maximises the ratio of subsequent terms in the polynomial P_F. Let $t_i = \binom{d}{i}$, the coefficient of λ^i in $P_{\overline{K}_d}$, and write $P_F = 1 + \sum_{i=1}^d r_i \lambda^i$. We have $(i+1)t_{i+1} = (d-i)t_i$ and $(i+1)r_{i+1} \le (d-i)r_i$ by counting independent sets of size $i+1$ (recall that F has at most d vertices).

To verify (4.23) we show that for each $1 \le k \le d$ the coefficient s_k of λ^k in the polynomial $(\lambda P'_{\overline{K}_d})(P_F - 1) - (\lambda P'_F)(P_{\overline{K}_d} - 1)$ is non-negative. We have

$$\begin{aligned} s_k &= \sum_{i=1}^{k-1} i t_i r_{k-i} - \sum_{i=1}^{k-1} i t_{k-i} r_i \\ &= \sum_{i=1}^{\lfloor k/2 \rfloor} (k-2i)(t_{k-i} r_i - t_i r_{k-i}) \,. \end{aligned}$$

Observe that each term in the above sum is non-negative by comparing the ratio of successive coefficients in $P_{\overline{K}_d}$ and P_F. Furthermore, if $P_F \ne P_{\overline{K}_d}$ then at least one s_k must be positive and (4.23) follows. It follows by weak duality (Theorem 1.3) that

$$\alpha^* \le \Lambda_1 = \alpha_{K_{d,d}}(\lambda). \tag{4.24}$$

Of course we in fact have equality in (4.24) as witnessed by the distribution associated to $K_{d,d}$: $p_\emptyset = \frac{1-(1+\lambda)^{-d}}{2-(1+\lambda)^{-d}}$, $p_{\overline{K_d}} = \frac{1}{2-(1+\lambda)^{-d}}$, $p_F = 0$ for all $F \in \mathcal{C}_d \setminus \{\emptyset, \overline{K}_d\}$. The strict inequality in (4.23) shows, by complementary slackness (Theorem 1.4), that any optimal solution must be supported only on the configurations $\emptyset$ and $\overline{K}_d$. Recall that by Claim 4.13, disjoint unions of $K_{d,d}$ are the only graphs which induce a distribution with this support. The result follows.

4.5 On the Ratio of the Maximum and Average Independent Set Size

In light of Corollary 4.3, showing that the average size of an independent set in a triangle-free graph with maximum degree d is at least $(1+o_d(1))\frac{\log d}{d}n$, we now raise the question of whether the largest independent set should be significantly larger. This gives a new way to pursue an upper bound on $R(3,k)$. There is always a gap between the maximum and average size of an independent set (since the empty set is an independent set), but in general the ratio of maximum to average size can be arbitrarily close to 1. For example, the complete graph K_n has maximum independent set size 1 with average size $n/(n+1)$.

We conjecture that such a narrow gap cannot occur in triangle-free graphs. The following two conjectures make this claim precise in different ways. We let $\overline{\alpha}(G)$ denote the average size of the independent sets in G and let $\alpha(G)$ denote the largest size of an independent set in G.

Conjecture 4.14. *For every triangle-free graph G,*

$$\frac{\alpha(G)}{\overline{\alpha}(G)} \geq 4/3.$$

Replacing 4/3 with any number strictly greater than 1 would give an improvement to the $R(3,k)$ bound. The graph with the smallest ratio $\alpha/\overline{\alpha}$ we have found is the triangle-free cyclic graph that exhibits the bound $R(3,9) \geq 36$ [49]. For this graph $\alpha/\overline{\alpha} = \frac{197136}{137585} = 1.43283\ldots$. We choose 4/3 since it is a nice fraction less than 1.43 and since it is the ratio of maximum to average size in a triangle. One might wonder if the extremal $R(3,k)$ graphs are good candidates for pushing the ratio $\alpha/\overline{\alpha}$ down to 1. However, for large k it may be the case that graphs arising from the triangle-free process are asymptotically extremal, as is conjectured in [40]. We believe that for such graphs the ratio $\alpha/\overline{\alpha}$ in fact converges to 2. This motivates the following conjecture.

Conjecture 4.15. *For every triangle-free graph G of minimum degree d,*

$$\frac{\alpha(G)}{\overline{\alpha}(G)} \geq 2 - o_d(1).$$

Lemma 4.16. *The following improvements to Shearer's upper bound on $R(3,k)$ would follow from the above conjectures and Corollary 4.3.*

1. *Conjecture 4.14 implies $R(3,k) \leq (3/4 + o(1))k^2/\log k$.*

2. *Conjecture 4.15 implies $R(3,k) \leq (1/2 + o(1))k^2/\log k$.*

Proof. Let G be a triangle-free graph on n vertices with no independent set of size k. Since vertex neighbourhoods in G are independent sets we see that G must have maximum degree less than k. Corollary 4.3 and Conjecture 4.14 would then imply that $\alpha(G) \geq (4/3 + o_k(1))\frac{\log k}{k}n$. However, we also have that $\alpha(G) < k$ and so $n \leq (3/4 + o(1))k^2/\log k$.

To show (2), we select a vertex in G with degree less than $k/\log^2 k$ and remove it along with all its neighbours. We then repeat this process until it is no longer possible and call the remaining graph G'. Since the selected vertices form an independent set in G, we can repeat the process at most k times and so $v(G') \geq n - k^2/\log^2 k$. If G' is empty then $n \leq k^2/\log^2 k$, otherwise G' is a graph of minimum degree at least $k/\log^2 k$. Since G' also has maximum degree k it follows from Corollary 4.3 and Conjecture 4.14 that $\alpha(G') \geq (2 + o_k(1))\frac{\log k}{k}(n - \frac{k^2}{\log^2 k})$. However, we also have that $\alpha(G') \leq \alpha(G) < k$ and so $n \leq (1/2 + o(1))k^2/\log k$. □

One possible approach to the above conjectures is via the following simple consequence of the proof of Proposition 4.11. For any graph G on n vertices

$$\alpha(G) = v(G) \lim_{\lambda \to \infty} \alpha_G(\lambda) = \overline{\alpha}(G) + \int_1^\infty \frac{\mathrm{var}_\lambda(|I|)}{\lambda}\, d\lambda.$$

In this chapter we gave a lower bound for $\overline{\alpha}(G)$, the expected size of an independent set drawn from a triangle-free graph according to the hard-core model. The above equation shows that one approach to the above conjectures would be to do the same for the variance.

5

Matchings and the Monomer-Dimer Model

This chapter is based on joint work with Ewan Davies, Will Perkins and Barnaby Roberts published in [24].

5.1 Introduction

Recall from Section 1.4 that the matching polynomial of a graph G is

$$M_G(\lambda) = \sum_{H \in \mathcal{M}(G)} \lambda^{|H|}$$

where $\mathcal{M}(G)$ is the set of all matchings of G (including the empty matching) and $|H|$ is the number of edges in the matching H. Just as in the hard-core model we can define a probability distribution over matchings:

$$\mathbb{P}_{G,\lambda}[H] = \frac{\lambda^{|H|}}{M_G(\lambda)}\,.$$

For a d-regular graph G, the *edge occupancy fraction*, or the dimer density, is the expected fraction of the edges of G in such a random matching:

$$\alpha_G^M(\lambda) := \frac{1}{e(G)} \sum_{H \in \mathcal{M}(G)} |H| \cdot \mathbb{P}[H] = \frac{\lambda}{e(G)} (\log M_G(\lambda))'\,.$$

We then have

$$\frac{1}{e(G)} \log M_G(\lambda) = \int_0^\lambda \frac{\alpha_G^M(t)}{t}\, dt. \tag{5.1}$$

Our next result is an upper bound on the edge occupancy fraction of any d-regular graph:

Theorem 5.1. *For all d-regular graphs G and all $\lambda > 0$, we have*

$$\alpha_G^M(\lambda) \le \alpha_{K_{d,d}}^M(\lambda)\,.$$

Moreover, the maximum is achieved only by disjoint unions of $K_{d,d}$'s.

It follows by (5.1) that $K_{d,d}$ (and thus also $H_{d,n}$) maximises $M_G(\lambda)^{1/e(G)}$ over all d-regular graphs for any $\lambda > 0$. This resolves Conjecture 7.1 in [45]. A corollary of Bregman's theorem [13] on the permanents of 0/1 matrices with given row sums is that the number of perfect matchings of a d-regular,

n-vertex bipartite graph is maximised by $H_{d,n}$, and this was extended by Kahn and Lovász to all d-regular graphs (see [45] for a full discussion). Our result on $M_G(\lambda)$ extends this: letting $\lambda \to \infty$ recovers the result for perfect matchings, while setting $\lambda = 1$ shows that $H_{d,n}$ maximises the total number of matchings of any d-regular graph on n vertices.

For a graph G, let $m_k(G)$ denote the number of matchings with k edges in G. The Upper Matching Conjecture of Friedland, Krop, and Markström [44] asserts that over all n-vertex d-regular graphs, $H_{d,n}$ should in fact maximise $m_k(G)$ for all k (when $2d$ divides n). Previous bounds towards this conjecture were given in [15, 54]; for d fixed and k linear in n, all previous bounds were off the conjectured values by a multiplicative factor exponential in n. In Section 5.3 we use Theorem 5.1 along with a theorem of Heilman and Lieb [52] to give a bound that is tight up to a factor of $7\sqrt{k}$ (which is at worst $7\sqrt{n/2}$), for all d.

Theorem 5.2. *For all d-regular graphs G on n vertices where $2d|n$,*

$$m_k(G) \leq 7\sqrt{k} \cdot m_k(H_{d,n}) \,.$$

Although the Upper Matching Conjecture remains open, Theorem 5.2 is strong enough to imply the Asymptotic Upper Matching Conjecture of Friedland, Krop, Lundow, and Markström [43]. We defer the precise statement of this conjecture to Section 5.3.

5.2 Proof of Theorem 5.1

The proof of Theorem 5.1 follows the same general approach as the proof of Theorem 4.1 from the previous chapter. Given a graph G, we express its edge occupancy fraction as a linear function of certain 'local probabilities' related to the monomer-dimer model on G. We then optimise the occupancy fraction subject to linear consistency constraints on these probabilities that must hold for all regular graphs.

First let us introduce some useful notation and terminology similar to that used in the previous chapter. Let G be a graph and let H be a matching in

G. We refer to an edge e of G as *covered* by H if an edge incident to e is in H. We say that e is *uncovered* by H otherwise. We may simply say that e is covered/uncovered if the matching H is clear from the context.

For a subset $S \subseteq E(G)$, we define the set

$$S^H := \{e \in S : e \text{ uncovered by } H \setminus S\}.$$

In other words, if we reveal H only on the edges outside of S, the edges of S^H are those that could potentially be in H.

By applying Lemma 4.8 from the previous chapter to the line graph of a graph G we recover the Gibbs property for the monomer-dimer model. In the following statement we identify a graph with its edge set.

Lemma 5.3. *Let G be a graph and let $U \subseteq S \subseteq E(G)$. Let H be a matching in G drawn from the monomer-dimer model at fugacity λ. Then for $T \in \mathcal{M}(U)$ we have (provided that $\mathbb{P}(S^H = U) > 0$)*

$$\mathbb{P}[H \cap U = T | S^H = U] = \frac{\lambda^{|T|}}{M_U(\lambda)}.$$

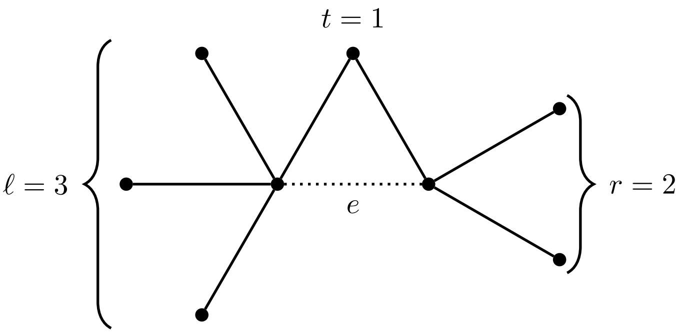
I.e. conditioned on the event that $S^H = U$, $H \cap U$ is distributed according to the monomer-dimer model on U at fugacity λ.

Now, let G be a d-regular graph on n vertices and let H be a random matching drawn from the monomer-dimer model on G at fugacity λ.

For an edge e, we define the *free neighbourhood* $F(e)$ of e to be the subgraph of G containing all of the edges incident to e that are uncovered by edges outside of both e and its incident edges. Note that when considering independent sets in the previous chapter, the free neighbourhood was empty if the random vertex v was in the independent set. Here the presence or absence in the matching of e or an edge adjacent to e does not affect $F(e)$.

Let us give each edge of G an arbitrary orientation that we fix throughout the proof i.e. for each edge $e \in G$, one endpoint of e is chosen to be the 'left side' and the other endpoint the 'right side'. The possible free neighbourhoods of an edge e are then completely defined by three parameters: $\ell(e), r(e), t(e) \in \{0, 1, \ldots, d-1\}$, counting the number of edges incident to the left side of e

in $F(e)$ with an endpoint of degree 1, the same for the right side, and the number of triangles formed by e and $F(e)$. We write $F(e) = (i, j, k)$ as a shorthand to denote the event that $(\ell(e), r(e), t(e)) = (i, j, k)$. An example is pictured below.



We denote the matching polynomial for such a free neighbourhood by $M_{i,j,k}$, where we can compute

$$M_{i,j,k}(\lambda) = 1 + (i + j + 2k)\lambda + \left[k^2 + k(i + j - 1) + ij\right]\lambda^2 .$$

Let $\mathbf{e}$ be an edge of G chosen uniformly at random and let $q(i, j, k)$ denote the probability that $F(\mathbf{e}) = (i, j, k)$ i.e.

$$q(i, j, k) = \frac{2}{dn} \sum_{e \in G} \mathbb{P}[F(e) = (i, j, k)].$$

We can write α_G^M as the expected fraction of edges incident to $\mathbf{e}$ that are in the random matching H:

$$\alpha_G^M(\lambda) = \frac{2}{dn} \sum_{e} \sum_{f \sim e} \frac{1}{2(d-1)} \mathbb{P}[f \in H] \tag{5.2}$$

$$= \frac{2}{dn} \sum_{e} \sum_{i,j,k} \frac{1}{2(d-1)} \left(\sum_{f \sim e} \mathbb{P}[f \in H | F(e) = (i, j, k)] \right) \mathbb{P}[F(e) = (i, j, k)]$$

$$= \frac{2}{dn} \sum_{e} \sum_{i,j,k} \frac{1}{2(d-1)} \frac{\lambda M'_{i,j,k}(\lambda)}{(\lambda + M_{i,j,k}(\lambda))} \mathbb{P}[F(e) = (i, j, k)] \tag{5.3}$$

$$= \sum_{i,j,k} \frac{\lambda M'_{i,j,k}(\lambda)}{2(d-1)(\lambda + M_{i,j,k}(\lambda))} q(i, j, k) .$$

For (5.2) we used that G is d-regular and for (5.3) we used Lemma 5.3: conditioned on the event $F(e) = (i, j, k)$, $H \cap (F(e) \cup e)$ is distributed according to the monomer-dimer model on $F(e) \cup e$ at fugacity λ. Note that the matching polynomial of $F(e) \cup e$ is $\lambda + M_{i,j,k}$ and so $\overline{\alpha}(i, j, k) := \frac{1}{2(d-1)} \frac{\lambda M'_{i,j,k}}{\lambda + M_{i,j,k}}$ is the expected fraction of neighbours of e that are in H conditioned on the event $F(e) = (i, j, k)$. With this notation the above expression can be written $\alpha_G^M = \sum_{i,j,k} \overline{\alpha}(i, j, k) q(i, j, k)$.

5.2.1 The Linear Program for Matchings

Our goal is to introduce linear consistency constraints on the $q(i, j, k)$ and then optimise $\alpha_G^M = \sum_{i,j,k} \overline{\alpha}(i, j, k) q(i, j, k)$ subject to these constraints. We could write multiple expressions for α_G^M, equate them, and solve the maximisation problem as we did for independent sets in Chapter 4. Using three expressions for α_G^M we were able to prove Theorem 5.1 for the case $d = 3$, in which the optimal distribution is supported on only three values: $q(0, 0, 0)$, $q(1, 1, 0)$, $q(2, 2, 0)$. But in general we need at least $d - 1$ constraints (in addition to the constraint that the $q(i, j, k)'s$ sum to one) as the distribution induced by $K_{d,d}$ is supported on d values.

Instead we write, for all t, two expressions for the probability that the number of uncovered neighbours on a randomly chosen side of a random edge is equal to t. We find the two expressions by choosing uniformly: a random edge **e**, a random side of **e** left or right, and **f**, a random neighbouring edge of **e** from the given side. We first calculate the probability that **e** has t uncovered neighbours on the side incident to **f**, then we calculate the probability that **f** has t uncovered neighbours on the side incident to **e**.

Given an edge e of G with free neighbourhood $F(e) = (i, j, k)$, e can have $0, 1, i+k-1$, or $i+k$ uncovered left neighbours; an edge f to the left of e can have $0, 1, i+k-2, i+k-1, i+k$, or $i+k+1$ uncovered neighbours on the side containing e (depending on whether f itself is in $F(e)$). To save space we use 'nbrs' as an abbreviation for 'neighbours' in some of the following displayed equations.

Let

$$\gamma^e_{i,j,k}(t) := \mathbb{P}[\mathbf{e} \text{ has } t \text{ uncovered left nbrs } | F(\mathbf{e}) = (i,j,k)]$$

and

$$\gamma^f_{i,j,k}(t) := \mathbb{P}[\mathbf{f} \text{ has } t \text{ uncovered nbrs on the side containing } \mathbf{e} | F(\mathbf{e}) = (i,j,k)]$$

where $\mathbf{f}$ is a uniformly chosen left neighbour of $\mathbf{e}$.

Claim 5.4. *Let $\beta_t = 1 + t\lambda$. Then we have*

$$\gamma^e_{i,j,k}(t) = \frac{1}{\lambda + M_{i,j,k}} \Big(\mathbf{1}_{t=0} \cdot \lambda + \mathbf{1}_{t=1} \cdot [i\lambda\beta_{j+k} + k\lambda\beta_{j+k-1}] + \mathbf{1}_{t=i+k} \cdot \beta_j + \mathbf{1}_{t=i+k-1} \cdot k\lambda \Big) \tag{5.4}$$

$$\begin{aligned} \gamma^f_{i,j,k}(t) = \frac{1}{(d-1)(\lambda + M_{i,j,k})} \Big(&\mathbf{1}_{t=0} \cdot [i\lambda\beta_{j+k} + k\lambda\beta_{j+k-1}] \\ &+ \mathbf{1}_{t=1} \cdot [(d-1)\lambda + (d-2)(i\lambda\beta_{j+k} + k\lambda\beta_{j+k-1})] \\ &+ \mathbf{1}_{t=i+k-2} \cdot [(i+k-1)k\lambda] + \mathbf{1}_{t=i+k-1} \cdot [(d-i-k)k\lambda + (i+k)j\lambda] \\ &+ \mathbf{1}_{t=i+k} \cdot [(d-1-i-k)j\lambda + (i+k)] + \mathbf{1}_{t=i+k+1} \cdot [d-1-i-k] \Big). \end{aligned} \tag{5.5}$$

Proof. We refer to edges of $F(\mathbf{e})$ that are incident to the left endpoint of $\mathbf{e}$ as left edges, we define right edges similarly, and we refer to edges of $F(\mathbf{e})$ that are in a triangle of $F(\mathbf{e}) \cup \mathbf{e}$ as triangle edges.

To compute the $\gamma^e_{i,j,k}$'s we consider the following disjoint events: 1) a left edge is in the matching 2) $\mathbf{e}$ is in the matching 3) no left edge or triangle edge is in the matching 4) no left edge is in the matching, but a right triangle edge is in the matching. These events happen with probability $\frac{i\lambda\beta_{j+k}+k\lambda\beta_{j+k-1}}{\lambda+M_{i,j,k}}$, $\frac{\lambda}{\lambda+M_{i,j,k}}$, $\frac{\beta_j}{\lambda+M_{i,j,k}}$, and $\frac{k\lambda}{\lambda+M_{i,j,k}}$ respectively (by Lemma 5.3). Under these events the number of uncovered neighbours of $\mathbf{e}$ is $1, 0, i+k$, and $i+k-1$ respectively. This gives (5.4).

To compute the $\gamma^f_{i,j,k}$'s we refine the above events to include the possible choices of $\mathbf{f}$: $\mathbf{f}$ can be an edge outside $F(\mathbf{e})$ with probability $(d-1-i-k)/(d-1)$; a non-triangle edge in $F(\mathbf{e})$ with probability $i/(d-1)$; a triangle edge in $F(\mathbf{e})$ with probability $k/(d-1)$. If a left edge is in the matching we choose it as $\mathbf{f}$ with probability $1/(d-1)$, and if a right triangle edge is in the matching we choose $\mathbf{f}$ adjacent to it with probability $1/(d-1)$. Computing the number of uncovered neighbours of $\mathbf{f}$ in each case gives (5.5). □

We are now in a position to build our extra constraints on the $q(i,j,k)'s$. Let $\mathbf{e}$ be an edge of G chosen uniformly at random and let $\mathbf{s}$ be a uniformly chosen side left/right of $\mathbf{e}$. Then, by conditioning on the free neighbourhood of $\mathbf{e}$ and the value of $\mathbf{s}$, for each $t \in \{0,\ldots,d-1\}$ we have,

$$\mathbb{P}[\mathbf{e} \text{ has } t \text{ uncovered nbrs on side } \mathbf{s}] = \sum_{i,j,k} q(i,j,k)\frac{1}{2}\left[\gamma^{e}_{i,j,k}(t) + \gamma^{e}_{j,i,k}(t)\right].$$

Let $\mathbf{f}$ be a neighbouring edge of $\mathbf{e}$ from the side $\mathbf{s}$ chosen uniformly at random and let $\mathbf{h}$ be the side of $\mathbf{f}$ which contains $\mathbf{e}$. Note that since G is d-regular, $(\mathbf{f},\mathbf{h})$ is also an edge chosen uniformly at random from G with a uniformly chosen side. It follows that

$$\mathbb{P}[\mathbf{e} \text{ has } t \text{ uncovered nbrs on side } \mathbf{s}] = \mathbb{P}[\mathbf{f} \text{ has } t \text{ uncovered nbrs on side } \mathbf{h}].$$

Again by conditioning on the free neighbourhood of $\mathbf{e}$ and the value of $\mathbf{s}$ we have,

$$\mathbb{P}[\mathbf{f} \text{ has } t \text{ uncovered nbrs on side } \mathbf{h}] = \sum_{i,j,k} q(i,j,k)\frac{1}{2}\left[\gamma^{f}_{i,j,k}(t) + \gamma^{f}_{j,i,k}(t)\right].$$

It follows that for each $t \in \{0,\ldots,d-1\}$ we have the constraint

$$\sum_{i,j,k} q(i,j,k)\frac{1}{2}\left[\gamma^{f}_{i,j,k}(t) + \gamma^{f}_{j,i,k}(t) - \gamma^{e}_{i,j,k}(t) - \gamma^{e}_{j,i,k}(t)\right] = 0.$$

This leads us to consider the following linear program.

$$\begin{aligned}
\max\ & \sum_{i,j,k} q(i,j,k)\overline{\alpha}(i,j,k) \\
\text{s.t.}\ & \sum_{i,j,k} q(i,j,k) = 1 \\
& \sum_{i,j,k} q(i,j,k)\frac{1}{2}\left[\gamma^{f}_{i,j,k}(t) + \gamma^{f}_{j,i,k}(t) - \gamma^{e}_{i,j,k}(t) - \gamma^{e}_{j,i,k}(t)\right] = 0 \ \forall\, t = 0,\ldots,d-2 \\
& q(i,j,k) \geq 0 \ \forall\ i,j,k\,.
\end{aligned}$$

Note that we omit the $t = d-1$ constraint since it is redundant. Let us denote the optimal solution to this linear program by α^* and note that $\alpha_G(\lambda) \leq \alpha^*$ for all d-regular graphs G. We expect the $q(i,j,k)$ distribution arising from $K_{d,d}$ (or a disjoint union of $K_{d,d}$'s) to be optimal. The following observation will be useful in guiding us toward a solution of our linear program.

Claim 5.5. *Suppose that the distribution $q(i,j,k)$ is supported only on triples with $i = j$ and $k = 0$. Then G must be a disjoint union of $K_{d,d}$'s.*

Proof. First note that G must be triangle free, else we could pick an edge in $k > 0$ triangles and the empty matching so that $q(d-1-k, d-1-k, k) > 0$. Let C be a connected component of G and let $e = \{u, v\}$ be an edge of C (where we suppose u is the 'left' vertex of e). Suppose that there exists an edge f in G that is incident to a vertex in $N(u)$ and not incident to any vertex in $N(v)$ (note that since C is triangle free, $N(u)$ and $N(v)$ are disjoint). Picking the edge e and the matching consisting only of the edge f would then show that $q(i, d-1, 0) > 0$ where $i = d-2$ or $d-3$. It follows that all edges that are incident to a vertex in $N(u)$ must also be incident to a vertex in $N(v)$. Since C is d-regular it must be the case that all edges between $N(u) \setminus \{v\}$ and $N(v) \setminus \{u\}$ are present and so C is isomorphic to $K_{d,d}$. □

The dual linear program is

$$\min \Lambda$$

$$\text{s.t. } \Lambda - \overline{\alpha}(i,j,k) + \sum_{t=0}^{d-2} \Lambda_t \frac{1}{2}\left[\gamma^f_{i,j,k}(t) + \gamma^f_{j,i,k}(t) - \gamma^e_{i,j,k}(t) - \gamma^e_{j,i,k}(t)\right] \geq 0 \ \forall\ i,j,k$$

where $\Lambda, \Lambda_0, \ldots, \Lambda_{d-2}$ are the dual variables. To show that $K_{d,d}$ is optimal, we find values for the dual variables so that the dual constraints hold with $\Lambda = \alpha^M_{K_{d,d}}$. To find such Λ_t's, we solve the system of equations generated by setting equality in the constraints corresponding to $i = j$ and $k = 0$.

With this choice of Λ_t's, we start by simplifying the form of the dual constraints with a substitution coming from equality in the $(i,j,k) = (0,0,0)$ constraint. The $(0,0,0)$ dual constraint has the simple form

$$\Lambda_0 - \Lambda_1 = \alpha^M_{K_{d,d}} .$$

Moreover, observe that from the $\mathbf{1}_{t=0}$ and $\mathbf{1}_{t=1}$ terms in $\gamma^e_{i,j,k}(t)$ and $\gamma^f_{i,j,k}(t)$, every dual constraint contains the term

$$\left[\overline{\alpha}(i,j,k) - \frac{\lambda}{(\lambda + M_{i,j,k})}\right](\Lambda_0 - \Lambda_1) = \left[\overline{\alpha}(i,j,k) - \frac{\lambda}{(\lambda + M_{i,j,k})}\right]\alpha^M_{K_{d,d}}\,.$$

With this simplification, we multiply through by $2(d-1)(\lambda + M_{i,j,k})$ and expand $\overline{\alpha}(i,j,k)$ terms to obtain the following form of the dual constraints (recall that we use β_t to denote $1 + t\lambda$).

$$\begin{aligned}\alpha^M_{K_{d,d}}\big[\lambda M'_{i,j,k} + 2(d-1)M_{i,j,k}\big] - \lambda M'_{i,j,k}& \\ + \Lambda_{i+k-2}\cdot(i+k-1)k\lambda& \\ + \Lambda_{i+k-1}\cdot[(d-i-k)k\lambda + (i+k)j\lambda - (d-1)k\lambda]& \\ + \Lambda_{i+k}\cdot[(d-1-i-k)j\lambda + i + k - (d-1)\beta_j]& \\ + \Lambda_{i+k+1}\cdot(d-1-i-k)& \\ + \Lambda_{j+k-2}\cdot(j+k-1)k\lambda& \\ + \Lambda_{j+k-1}\cdot[(d-j-k)k\lambda + (j+k)i\lambda - (d-1)k\lambda]& \\ + \Lambda_{j+k}\cdot[(d-1-j-k)i\lambda + j + k - (d-1)\beta_i]& \\ + \Lambda_{j+k+1}\cdot(d-1-j-k) \geq 0\,.&\end{aligned} \tag{5.6}$$

The $(i,i,0)$ equality constraints now read

$$\alpha^M_{K_{d,d}}\beta_i\big(\beta_i + \tfrac{i\lambda}{d-1}\big) - \tfrac{i\lambda\beta_i}{d-1} + \Lambda_{i-1}\tfrac{i^2\lambda}{d-1} - \Lambda_i\tfrac{d-1-i+i^2\lambda}{d-1} + \Lambda_{i+1}\tfrac{d-1-i}{d-1} = 0\,. \tag{5.7}$$

With this we can write Λ_{i+k+1} in terms of Λ_{i+k} and Λ_{i+k-1}, and similarly for Λ_{j+k+1}. Substituting this into (5.6) and dividing by λ we derive the simplified form of the dual constraints:

$$\begin{aligned}\lambda\big[(i-j)^2 + 2k\big](1 - d\alpha^M_{K_{d,d}})& \\ + \Lambda_{i+k-2}(i+k-1)k + \Lambda_{i+k-1}[k + (i+k)(j-i-2k)]& \\ + \Lambda_{i+k}(i+k)(i+k-j)& \\ + \Lambda_{j+k-2}(j+k-1)k + \Lambda_{j+k-1}[k + (j+k)(i-j-2k)]& \\ + \Lambda_{j+k}(j+k)(j+k-i) \geq 0\,.&\end{aligned} \tag{5.8}$$

Write $L(i,j,k)$ for the LHS of this inequality.

The constraint for $t = d-1$ was omitted, but we nonetheless introduce $\Lambda_{d-1} := 0$ in order to simplify the presentation of the argument. The $(d-1, d-1, 0)$ equality constraint gives Λ_{d-2} directly:

$$\Lambda_{d-2} = \frac{1}{(d-1)\lambda}\left[\lambda + (d-1)\lambda^2 - \alpha^M_{K_{d,d}}\beta_{d-1}\beta_d\right].$$

With Λ_{d-1}, Λ_{d-2}, and the recurrence relation (5.7) the dual variables are fully determined.

We now reduce the problem of showing that the dual constraints (5.8) corresponding to triples (i,j,k) with $k > 0$ or $i \neq j$ hold with strict inequality to showing that a particular function is increasing. We go on to prove this fact in Claims 5.6 and 5.7.

Putting $k = 0$ into (5.8) gives:

$$\begin{aligned}\frac{L(i,j,0)}{(j-i)} &= \lambda(j-i)(1 - d\alpha^M_{K_{d,d}}) + i\Lambda_{i-1} - i\Lambda_i - j\Lambda_{j-1} + j\Lambda_j \\ &= F_d(j) - F_d(i)\end{aligned}$$

where

$$F_d(t) := t\left[\lambda(1 - d\alpha^M_{K_{d,d}}) + \Lambda_t - \Lambda_{t-1}\right]. \tag{5.9}$$

From (5.8) we obtain

$$\begin{aligned}L(i-1, j-1, k+1) - L(i,j,k) =& F_d(i+k) - F_d(i+k-1) \\ &+ F_d(j+k) - F_d(j+k-1).\end{aligned}$$

Therefore, if $F_d(t)$ is strictly increasing, we have $L(i,j,0) > 0$ for $i \neq j$, and $L(i-1,j-1,k+1) > L(i,j,k) > \cdots > L(i+k, j+k, 0) \geq 0$.

We first find an explicit expression for $F_d(t)$. Recall that we write $M_{K_{t,t}}$ for the matching polynomial of the graph $K_{t,t}$, that is $M_{K_{t,t}}(\lambda) = \sum_{i=0}^{t}\binom{t}{i}^2 i!\lambda^t$.

Claim 5.6. *For all $d \geq 2$ and $1 \leq t \leq d-1$,*

$$F_d(t) = \frac{t(d-1)}{M_{K_{d,d}}}\sum_{\ell=t-1}^{d-2}\frac{(d-1-t)!}{(\ell+1-t)!}\lambda^{d-\ell}M_{K_{\ell,\ell}}. \tag{5.10}$$

Proof. We will use the following two facts:

$$M_{K_{d,d}} - \beta_{2d-1}M_{K_{d-1,d-1}} + (d-1)^2\lambda^2 M_{K_{d-2,d-2}} = 0 \tag{5.11}$$

$$\alpha^M_{K_{d,d}} = \frac{\lambda M_{K_{d-1,d-1}}}{M_{K_{d,d}}}. \tag{5.12}$$

The first is a Laguerre polynomial identity, verifiable by hand; the second is also a short calculation. The equality dual constraint (5.7) implies:

$$(d-1-t)F_d(t+1) = (t+1)[t\lambda F_d(t) + (d-1)\lambda - (d-1)\alpha^M_{K_{d,d}}\beta_{d+t}]\,.$$

We first show that the right hand side of (5.10) satisfies the above recurrence relation. Using (5.12) this amounts to showing that the following expression is equal to zero for all $d \geq 2$ and $1 \leq t \leq d-1$:

$$\begin{aligned}\Phi_d(t) :=&(d-1-t)!\left(\sum_{\ell=t}^{d-2}\frac{\lambda^{d-\ell}M_{K_{\ell,\ell}}}{(\ell-t)!} - t^2\sum_{\ell=t-1}^{d-2}\frac{\lambda^{d+1-\ell}M_{K_{\ell,\ell}}}{(\ell+1-t)!}\right)\\ &- \lambda(M_{K_{d,d}} - \beta_{d+t}M_{K_{d-1,d-1}})\,.\end{aligned}$$

We proceed by induction on d. Note that when $d=2$, $\Phi_2(1)$ is easily verified to be zero. Note that

$$\Phi_{d+1}(t) = \lambda\Big((d-t)\Phi_d(t) - M_{K_{d+1,d+1}} + \beta_{2d+1}M_{K_{d,d}} - d^2\lambda^2 M_{K_{d-1,d-1}}\Big)\,.$$

By the induction hypothesis and (5.11) the result follows. To complete the proof of the claim it suffices to show that (5.10) holds for $t = d-1$. Recalling that

$$\begin{aligned}\Lambda_{d-1} &= 0\\ \Lambda_{d-2} &= \frac{1}{d-1} + \lambda - \frac{\alpha^M_{K_{d,d}}}{(d-1)\lambda}\beta_d\beta_{d-1}\,,\end{aligned}$$

substituting into (5.9), and using (5.11) and (5.12) we have

$$\begin{aligned}F_d(d-1) &= (d-1)\left[\lambda(1-d\alpha^M_{K_{d,d}}) - \frac{1}{d-1} - \lambda + \frac{\alpha^M_{K_{d,d}}}{(d-1)\lambda}\beta_d\beta_{d-1}\right]\\ &= \frac{\alpha^M_{K_{d,d}}}{\lambda}\beta_{2d-1} - 1\\ &= \frac{1}{M_{K_{d,d}}}\left[\beta_{2d-1}M_{K_{d-1,d-1}} - M_{K_{d,d}}\right]\\ &= \frac{(d-1)^2\lambda^2 M_{K_{d-2,d-2}}}{M_{K_{d,d}}}\,,\end{aligned}$$

verifying (5.10) for $t = d-1$. $\square$

Using Claim 5.6 we prove the following.

Claim 5.7. *$F_d(t)$ is strictly increasing as a function of t.*

Proof. To prove that $F_d(t)$ is increasing, we show that

$$\begin{aligned} R_d(t) &:= \frac{M_{K_{d,d}}}{(d-1)} \cdot \frac{F_d(t+1) - F_d(t)}{(d-2-t)!} \\ &= (t+1) \sum_{\ell=t}^{d-2} \frac{\lambda^{d-\ell}}{(\ell-t)!} M_{K_{\ell,\ell}} - t(d-1-t) \sum_{\ell=t-1}^{d-2} \frac{\lambda^{d-\ell}}{(\ell+1-t)!} M_{K_{\ell,\ell}} \end{aligned}$$

is positive for each t with $1 \leq t \leq d-2$. We do this by fixing t and inducting on d from $t+2$ upwards. A useful inequality will be $M_{K_{t,t}} > t\lambda M_{K_{t-1,t-1}}$ which comes from only counting matchings of $K_{t,t}$ that use a specific vertex. Iterating this inequality we obtain

$$M_{K_{t,t}} > \frac{t!}{\ell!} \lambda^{t-\ell} M_{K_{\ell,\ell}} \text{ for } 0 \leq \ell \leq t-1\,. \tag{5.13}$$

For the base case of our induction, $d = t+2$, we have $R_d(d-2) = \lambda^2 \left[M_{K_{d-2,d-2}} - (d-2)\lambda M_{K_{d-3,d-3}}\right]$ which by (5.13) is positive.

For the inductive step we have

$$R_{d+1}(t) = \lambda \left[R_d(t) + \frac{\lambda}{(d-1-t)!} M_{K_{d-1,d-1}} - \sum_{\ell=t-1}^{d-2} \frac{t\lambda^{d-\ell}}{(\ell-t+1)!} M_{K_{\ell,\ell}} \right]$$

and so it is sufficient to show

$$\sum_{\ell=t-1}^{d-2} \frac{t\lambda^{d-\ell}}{(\ell+1-t)!} M_{K_{\ell,\ell}} < \frac{\lambda}{(d-1-t)!} M_{K_{d-1,d-1}}\,. \tag{5.14}$$

We use the inequality (5.13) in each term of the sum to see that the LHS of (5.14) is less than

$$\sum_{\ell=t-1}^{d-2} \frac{t\ell!\lambda}{(\ell+1-t)!(d-1)!} M_{K_{d-1,d-1}}$$

and so

$$\begin{aligned}\sum_{\ell=t-1}^{d-2}\frac{t\lambda^{d-\ell}}{(\ell+1-t)!}M_{K_{\ell,\ell}} &< \sum_{\ell=t-1}^{d-2}\frac{t\ell!\lambda}{(\ell+1-t)!(d-1)!}M_{K_{d-1,d-1}}\\ &= \frac{\lambda M_{K_{d-1,d-1}}}{(d-1-t)!}\cdot\sum_{\ell=t-1}^{d-2}\frac{t\ell!(d-1-t)!}{(\ell+1-t)!(d-1)!}\\ &= \frac{\lambda M_{K_{d-1,d-1}}}{(d-1-t)!}\cdot\binom{d-1}{t}^{-1}\cdot\sum_{\ell=t-1}^{d-2}\binom{\ell}{t-1}\\ &= \frac{\lambda M_{K_{d-1,d-1}}}{(d-1-t)!},\end{aligned}$$

therefore (5.14) holds as required. □

This completes the proof of dual feasibility and hence

$$\alpha^* \le \alpha^M_{K_{d,d}}(\lambda)$$

by weak duality (Theorem 1.3). Strict inequality in the dual constraints outside of the $(i,i,0)$ constraints implies, by complementary slackness (Theorem 1.4), that the support of any optimal solution in the primal is contained in the set of $(i,i,0)$ configurations. Theorem 5.1 follows since the distribution arising from $K_{d,d}$ is optimal and disjoint unions of $K_{d,d}$'s are the only graphs which induce a distribution supported on the set of $(i,i,0)$ configurations (Claim 5.5).

5.3 Matchings of a Given Size

In this section we prove Theorem 5.2 and show how the Asymptotic Upper Matching Conjecture of Friedland, Krop, Lundow, and Markström [43] follows as a corollary.

Recall that if H is a random matching drawn from the monomer-dimer model on G at fugacity λ then we have

$$\mathbb{E}|H| = \frac{\lambda M_G'(\lambda)}{M_G(\lambda)}. \tag{5.15}$$

By a calculation identical to the one given in the proof of Theorem 4.11 from the previous chapter we also have the following expression for the variance:

$$\mathrm{var}|H| = \lambda \cdot \frac{d}{d\lambda}\mathbb{E}|H|\,. \tag{5.16}$$

We will make use of the celebrated Heilman-Lieb Theorem [52] which asserts that for any graph G, the roots of the matching polynomial M_G are all real. In particular, if $\nu = \nu(G)$ is the maximum number of edges in a matching in G, then we can write

$$M_G(\lambda) = \prod_{i=1}^{\nu}(1+\lambda r_i), \tag{5.17}$$

where all the r_i's are real.

Moreover, if a probability distribution has a generating function with the form (5.17) (properly normalised), then it must be the distribution of ν independent Bernoulli random variables with parameters $\frac{\lambda r_i}{1+\lambda r_i}$, $i = 1, \ldots \nu$. This gives a tremendous amount of structure to the distribution of $|H|$, the size of a random matching drawn from the monomer-dimer model on G: it has the distribution of the sum of independent Bernoulli random variables (or in other words, the sequence $m_0(G), \ldots m_\nu(G)$ is a *Polya frequency sequence* [71]). In particular we have that the sequence $m_0(G), \ldots, m_\nu(G)$ is log-concave, that is

$$m_k(G)^2 \geq m_{k-1}(G)m_{k+1}(G) \;\text{ for }\; k = 1, \ldots, \nu - 1\,.$$

This structure will allow us to make use of the following probabilistic result of Darroch.

Lemma 5.8 ([22] Theorem 4)**.** *Let Z be the sum of n independent (but not necessarily identically distributed) Bernoulli random variables with $\mathbb{E}Z = \mu$. Let m be the mode of Z, then*

$$|\mu - m| < 1.$$

We start with a bound on the variance of the size of a random matching in a graph G.

Lemma 5.9. *Let G be a graph and let H be a random matching drawn from the monomer-dimer model on G at fugacity $\lambda > 0$, then*

$$var|H| \le \mathbb{E}|H| .$$

Proof. Let $\nu := \nu(G)$. Recall that by the Heilman-Lieb Theorem we may write

$$M_G(\lambda) = \prod_{i=1}^{\nu}(1 + \lambda r_i) ,$$

where the r_i are real. Moreover since the coefficients of M_G are positive, all roots of M_G must be negative and so the r_i are positive. By (5.15) and (5.16) we then have

$$\mathbb{E}|H| = \sum_{i=1}^{\nu} \frac{\lambda r_i}{1 + \lambda r_i} \quad \text{and} \quad \text{var}|H| = \sum_{i=1}^{\nu} \frac{\lambda r_i}{(1 + \lambda r_i)^2}.$$

Since each of the denominators is greater than 1 the result is clear. □

Lemma 5.10. *Let G be a graph. Then for all $1 \le k \le \nu(G)$, there exists a λ so that*

$$\frac{m_k(G)\lambda^k}{M_G(\lambda)} = \mathbb{P}_{G,\lambda}[|H| = k] > \frac{1}{7\sqrt{k}} .$$

Proof. Let $\nu = \nu(G)$. Choose λ so that $\mathbb{P}[|H| = k-1] = \mathbb{P}[|H| = k]$. Since the probability distribution of $|H|$ is log-concave, it follows that $\mathbb{P}[|H| = k]$ is maximal (i.e. both $k-1$ and k are modal values of $|H|$). Darroch's rule (Lemma 5.8) then implies that $k - 1 < \mathbb{E}|H| < k$. It then follows from Lemma 5.9 that $\text{var}|H| \le k$. By Chebyshev's inequality, with probability at least 2/3 the size of H is one of at most $\left\lceil 2\sqrt{3}\sqrt{\text{var}|H|}\right\rceil \le \left\lceil 2\sqrt{3k}\right\rceil$ values. It follows that

$$\mathbb{P}[|H| = k] \ge \frac{2}{3\left\lceil 2\sqrt{3k}\right\rceil} \ge \frac{1}{7\sqrt{k}} .$$

□

Proof of Theorem 5.2. Choosing λ according to Lemma 5.10 we have:

$$m_k(G)\lambda^k \le M_G(\lambda) \le M_{H_{d,n}}(\lambda) \le 7\sqrt{k} \cdot m_k(H_{d,n})\lambda^k ,$$

where for the second inequality we used Theorem 5.1. □

As a consequence, we prove the Asymptotic Upper Matching Conjecture [43]. To state the conjecture precisely we must first introduce some notation.

Fix d and consider an infinite sequence of d-regular graphs $\mathcal{G}_d = G_1, G_2, \ldots$ where $v(G_n) \to \infty$. For any $\rho \in [0,1]$, the ρ-monomer entropy is

$$h_{\mathcal{G}_d}(\rho) = \sup_{\{k_n\}} \limsup_{n\to\infty} \frac{\log m_{k_n}(G_n)}{v(G_n)},$$

where the supremum is taken over all integer sequences $\{k_n\}$ with $\frac{2k_n}{v(G_n)} \to \rho$. Let $\mathcal{H}_d$ denote the sequence $H_{d,2d}, H_{d,4d}, H_{d,6d}, \ldots$. The following theorem is conjectured in [43] (Conjecture 7.2) where it is referred to as the Asymptotic Upper Matching Conjecture.

Theorem 5.11. *$\mathcal{G}_d = G_1, G_2, \ldots$ be a sequence of d-regular graphs where $v(G_n) \to \infty$. Then for any $\rho \in [0,1]$ we have*

$$h_{\mathcal{G}_d}(\rho) \le h_{\mathcal{H}_d}(\rho).$$

In fact the conjecture is made for sequences of bipartite regular graphs and the authors remark that it is plausible the bipartite restriction is not necessary. We show that this is indeed the case.

Proof. Assume $\rho > 0$ since for $\rho = 0$ the result is trivially true. Let $\{k_n\}$ be a sequence of integers with $\frac{2k_n}{v(G_n)} \to \rho$. Assume for the sake of contradiction that $\limsup \frac{\log m_{k_n}(G_n)}{v(G_n)} > h_{\mathcal{H}_d}(\rho) + \epsilon$ for some $\epsilon > 0$. Take N large enough that for all $n_1 \ge N$, divisible by $2d$, $\frac{\log m_{\lfloor \rho n_1/2 \rfloor}(H_{d,n_1})}{n_1} < h_{\mathcal{H}_d}(\rho) + \epsilon/2$. Now take some n with $v(G_n) \ge N$ and $\frac{\log m_{k_n}(G_n)}{v(G_n)} > h_{\mathcal{H}_d}(\rho) + \epsilon$, and let $n_1 = 2d \cdot \lceil v(G_n)/(2d) \rceil$. By Lemma 5.10, we choose λ so that $m_{\lfloor \rho n_1/2 \rfloor}(H_{d,n_1}) \lambda^{\lfloor \rho n_1/2 \rfloor} > \frac{1}{7\sqrt{\rho n_1/2}} M_{H_{d,n_1}}(\lambda)$. Note that since $\rho > 0$, such λ is bounded away from 0 as $n_1 \to \infty$. Then we have

$$\begin{aligned}\frac{\log M_{G_n}(\lambda)}{v(G_n)} \ge \frac{\log(m_{k_n}(G_n)\lambda^{k_n})}{v(G_n)} &> \frac{k_n}{v(G_n)} \log \lambda + h_{\mathcal{H}_d}(\rho) + \epsilon \\ &= \frac{\rho}{2} \log \lambda + h_{\mathcal{H}_d}(\rho) + \epsilon + o(1) \text{ as } n \to \infty,\end{aligned}$$

and

$$\begin{aligned}
\frac{\log M_{K_{d,d}}(\lambda)}{2d} &= \frac{\log M_{H_{d,n_1}}(\lambda)}{n_1} \\
&< \frac{\log\left(7\sqrt{\rho n_1/2}\cdot m_{\lfloor \rho n_1/2\rfloor}(H_{d,n_1})\lambda^{\lfloor \rho n_1/2\rfloor}\right)}{n_1} \\
&< \frac{\log(7\sqrt{\rho n_1/2})}{n_1} + \frac{\lfloor \rho n_1/2\rfloor}{n_1}\log\lambda + h_{\mathcal{H}_d}(\rho) + \epsilon/2 \\
&= \frac{\rho}{2}\log\lambda + h_{\mathcal{H}_d}(\rho) + \epsilon/2 + o(1)\,.
\end{aligned}$$

However, this contradicts Theorem 5.1. □

Although the Upper Matching Conjecture remains open, we venture to make the following even stronger conjecture.

Conjecture 5.12. *Let G be a d-regular, n-vertex graph where $2d$ divides n. Then for all k, the ratio $\frac{m_k(G)}{m_{k-1}(G)}$ is maximised by $H_{d,n}$.*

This conjecture is in fact strong enough to imply Theorem 5.1. The relation to the work here is that Conjecture 5.12 can be stated as follows: the expected number of edges incident to a uniformly random matching of size k is minimised by $H_{d,n}$. Theorem 5.1 shows that such a statement is true when the random matching is chosen according to the monomer-dimer model instead of uniformly over those of a given size.

Bibliography

[1] M. Ajtai, J. Komlós, and E. Szemerédi. A note on Ramsey numbers. *J. Combin. Theory Ser. A*, 29(3):354–360, 1980.

[2] P. Allen, G. Brightwell, and J. Skokan. Ramsey-goodness—and otherwise. *Combinatorica*, 33(2):125–160, 2013.

[3] N. Alon. Independent sets in regular graphs and sum-free subsets of finite groups. *Israel J. Math.*, 73(2):247–256, 1991.

[4] F. S. Benevides, T. Łuczak, A. Scott, J. Skokan, and M. White. Monochromatic cycles in 2-coloured graphs. *Combin. Probab. Comput.*, 21(1-2):57–87, 2012.

[5] F. S. Benevides and J. Skokan. The 3-colored Ramsey number of even cycles. *J. Combin. Theory Ser. B*, 99(4):690–708, 2009.

[6] T. Bohman. The triangle-free process. *Adv. Math.*, 221(5):1653–1677, 2009.

[7] T. Bohman and P. Keevash. Dynamic concentration of the triangle-free process. In *The Seventh European Conference on Combinatorics, Graph Theory and Applications*, volume 16 of *CRM Series*, pages 489–495. Ed. Norm., Pisa, 2013.

[8] B. Bollobás. Three-graphs without two triples whose symmetric difference is contained in a third. *Discrete Math.*, 8:21–24, 1974.

[9] J. A. Bondy. Pancyclic graphs. I. *J. Combinatorial Theory Ser. B*, 11:80–84, 1971.

[10] J. A. Bondy and P. Erdős. Ramsey numbers for cycles in graphs. *J. Combinatorial Theory Ser. B*, 14:46–54, 1973.

[11] R. C. Bose and S. Chowla. Theorems in the additive theory of numbers. *Comment. Math. Helv.*, 37:141–147, 1962/1963.

[12] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004.

[13] L. M. Brègman. Certain properties of nonnegative matrices and their permanents. *Dokl. Akad. Nauk SSSR*, 211:27–30, 1973.

[14] S. A. Burr and P. Erdős. On the magnitude of generalized Ramsey numbers for graphs. pages 215–240. Colloq. Math. Soc. János Bolyai, Vol. 10, 1975.

[15] T. Carroll, D. Galvin, and P. Tetali. Matchings and independent sets of a fixed size in regular graphs. *J. Combin. Theory Ser. A*, 116(7):1219–1227, 2009.

[16] C. Chvátal, V. Rödl, E. Szemerédi, and W. T. Trotter, Jr. The Ramsey number of a graph with bounded maximum degree. *J. Combin. Theory Ser. B*, 34(3):239–243, 1983.

[17] L. H. Clark, J. C. George, and T. D. Porter. On the number of 1-factors in the n-cube. In *Proceedings of the Twenty-eighth Southeastern International Conference on Combinatorics, Graph Theory and Computing (Boca Raton, FL, 1997)*, volume 127, pages 67–69, 1997.

[18] E. J. Cockayne and P. J. Lorimer. The Ramsey number for stripes. *J. Austral. Math. Soc.*, 19:252–256, 1975.

[19] D. Conlon. A new upper bound for diagonal Ramsey numbers. *Ann. of Math. (2)*, 170(2):941–960, 2009.

[20] J. Cooper, K. Dutta, and D. Mubayi. Counting independent sets in hypergraphs. *Combin. Probab. Comput.*, 23(4):539–550, 2014.

[21] J. Cooper and D. Mubayi. Counting independent sets in triangle-free graphs. *Proc. Amer. Math. Soc.*, 142(10):3325–3334, 2014.

[22] J. N. Darroch. On the distribution of the number of successes in independent trials. *Ann. Math. Statist.*, 35:1317–1321, 1964.

[23] E. Davies, M. Jenssen, W. Perkins, and B. Roberts. Extremes of the internal energy of the potts model on cubic graphs. *Random Structures Algorithms*, To appear.

[24] E. Davies, M. Jenssen, W. Perkins, and B. Roberts. Independent sets, matchings, and occupancy fractions. *J. London Math. Soc.*, To appear.

[25] E. Davies, M. Jenssen, W. Perkins, and B. Roberts. On the average size of independent sets in triangle-free graphs. *Proc. Amer. Math. Soc.*, To appear.

[26] A. N. Day and J. R. Johnson. Multicolour Ramsey numbers of odd cycles. *J. Combin. Theory Ser. B*, 124:56–63, 2017.

[27] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53:292–294, 1947.

[28] P. Erdős, R. J. Faudree, C. C. Rousseau, and R. H. Schelp. Generalized Ramsey theory for multiple colors. *J. Combinatorial Theory Ser. B*, 20(3):250–264, 1976.

[29] P. Erdős and M. Simonovits. A limit theorem in graph theory. *Studia Sci. Math. Hungar*, 1:51–57, 1966.

[30] P. Erdős and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc*, 52(1087-1091):1, 1946.

[31] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.

[32] P. Erdős and T. Gallai. On maximal paths and circuits of graphs. *Acta Math. Acad. Sci. Hungar*, 10:337–356 (unbound insert), 1959.

[33] P. Erdős and R. L. Graham. On partition theorems for finite graphs. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vol. I*, pages 515–527. Colloq. Math. Soc. János Bolyai, Vol. 10. North-Holland, Amsterdam, 1975.

[34] R. J. Faudree and R. H. Schelp. All Ramsey numbers for cycles in graphs. *Discrete Math.*, 8:313–329, 1974.

[35] D. Ferguson. The Ramsey number of mixed-parity cycles i. *arXiv preprint arXiv:1508.07154*, 2015.

[36] D. Ferguson. The Ramsey number of mixed-parity cycles ii. *arXiv preprint arXiv:1508.07171*, 2015.

[37] D. Ferguson. The Ramsey number of mixed-parity cycles iii. *arXiv preprint arXiv:1508.07176*, 2015.

[38] A. Figaj and T. Łuczak. The Ramsey number for a triple of long even cycles. *J. Combin. Theory Ser. B*, 97(4):584–596, 2007.

[39] A. Figaj and T. Łuczak. The Ramsey number for a triple of large cycles. *ArXiv e-prints*, September 2007.

[40] G. Fiz Pontiveros, S. Griffiths, and R. Morris. The triangle-free process and $R(3,k)$. *Mem. Amer. Math. Soc.*, To appear.

[41] P. Frankl and Z. Füredi. Extremal problems and the Lagrange function for hypergraphs. *Bull. Inst. Math. Acad. Sinica*, 16(4):305–313, 1988.

[42] P. Frankl and Z. Füredi. Extremal problems whose solutions are the blowups of the small Witt-designs. *J. Combin. Theory Ser. A*, 52(1):129–147, 1989.

[43] S. Friedland, E. Krop, P. H. Lundow, and K. Markström. On the validations of the asymptotic matching conjectures. *J. Stat. Phys.*, 133(3):513–533, 2008.

[44] S. Friedland, E. Krop, and K. Markström. On the number of matchings in regular graphs. *Electron. J. Combin.*, 15(1):Research Paper 110, 28, 2008.

[45] D. Galvin. Three tutorial lectures on entropy and counting. *arXiv preprint arXiv:1406.7872*, 2014.

[46] D. Galvin and P. Tetali. On weighted graph homomorphisms. In *Graphs, morphisms and statistical physics*, volume 63 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 97–104. Amer. Math. Soc., Providence, RI, 2004.

[47] L. Gerencsér and A. Gyárfás. On Ramsey-type problems. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math*, 10:167–170, 1967.

[48] N. Graham and F. Harary. The number of perfect matchings in a hypercube. *Appl. Math. Lett.*, 1(1):45–48, 1988.

[49] C. M. Grinstead and S. M. Roberts. On the Ramsey numbers $R(3, 8)$ and $R(3, 9)$. *J. Combin. Theory Ser. B*, 33(1):27–51, 1982.

[50] A. Gyárfás, M. Ruszinkó, G. N. Sárközi, and E. Szemerédi. Three-color Ramsey numbers for paths. *Combinatorica*, 27(1):35–69, 2007.

[51] D. Hefetz and P. Keevash. A hypergraph Turán theorem via Lagrangians of intersecting families. *J. Combin. Theory Ser. A*, 120(8):2020–2038, 2013.

[52] O. J. Heilmann and E. H. Lieb. Theory of monomer-dimer systems. *Comm. Math. Phys.*, 25:190–232, 1972.

[53] D. Hilbert. Über die irreduzibilität ganzer rationaler funktionen mit ganzzahligen koeffizienten. *J. Reine Angew. Math.*, 110:104–129, 1892.

[54] L. Ilinca and J. Kahn. Asymptotics of the upper matching conjecture. *J. Combin. Theory Ser. A*, 120(5):976–983, 2013.

[55] J. Kahn. An entropy approach to the hard-core model on bipartite graphs. *Combin. Probab. Comput.*, 10(3):219–237, 2001.

[56] G. O. H. Katona. Extremal problems for hypergraphs. pages 13–42. Math. Centre Tracts, No. 56, 1974.

[57] G. O. H. Katona, T. Nemetz, and M. Simonovits. On a problem of Turán in the theory of graphs. *Mat. Lapok*, 15:228–238, 1964.

[58] P. Keevash. Hypergraph Turán problems. In *Surveys in combinatorics 2011*, volume 392 of *London Math. Soc. Lecture Note Ser.*, pages 83–139. Cambridge Univ. Press, Cambridge, 2011.

[59] P. Keevash. The existence of designs. *arXiv preprint arXiv:1401.3665*, 2014.

[60] J. H. Kim. The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$. *Random Structures Algorithms*, 7(3):173–207, 1995.

[61] Y. Kohayakawa, M. Simonovits, and J. Skokan. The 3-colored Ramsey number of odd cycles. In *Proceedings of GRACO2005*, volume 19 of *Electron. Notes Discrete Math.*, pages 397–402 (electronic), Amsterdam, 2005. Elsevier.

[62] J. Komlós and M. Simonovits. Szemerédi's regularity lemma and its applications in graph theory. In *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, volume 2 of *Bolyai Soc. Math. Stud.*, pages 295–352. János Bolyai Math. Soc., Budapest, 1996.

[63] C. Lee. Ramsey numbers of degenerate graphs. *Ann. of Math.*, 185:791–829, 2017.

[64] L. Lovász and M. D. Plummer. *Matching theory*, volume 121 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam; North-Holland Publishing Co., Amsterdam, 1986. Annals of Discrete Mathematics, 29.

[65] T. Łuczak. $R(C_n, C_n, C_n) \leq (4 + o(1))n$. *J. Combin. Theory Ser. B*, 75(2):174–187, 1999.

[66] T. Łuczak, M. Simonovits, and J. Skokan. On the multi-colored Ramsey numbers of cycles. *J. Graph Theory*, 69(2):169–175, 2012.

[67] W. Mantel. Problem 28. *Wiskundige Opgaven*, 10(60-61):320, 1907.

[68] T. S. Motzkin and E. G. Straus. Maxima for graphs and a new proof of a theorem of Turán. *Canad. J. Math.*, 17:533–540, 1965.

[69] P. R. J. Östergård and V. H. Pettersson. Enumerating perfect matchings in n-cubes. *Order*, 30(3):821–835, 2013.

[70] G. Perarnau and W. Perkins. Counting independent sets in cubic graphs of given girth. *arXiv preprint arXiv:1611.01474*, 2016.

[71] J. Pitman. Probabilistic bounds on the coefficients of polynomials with only real zeros. *J. Combin. Theory Ser. A*, 77(2):279–303, 1997.

[72] F. P. Ramsey. On a Problem of Formal Logic. *Proc. London Math. Soc.*, 30(1):264–286, 1930.

[73] A. A. Razborov. On 3-hypergraphs with forbidden 4-vertex configurations. *SIAM J. Discrete Math.*, 24(3):946–963, 2010.

[74] J. K. Roberts. Some properties of adsorbed films of oxygen on tungsten. *Proc. R. Soc. A*, 152(876):464–477, 1935.

[75] V. Rosta. On a Ramsey-type problem of J. A. Bondy and P. Erdős. I, II. *J. Combinatorial Theory Ser. B*, 15:94–104; ibid. 15 (1973), 105–120, 1973.

[76] W. Rudin. *Principles of mathematical analysis.* McGraw-Hill Book Co., New York-Auckland-Düsseldorf, third edition, 1976. International Series in Pure and Applied Mathematics.

[77] J. B. Shearer. A note on the independence number of triangle-free graphs. *Discrete Math.*, 46(1):83–87, 1983.

[78] A. F. Sidorenko. On the maximal number of edges in a homogeneous hypergraph that does not contain prohibited subgraphs. *Mat. Zametki*, 41(3):433–455, 459, 1987.

[79] A. F. Sidorenko. Solution of a problem of bollobás on 4-graphs. *Mat. Zametki*, 41:433–455, 1987.

[80] A. F. Sidorenko. Asymptotic solution for a new class of forbiddenr-graphs. *Combinatorica*, 9(2):207–215, 1989.

[81] J. Singer. A theorem in finite projective geometry and some applications to number theory. *Trans. Amer. Math. Soc.*, 43(3):377–385, 1938.

[82] J. Spencer. Ramsey's theorem—a new lower bound. *J. Combinatorial Theory Ser. A*, 18:108–115, 1975.

[83] E. Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.

[84] P. Turán. On an extremal problem in graph theory. *Mat. Fiz. Lapok*, 48(436-452):137, 1941.

[85] B. L. Van der Waerden. Beweis einer baudetschen vermutung. *Nieuw Arch. Wisk*, 15(2):212–216, 1927.

[86] Y. Zhao. The bipartite swapping trick on graph homomorphisms. *SIAM J. Discrete Math.*, 25(2):660–680, 2011.

Appendices

A Completing the Proof of Theorem 2.1

Here we verify the assertion made at the end of the proof of Theorem 2.1 from Chapter 2 which amounts to a simple yet tedious calculation. Recall that for fixed integer $r \geq 4$ we defined the following function on the interval $(0,1]$:

$$f_r(t) = 24t^3(\lfloor\mu\rfloor + (\mu - \lfloor\mu\rfloor)^3)\prod_{i=1}^{r-4}(1-it).$$

where $\mu = (1-(r-3)t)/4t$.

Lemma A. *The function $f_r(t)$ is maximised on the interval $(0,1]$ at $t = 1/(r+1)$ for $r = 4,5,6$ and 7.*

Proof. Let

$$g_r(t) := 24t^3\mu\prod_{i=1}^{r-4}(1-it) = 6t^2\prod_{i=1}^{r-3}(1-it)\,.$$

Since $\lfloor\mu\rfloor + (\mu - \lfloor\mu\rfloor)^3 \leq \mu$ we have that $f_r(t) \leq g_r(t)$ for $t \in (0,1]$. For $r = 4,5$ and 6 we show that $g_r(t)$ is increasing on the interval $(0, 1/(r+1)]$, that $g_r(1/(r+1)) = f_r(1/(r+1))$ and that $f_r(t)$ is decreasing on the interval $[1/(r+1), 1]$. This proves the proposition in the cases $r = 4,5$ and 6. The case $r = 7$ is slightly more delicate and we leave this case until last.
Let us show that the derivative of $g_r(t)$ is positive on $(0, 1/(r+1)]$:

$$\frac{d}{dt}g_r(t) = 6\left[2t - t^2\left(\frac{1}{1-t} + \frac{2}{1-2t} + \ldots + \frac{r-3}{1-(r-3)t}\right)\right]\prod_{i=1}^{r-3}(1-it)$$

which is positive if and only if

$$h_r(t) := 2 - t\left(\frac{1}{1-t} + \frac{2}{1-2t} + \ldots + \frac{r-3}{1-(r-3)t}\right) > 0$$

since $6t\prod_{i=1}^{r-3}(1-it) > 0$ on $(0, 1/(r+1)]$. Clearly $h_r(t)$ is decreasing on $(0, 1/(r+1)]$ and so it suffices to show that $h_r(1/(r+1)) > 0$. A simple calculation reveals that this is indeed the case for $r = 4, 5$ and 6. Showing that $g_r(1/(r+1)) = f_r(1/(r+1))$ is simply the observation that $\mu = 1$, an integer, when $t = 1/(r+1)$. Finally, if $t > 1/(r+1)$ then $\lfloor \mu \rfloor = 0$ and so $f_r(t) = \frac{3}{8}(1-(r-3)t)^3\prod_{i=1}^{r-4}(1-it)$, a decreasing function of t. This deals with the cases $r = 4, 5$ and 6.

For the case $r = 7$ we follow a similar procedure however it is not the case that $g_7(t)$ is increasing on $(0, 1/8)$. Instead we show that $g_7(t)$ is increasing on $(0, 1/12]$, that $g_7(1/12) = f_7(1/12)$, that $f_7(t)$ is convex on $(1/12, 1/8)$ with $f_7(1/8) > f_7(1/12)$ and finally that $f_7(t)$ is decreasing on $(1/8, 1]$. These observations are enough to deal with the case $r = 7$.

To show that $g_7(t)$ is increasing on $(0, 1/12]$ it suffices to observe, by the above, that $h_7(1/12) > 0$. Showing that $g_7(1/12) = f_7(1/12)$ is simply the observation that $\mu = 2$, an integer, when $r = 7, t = 1/12$. To show that $f_7(t)$ is convex on (1/12,1/8) we show that $f_7''(t) > 0$ on this interval. Note that for $r = 7, t \in (1/12, 1/8)$ we have that $\lfloor \mu \rfloor = 1$ so that $f_7(t)$ is a degree six polynomial so that by Taylor's theorem:

$$f_7''(t) = \sum_{k=0}^{4}\frac{1}{k!}f_7^{(k+2)}(1/8)(t-1/8)^k.$$

Elementary, yet tedious, calculation shows that $(-1)^k f_7^{(k+2)}(1/8) > 0$ for $k = 0, 1, 2, 3$ and 4. It follows therefore that $f_7''(t) > 0$ for $t \in (1/12, 1/8)$. A final calculation shows that $f_7(1/8) > f_7(1/12)$ and above we showed that $f_r(t)$ is decreasing on $(1/(r+1), 1]$ for $r \geq 4$. $\square$

B Proof of Lemmas 3.46 and 3.47

In this section we present the proofs of Lemmas 3.46 and 3.47 from Chapter 3. We use the following simple property of regular pairs which appears as Lemma 5 in [39].

Lemma B. *Let $1/m \ll \delta \ll d$ and let $G = (V_1, V_2)$ be a (δ, d)-super-regular pair with $|V_1| = |V_2| = m$. Then for each pair $u \in V_1$, $w \in V_2$, G contains a uw-path of length ℓ for each odd $3 \leq \ell \leq 2(1 - 5\delta)m$.*

Lemma 3.46. *Let $q \geq 4$ and suppose that $\frac{1}{m} \ll \delta \ll d$. Let F be a connected matching of order q such that every vertex of F is incident to a matching edge and let H be a (δ, d, m)-super-regular blow-up of F. Then the following holds:*

If $i, j \in V(F)$ and there is an ij-path of length r in F, then for every pair of vertices $u \in U_i$, $w \in U_j$, there exists a uw-path of length ℓ in H for each $3q \leq \ell \leq (1 - 6\delta)qm$ such that $\ell \equiv r \pmod 2$.

Proof. Take $i, j \in V(F)$ and let $u \in U_i$, $w \in U_j$. Let T be a spanning tree of F which includes every matching edge of F. Note that T contains a closed walk $W = y_0 \dots y_p$, where $y_1 = y_p = i$ and W covers each edge of T exactly twice, in particular $p = 2(q-1)$ (note that $q = v(F)$). Using basic properties of regular pairs we can find a path $\widetilde{W} = w_0 \dots w_p$ in H where $u = w_0$ and $w_t \in U_{y_t}$ for all t. Let $P = x_0 \dots x_r$ be a path of length r in F where $x_0 = i, x_r = j$. Again, using basic properties of regular pairs we can find a path $\widetilde{P} = v_0 \dots v_r$ in H where $v_0 = w_p$, $v_r = w$, $v_t \in U_{x_t}$ for all t and $\widetilde{P}$ intersects $\widetilde{W}$ only in the vertex w_p. Letting $Q = \widetilde{W}\widetilde{P}$, it follows that Q is a uw-path in H of length $r + p = r + 2(q-1) \equiv r \pmod 2$. Suppose that $\{a, b\}$ is a matching edge of F so that (U_a, U_b) is (δ, d)-super-regular in H. Note that Q visits each set U_i in H at most 3 times and so there exist $U'_a \subseteq U_a \backslash Q$, $U'_b \subseteq U_b \backslash Q$ such that $|U'_a| = |U'_b| = m - 3$. Note that (U'_a, U'_b) is certainly $(2\delta, d/2)$-super-regular by Fact 3.49. By construction, we may pick consecutive vertices w_t, w_{t+1} of $\widetilde{W}$ (and hence Q) such that $w_t \in U_a, w_{t+1} \in U_b$. By super-regularity we may then pick vertices $u_a \in N(w_{t+1}) \cap U'_a$, $u_b \in N(w_t) \cap U'_b$ such that $\{u_a, u_b\}$ is an edge of H. Applying Lemma B to (U'_a, U'_b) and vertices u_a, u_b, it follows that we can find a $q_t q_{t+1}$-path in H which intersects Q only at its endpoints and we can choose this path to have any odd length $1 \leq \ell \leq 2(1 - 5\delta)(m - 3) + 2$. Note that letting such a path replace the edge $\{q_t, q_{t+1}\}$ in Q does not change the parity of the length of Q. Applying the same argument to each matching edge of F we see that H

contains uw-paths of each length $r+2(q-1) \leq \ell \leq r+2(q-1)+\frac{q}{2}\cdot 2(1-6\delta)m$ for which $\ell \equiv r \pmod 2$. The result follows. $\square$

Lemma 3.47. *Let $q \geq 4$ and let $\frac{1}{m} \ll \delta \ll d$. Let F be an odd connected matching of order q and suppose that H is a (δ, m)-regular blow-up of F with minimum density d. Then H contains a cycle of length ℓ for each odd $3q \leq \ell \leq (1-6\delta)qm$*

Proof. Since F is non-bipartite it contains an odd cycle C. Since the largest matching in F has $q/2$ edges it follows that $|C| \leq q+1$. Let $T \subseteq F$ be a minimal tree that contains every matching edge of F. It is easy to show that T must have $< 2q$ vertices. Let W be a closed walk in T which traverses each edge of T precisely twice (so in particular W has even length). Since W and C must intersect, we can augment the walk W by C to obtain a closed walk $W' = x_1 \dots x_p x_1$ in F where p is odd and $p \leq 3q$ by the above. Note that by Facts 3.49 and 3.50, we can find $H' \subseteq H$ such that H' is a $(2\delta, d/2, (1-\delta)m)$-super-regular blowup of F. Let U_j denote the vertex class of H' corresponding to the vertex j in F for each $j \in V(F)$. Using basic properties of regular pairs, we can find an odd cycle $D = v_1 \dots v_p v_1$ in H' where $v_j \in U_{x_j}$ for all j.

Suppose that $\{a, b\}$ is a matching edge of F so that (U_a, U_b) is $(2\delta, d/2)$-super-regular. By construction, we may pick consecutive vertices v_t, v_{t+1} of D such that $v_t \in U_a, v_{t+1} \in U_b$. Note that D visits each set U_i in H at most 3 times. We may therefore apply Lemma B as we did in the proof of Lemma 3.46 to find a $v_t v_{t+1}$-path Q in H' such that Q intersects D only at its endpoints and we can choose Q to have any odd length $1 \leq \ell \leq 2(1-5\delta)[(1-\delta)m - 3] + 2$. Applying the same argument to each matching edge of F we see that H contains an odd cycle of each odd length $p \leq \ell \leq p + \frac{q}{2} \cdot 2(1-6\delta)m$. The result follows.

$\square$