Partitions of Combinatorial Structures

Viresh Patel

A thesis submitted for the degree of
Doctor of Philosophy

Department of Mathematics

London School of Economics
and Political Science

March 2009

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 solely 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 extremal, structural, and algorithmic problems involving the partitioning of combinatorial structures.

We begin by considering problems from the theory of graph cuts. It is well known that every graph has a cut containing at least half its edges. We conjecture that (except for one example), given any two graphs on the same vertex set, we can partition the vertices so that at least half the edges of each graph go across the partition. We give a simple algorithm that comes close to proving this conjecture. We also prove, using probabilistic methods, that the conjecture holds for certain classes of graphs.

We consider an analogue of the graph cut problem for posets and determine which graph cut results carry over to posets. We consider both extremal and algorithmic questions, and in particular, we show that the analogous maxcut problem for posets is polynomial-time solvable in contrast to the maxcut problem for graphs, which is NP-complete.

Another partitioning problem we consider is that of obtaining a regular partition (in the sense of the Szemerédi Regularity Lemma) for posets, where the partition respects the order of the poset. We prove the existence of such order-preserving, regular partitions for both the comparability graph and the covering graph of a poset, and go on to derive further properties of such partitions.

We give a new proof of an old result of Frankl and Füredi, which characterises all 3-uniform hypergraphs for which every set of 4 vertices spans exactly 0 or 2 edges. We use our new proof to derive a corresponding stability result.

We also look at questions concerning an analogue of the graph linear extension problem for posets.

Acknowledgements

Firstly, I would like to thank my supervisors, Graham Brightwell and Jan van den Heuvel, who have been a constant source of guidance and encouragement. Much of this work is the outcome of the many stimulating discussions we have had, and working with them has been both instructive and rewarding. I am also grateful to them for their valuable criticisms of my work, which have contributed greatly to its improvement.

I would also like to thank the members of the Maths Department at LSE for creating an informal, friendly atmosphere in which to work. In particular, I would like to thank Jozef Skokan, who introduced me to new problems; Peter Allen, who was always willing to share and discuss ideas; Raju Chintilapathy, who was always available for a friendly chat; and David Scott, Simon Jolly, and Jackie Everid for their administrative support.

I am grateful for the financial support provided by the EPSRC, the LSE, and the Mathematics Department.

Outside of LSE, I am grateful to Pooya Farshim for all his advice and for our many internet conversations, a most welcome distraction! I am also indebted to the many lecturers and teachers who have inspired and encouraged me throughout my education.

Finally, I would like to thank my parents, Jaimini and Suryakant Patel, and my brother, Kishan, for all their love and support. This thesis is dedicated to them.

Contents

List of Figures

Chapter 1

Introduction

There are many problems in combinatorics related to the partitioning of structures. Those studied in this thesis include graph cut problems; the problem of obtaining regular (in the sense of the Szemerédi Regularity Lemma) partitions for graphs and other combinatorial structures; and hypergraph problems with colouring constraints. Within these areas, one can ask different types of questions, for example, extremal questions, structural questions, or algorithmic questions.

Extremal questions ask us to find best possible bounds on real-valued functions that take as their inputs combinatorial structures having certain properties. For example, we might ask, what is the maximum number of edges in a triangle-free graph on n vertices? Here, our function takes a triangle-free graph as an input and outputs the number of edges in the graph. Any graph for which the maximum is achieved is called an extremal graph for the problem. As we shall see in Section 1.4, there is a unique extremal graph for this problem.

Structural questions ask for some sort of description or classification of all combinatorial structures having a certain property. For example, what can we say about triangle-free graphs having *close* to the maximum number of edges? It turns out that for such graphs, we can alter (add and delete) a *small* number of edges to obtain the extremal graph. Thus, triangle-free graphs having close to the maximum number of edges are structurally close to the extremal graph. Again, this is made precise in Section 1.4.

Algorithmic questions ask if we can devise a general method (an algorithm) for quickly checking if a given input combinatorial structure contains a certain substructure. For example, we might ask for an efficient method to check whether a (general) graph contains any triangles or not. A naive way to do this would be to check all triples of vertices in the input graph for triangles. This is in fact an efficient algorithm in a certain sense, as we shall discuss. On the other hand, we might ask for an efficient method to determine whether a graph is 3-colourable (that is, whether the vertices of the graph can be partitioned into three classes such that no class contains an edge). There is no known efficient algorithm for solving this problem, and furthermore, it is believed that no such efficient algorithm exists.

Often, these types of questions are interrelated; for example, the guaranteed existence of a certain substructure may lead to an extremal result or the construction of an algorithm. We shall investigate questions of these three types in the context of partitions of combinatorial structures.

We outline how the rest of this chapter is arranged. In Section 1.1, we discuss extremal and algorithmic aspects of the theory of graph cuts, as background to and in preparation for Chapters 2 and 3. We give the basic definitions and notation, and we state and prove some simple results to give a flavour of what is to come.

In Section 1.2, we introduce partially ordered sets or posets, giving notation and elementary results in preparation for Chapters 3, 4, and 5.

In Section 1.3, we turn to the Szemerédi Regularity Lemma. We develop notation and terminology, and we state the Regularity Lemma, all in preparation for Chapter 5. This section is a gentle introduction to the notion of regular partitions and the Regularity Lemma.

In Section 1.4, we introduce extremal hypergraph theory. We outline some of the main themes and results in this broad area of combinatorics, placing in context the results of Chapter 6.

Finally in Section 1.5, we give summaries of the results presented in each chapter of the thesis.

Chapters 2 and 3 are based on published work [46] and [47] respectively. Chapter 4, which is joint work (in equal part) with Graham Brightwell, is being submitted for publication.

1.1 Graph Cuts

All graphs are understood to be finite, simple, and undirected unless otherwise stated. One of the most basic things we can do with a graph $G = (V, E)$ is to partition its vertex set and ask questions about how the edges of the graph relate to the partition. For example, given a bipartition of V, that is, a partition of V into two parts V_1 and V_2, we might consider the set of edges that have one end in V_1 and the other in V_2. This set of edges is denoted by $E_G(V_1, V_2)$, and any set of edges generated in this way is called a *cut* of G. Let $e_G(V_1, V_2) = |E_G(V_1, V_2)|$.

Two algorithmic problems associated with cuts, known as the mincut and maxcut problems, ask us to find a cut in the input graph of smallest and largest size respectively. (For the mincut problem, we do not permit the trivial partition ϕ, V, which yields the empty cut.) Both of these problems and their variants have found a great many applications to real-world problems, and whilst we are not concerned with these applications here, we mention a few for completeness.

We note that, for a given connected graph, deleting the edges of any cut disconnects the graph. Thus the size of the smallest cut, known as the *edge connectivity* of a graph, gives the minimum number of edges that need to be deleted in order to disconnect the graph. Thinking of a graph as a model for a network, for example, a transport network or a communications network, the connectivity is a possible measure of how robust a network is. Finding the smallest non-empty cut allows us to determine the weakest points in a communications network or the sites of most congestion in a transport network, and therefore gives an indication of where best to place new links.

Applications of the maxcut problem are less direct: amongst others, they include cluster analysis, the design of VLSI (very large scale integration) circuits, and statistical physics (see [4] for more on the latter two). Cluster analysis, for example, is an exploratory data analysis tool for classification problems. The problem is to find algorithms that partition a set of data points into groups, where points within the same group should be more closely related than points in different groups. The aim is to reveal new structures or classifications within the data. Thinking of the data points as vertices of a graph, with an edge indicating that two data points are unrelated or weakly related, finding a maximum sized cut in the graph gives a partition for which data points in different parts are as unrelated as possible.

Of course, problems such as maxcut and mincut are often too simple to adequately model real-world problems, which in general have many constraints. However, solutions or approximate solutions to maxcut and mincut can still provide heuristics for the real-world problems.

One way to determine the size of the largest and smallest cuts of an n-vertex graph would be to enumerate all 2^n bipartitions of its vertices and check the sizes of the largest and smallest cuts. The running time for such an algorithm grows very quickly with the size of the graph. Thus, large improvements in computational speed permit only relatively small improvements in the size of the graphs for which the algorithm terminates in a reasonable time. Clearly, we would like to have a more efficient way of solving the maxcut and mincut problems. In this sense, there is a significant difference between the two problems: the mincut problem can be solved efficiently, whereas the maxcut problem is intractable. The mincut problem is *polynomial-time solvable*, whereas the maxcut problem is *NP-hard.* We do not give the precise technical definitions for these terms; instead we give working definitions, which will suffice for the purposes of this thesis. For a more comprehensive treatment, see, for instance, [26].

A problem is polynomial-time solvable if there exists an algorithm that solves the problem and whose running time is bounded by a polynomial in the length of the input, where the running time is the number of elementary operations performed during the course of the algorithm. See the first proof of Proposition 1.1.1 and the comments that follow for an analysis of the running time of an algorithm.

For the maxcut and mincut problems, the input is simply a graph, and we may assume that any reasonable encoding of a graph on n vertices has length that is polynomially bounded in n. In this case, a polynomial-time algorithm is simply one whose running time is bounded by a polynomial in n. The mincut problem can be solved by such a polynomial-time algorithm (for example, the Ford-Fulkerson algorithm based on their max-flow min-cut theorem [20]). The maxcut problem in contrast is NP-hard [26]. Since NP-hardness is mentioned only to place results of the thesis in context, we do not define it here, but refer the reader to [26], where it is treated comprehensively. We indicate only that it is widely believed that a problem which is NP-hard is not polynomial-time solvable. This is a direct consequence of the famous conjecture that P $\neq$ NP.

We now turn to the extremal theory of graph cuts. Given a graph $G = (V, E)$, let $c(G)$ be the size of the largest cut in G. We would like to find a lower bound on $c(G)$ in terms of $m := |E|$. The best possible, or *extremal*, bound on $c(G)$ in terms of m is given by determining the value of

$$c(m) := \min_{G:G \text{ has } m \text{ edges}} c(G).$$

(Note that there are only finitely many graphs with m edges up to addition of isolated vertices.) A graph on m edges for which $c(G) = c(m)$ is known as an *extremal* graph (for this problem). It is a graph that is worst in terms of the proportion of edges in its largest cut. A very elementary and well-known bound for $c(m)$ is given by the following proposition; we give two proofs of the proposition to illustrate the different techniques available to us.

Proposition 1.1.1 *For every graph $G = (V, E)$, there exists a bipartition of V into parts V_1 and V_2 such that*

$$e_G(V_1, V_2) \geq \frac{1}{2}|E|,$$

or equivalently, for all $m \in \mathbb{N}$, we have $c(m) \geq \frac{1}{2}m$.

Before we give the proofs, let us fix some notation. Given a graph $G = (V, E)$, with $A, B \subseteq V$, let

$$E_G(A) = \{ab \in E : a, b \in A\} \quad \text{and} \quad e_G(A) = |E_G(A)|.$$

As before, but now allowing A, B to be arbitrary subsets of V (not just a partition), let

$$E_G(A, B) = \{ab \in E : a \in A, b \in B\} \quad \text{and} \quad e_G(A, B) = |E_G(A, B)|.$$

If A consists of a single vertex x, then we write $E_G(x, B)$ and $e_G(x, B)$ rather than $E_G(\{x\}, B)$ and $e_G(\{x\}, B)$, and we refer to $e_G(x, B)$ as the degree of x in B. The *degree* of x (in G) is simply $e_G(x, V)$. The subscript is dropped when it is clear which graph is being referred to.

Proof (constructive) The graph $G = (V, E)$ is given. Pick any bipartition U_1, U_2 of V. If such a vertex exists, let x be a vertex of U_1 for which

$e(x, U_2) < e(x, U_1)$, or a vertex of U_2 for which $e(x, U_1) < e(x, U_2)$. Create a new bipartition, U_1', U_2' by moving x from U_1 to U_2 if $x \in U_1$ and moving x from U_2 to U_1 if $x \in U_2$. Our choice of x ensures that

$$e(U_1', U_2') > e(U_1, U_2).$$

We iterate this process of moving vertices to increase the cut size, and since the cut size cannot increase indefinitely, the process is guaranteed to terminate. Let V_1, V_2 be the final partition of V: for this partition, we have that

$$e(x, V_2) \geq e(x, V_1) \ \forall x \in V_1 \quad \text{and} \quad e(y, V_1) \geq e(y, V_2) \ \forall y \in V_2,$$

that is,

$$e(x, V_2) \geq \frac{1}{2} e(x, V) \ \forall x \in V_1 \quad \text{and} \quad e(y, V_1) \geq \frac{1}{2} e(y, V) \ \forall y \in V_2.$$

Summing over all vertices, we have

$$\begin{aligned} 2|E| = \sum_{v \in V} e(v, V) &= \sum_{x \in V_1} e(x, V) + \sum_{y \in V_2} e(y, V) \\ &\leq \sum_{x \in V_1} 2e(x, V_2) + \sum_{y \in V_2} 2e(y, V_1) = 4e(V_1, V_2). \end{aligned}$$

Thus $e(V_1, V_2) \geq \frac{1}{2}|E|$, proving the proposition. □

Note that the proof gives a polynomial-time algorithm for constructing a cut containing at least half the edges of the graph. Let us see why this is true. Assuming $V = [n]$, where $[n] = \{1, \ldots, n\}$, one way to encode our graph is as an $n \times n$ matrix whose $(i, j)^{\text{th}}$ entry is 1 if ij is an edge of G and is 0 otherwise. Thus the input length is n^2. For a given vertex x, determining $e(x, U_1)$ requires us to check the presence of all edges in $E(x, U_1)$: this requires $O(n)$ elementary operations. Thus checking which of $e(x, U_1)$ or $e(x, U_2)$ is greater and moving x if necessary requires $O(n)$ operations. In the worst case, we check at most n vertices before one is moved or the process terminates. Each time a vertex is moved, the size of the cut increases by at least 1; hence in the worst case, we move vertices a total of at most $|E| < n^2$ times. Thus the total running time is $O(n) \cdot n \cdot n^2 = O(n^4)$, which is polynomially bounded in n^2 as required.

Our second proof only guarantees the existence of the desired cut, but does not give an explicit construction. We shall make use of both techniques when we consider the problems of Chapters 2 and 3.

Throughout, $\mathbb{P}(E)$ denotes the probability of an event E, and $\mathbb{E}(X)$ denotes the expectation of a random variable X.

Proof (probabilistic) Given the graph $G = (V, E)$, choose a subset V_1 of V uniformly at random, and let V_2 be its complement. That is, V_1 is constructed by tossing a fair coin for each vertex independently to determine whether it belongs to V_1 or V_2. We have that

$$\mathbb{E}(e_G(V_1, V_2)) = \sum_{e \in E} \mathbb{P}(e \in E_G(V_1, V_2)) = \sum_{e \in E} \frac{1}{2} = \frac{1}{2}|E|.$$

Since the average size of a random cut is $\frac{1}{2}|E|$, there must exist a cut of size at least $\frac{1}{2}|E|$. □

We have shown that $c(m) \geq \frac{1}{2}m$ for all $m \in \mathbb{N}$. The constant $\frac{1}{2}$ cannot be improved. To see this, let K_n be the complete graph on n vertices, that is, the graph on n vertices with all edges present. It has $\binom{n}{2}$ edges, and the largest cut is given by bipartitioning the vertices of K_n as evenly as possible. Thus, if n is odd and $m = \binom{n}{2}$, then

$$c(m) \leq c(K_n) = \left(\frac{n+1}{2}\right)\left(\frac{n-1}{2}\right) = \frac{1}{2}\binom{n}{2} + \frac{n-1}{4}.$$

Solving the quadratic $m = \binom{n}{2}$ for n in terms of m, we obtain

$$c(m) \leq c(K_n) = \frac{1}{2}m + \sqrt{\frac{m}{8} + \frac{1}{64}} - \frac{1}{8}.$$

Although the constant $\frac{1}{2}$ cannot be improved, we can improve the lower bound on $c(m)$ with lower order terms. Edwards [13, 14] proved that the upper bound for $c(m)$ given above is in fact also a lower bound for *all* m. We give a short (adapted) proof of Alon [2] and Hofmeister and Lefmann [32] of this result.

Theorem 1.1.2 (Edwards [13, 14]) *For every $m \in \mathbb{N}$, we have*

$$c(m) \geq \frac{1}{2}m + \sqrt{\frac{m}{8} + \frac{1}{64}} - \frac{1}{8}.$$

Note that this is an extremal result because the bound is sharp for infinitely many values of m. We prove this result by considering weighted graphs, for which we now set up notation. Given a set of vertices V, let $V^{(2)}$ denote the set of unordered pairs of V. A *weighted graph* is a pair $G = (V, w)$, where w is a function from $V^{(2)}$ to the non-negative reals. We call w a *weighting* of G, and for $ab \in E$, we call $w(ab)$ the *weight* of ab. The *total weight* of the graph is defined to be

$$w(G) = \sum_{ab \in V^{(2)}} w(ab).$$

For V_1, V_2 a partition of V, we define the *weight across* V_1, V_2 to be

$$w(V_1, V_2) = \sum_{a \in V_1,\, b \in V_2} w(ab).$$

Note that a normal graph $G = (V, E)$ can be thought of as a weighted graph $G = (V, w)$ having a *standard weighting* of 1 on the edges and 0 on the non-edges, and in this case,

$$w(G) = |E| \quad \text{and} \quad w(V_1, V_2) = e_G(V_1, V_2).$$

This gives a natural generalisation of graph cut problems to weighted graphs.

Proof (of Theorem 1.1.2) Let $G = (V, E)$ be a graph with m edges. Let w be the standard weighting for this graph. We compress G iteratively in the following way. Whenever we find two vertices $x, y \in V$ with $w(xy) = 0$, we replace them with a vertex z, and for every $a \in V \setminus \{x, y\}$, we set $w(za) = w(xa) + w(ya)$. We repeat this process until we obtain the weighted graph $G' = (V', w')$ in which every element of $V'^{(2)}$ has a positive weighting. In fact, we have $w'(e) \geq 1$ for every $e \in V'^{(2)}$. Also, we have $w'(G') = m$, and setting $n' = |V'|$, we obtain that $m \geq \binom{n'}{2}$.

Any partition V_1', V_2' of V' can be extended to a partition V_1, V_2 of V by reversing the iterative process described above, replacing z with x and y at each stage, and keeping x and y in the same part as z. At each stage in this reverse process, the weight across the partition does not change; hence we obtain

$$w(V_1, V_2) = w'(V_1', V_2'),$$

and so it is sufficient to find a partition V_1', V_2' of V' such that

$$w'(V_1', V_2') \geq \frac{1}{2}m + \sqrt{\frac{m}{8} + \frac{1}{64}} - \frac{1}{8}.$$

Let U_1, U_2 be a random equipartition of V' (an equipartition is a partition where the parts differ in size by at most one element), and let

$$K(U_1, U_2) = \{ab \in V'^{(2)} : a \in U_1, b \in U_2\}.$$

Since $|K(U_1, U_2)| = \lfloor n'^2/4 \rfloor$, and each of the $\binom{n'}{2}$ elements of $V'^{(2)}$ is equally likely to occur as an element of $K(U_1, U_2)$, we have that

$$\mathbb{P}(ab \in K(U_1, U_2)) = \frac{\lfloor n'^2/4 \rfloor}{\binom{n'}{2}}.$$

We have that

$$\begin{aligned}
\mathbb{E}(w'(U_1, U_2)) &= \sum_{ab \in V'^{(2)}} \mathbb{P}(ab \in K(U_1, U_2)) w'(ab) \\
&= \sum_{ab \in V'^{(2)}} \frac{\lfloor n'^2/4 \rfloor}{\binom{n'}{2}} w'(ab) \\
&\geq \frac{n'^2 - 1}{2n'(n' - 1)} m \\
&= \frac{n' + 1}{2n'} m \\
&= \frac{1}{2}m + \frac{m}{2n'} \\
&\geq \frac{1}{2}m + \frac{n' - 1}{4} \quad \text{using } m \geq \binom{n'}{2} \\
&\geq \frac{1}{2}m + \sqrt{\frac{m}{8} + \frac{1}{64}} - \frac{1}{8},
\end{aligned}$$

where the last inequality follows by solving the inequality $m \geq \binom{n'}{2}$ for n'. We are now guaranteed the existence of the desired partition V_1', V_2'. □

The results we have discussed can easily be generalised to *k-partitions* (partitions into k parts), which we now discuss briefly. Given a graph $G = (V, E)$ and k subsets $V_1, \ldots, V_k$ of V, we define

$$E_G(V_1, \ldots, V_k) = \{ab \in E : a \in V_i, b \in V_j, i \neq j\}.$$

When $V_1, \ldots, V_k$ is a k-partition, $E_G(V_1, \ldots, V_k)$ is called a *k-cut* of G. We let

$$e_G(V_1, \ldots, V_k) = |E_G(V_1, \ldots, V_k)|.$$

We have the following proposition corresponding to Proposition 1.1.1.

Proposition 1.1.3 *For every graph $G = (V, E)$, there exists a k-partition of V into parts $V_1, \ldots, V_k$ such that*

$$e_G(V_1, \ldots, V_k) \geq \left(1 - \frac{1}{k}\right)|E|.$$

Proof (probabilistic) Given the graph $G = (V, E)$, construct a random k-partition $V_1, \ldots, V_k$ by assigning each vertex of V independently and uniformly at random to one of $V_1, \ldots, V_k$. We have that

$$\begin{aligned} \mathbb{E}(e_G(V_1, \ldots, V_k)) &= \sum_{e \in E} \mathbb{P}(e \in E_G(V_1, \ldots, V_k)) = \sum_{e \in E} \left(1 - \frac{1}{k}\right) \\ &= \left(1 - \frac{1}{k}\right)|E|. \end{aligned}$$

Since the average size of a random k-cut is $(1 - \frac{1}{k})|E|$, there must exist a k-cut of size at least $(1 - \frac{1}{k})|E|$. □

Again, this result is extremal in the sense that the constant $(1 - \frac{1}{k})$ cannot be improved. However, as before, the result can be improved by adding lower order terms. The following result generalises Theorem 1.1.2.

Theorem 1.1.4 (Bollobás, Scott [5]) *For every graph $G = (V, E)$ with $m = |E|$, there exists a k-partition of V into parts $V_1, \ldots, V_k$ such that*

$$e_G(V_1, \ldots, V_k) \geq \left(1 - \frac{1}{k}\right)m + \frac{1}{2}\left(1 - \frac{1}{k}\right)\sqrt{2m + 1/4} - \frac{k^2 - 2k + 2}{8k}.$$

(Note that the statement of this theorem appearing in [5] and [54] has a small error.)

We have discussed what may be regarded as the first problems in extremal graph cut theory. There are many other problems derived from these ones. For example, we can formulate problems in which we consider more than one graph, special types of graphs, or special types of partitions. Indeed, such problems are what we consider in Chapters 2 and 3. For a survey on extremal graph cut problems, see [54].

1.2 Posets and Ordered Partitions

Chapters 3, 4, and 5 are concerned with partially ordered sets or *posets*, and more specifically with what we call *ordered partitions* of posets. We take this opportunity to give some basic definitions, to fix notation, and to give some fundamental partitioning results for posets. We begin with some basic notation for posets.

A poset P is a pair $(X, \prec)$, where X is a set called the *ground set* of P, and where $\prec$, called the *order* of P, is a binary relation on X satisfying the following properties:

1. $\prec$ is irreflexive: for all $a \in X$, we do not have $a \prec a$;

2. $\prec$ is transitive: for all $a, b, c \in X$, if $a \prec b$ and $b \prec c$, then $a \prec c$.

We sometimes write $a \succ b$ to mean $b \prec a$, and we write $a \preceq b$ to mean that either $a \prec b$ or $a = b$. All posets considered will be finite, that is, X will be finite. A poset $Q = (X, \prec_Q)$ is called the *dual* of a poset $P = (X, \prec_P)$ if for every $a, b \in X$, we have that $a \prec_Q b$ if and only if $b \prec_P a$.

Given a poset $P = (X, \prec)$ and $a, b \in X$, we say that a and b are *comparable* in P (written $a \perp b$) if either $a \prec b$ or $b \prec a$, and we say that a and b are *incomparable* in P (written $a \parallel b$) if a and b are not comparable in P. We say that b *covers* a in P if $a \prec b$ and there is no $c \in X$ such that $a \prec c \prec b$. Three graphs naturally associated with a poset $P = (X, \prec)$ are its *comparability graph*, denoted by $\mathrm{Com}(P)$; its *incomparability graph*, denoted by $\mathrm{Inc}(P)$; and its covering graph, denoted by $\mathrm{Cov}(P)$. Each of these graphs is an undirected graph on X. The edges of $\mathrm{Com}(P)$ are the pairs ab for which $a \perp b$; the edges of $\mathrm{Inc}(P)$ are the pairs ab for which $a \parallel b$; and the edges of $\mathrm{Cov}(P)$ are the pairs ab for which b covers a or a covers b.

A *chain* is a poset in which every pair of elements of the ground set is comparable, and as a consequence, the elements are ordered linearly. An *antichain* is a poset in which every pair of elements of the ground set is incomparable.

Fix a poset $P = (X, \prec)$. For $A \subseteq X$, the poset $Q = (A, \prec)$ is the poset with the same order as P, but restricted to A. Sometimes we refer to such posets simply by their ground sets; for instance, $C \subseteq X$ is referred to as chain of P if $(C, \prec)$ is a chain. Similarly $A \subseteq X$ is referred to as an antichain of P if $(A, \prec)$ is an antichain. The number of elements in the largest chain of P is called the *height* of P and is denoted by $ht(P)$. The number of elements in the largest antichain of P is called the *width* of P and is denoted by $wd(P)$.

We now turn to the very simple and natural notion of an *ordered partition.* Given a poset $P = (X, \prec)$, a partition $X_1, \ldots, X_k$ of X is called an ordered partition (of P), or an ordered k-partition (of P) if, whenever $a \in X_i$ and $b \in X_j$ with $a \prec b$, then $i \leq j$. Taking the contrapositive, $X_1, \ldots, X_k$ is an ordered partition of P if, whenever $a \in X_i$ and $b \in X_j$ with $i < j$, then either $a \prec b$ or $a \parallel b$. We shall interchange between these definitions without mention. An ordered partition of P is thus a partition of the ground set respecting the order of P. Such a simple idea undoubtedly pervades the literature on posets, although it is usually presented inconspicuously in alternative equivalent forms. For us, however, ordered partitions are highlighted and form a running theme through a significant part of the thesis.

Let us describe some of the alternative forms in which ordered partitions may appear. For a poset $P = (X, \prec)$, $U \subseteq X$ is called an *up-set* of P if, whenever $u \in U$ with $u' \succ u$, then $u' \in U$. Similarly $D \subseteq X$ is called a *down-set* of P if, whenever $d \in D$ with $d' \prec d$, then $d' \in D$. Note that if U is an up-set of P, then $U^c = X \backslash U$ is a down-set of P, and furthermore U^c, U forms an ordered 2-partition of P. Conversely, if X_1, X_2 is an ordered 2-partition of P, then X_1 is a down-set and X_2 is an up-set of P. More generally, we have the following correspondence.

Proposition 1.2.1 *Let $P = (X, \prec)$ be a poset.*

(i) If $\phi = D_0 \subseteq D_1 \subseteq D_2 \subseteq \cdots \subseteq D_k = X$ is a nested sequence of down-sets of P, then the partition of $X_1, \ldots, X_k$ given by

$$X_i = D_i \backslash D_{i-1}$$

is an ordered k-partition of P.

(ii) Conversely, if $X_1, \ldots, X_k$ is an ordered k-partition of P, then the sets $D_1 \subseteq \cdots \subseteq D_k$ given by

$$D_i = \bigcup_{j \leq i} X_j$$

form a nested sequence of down-sets of P.

Proof For (i), it is clear that $X_1, \dots, X_k$ is a partition of X. Take $a \in X_i$, $b \in X_j$, and $a \prec b$. We cannot have $j < i$, otherwise we have $b \in D_j$ and $a \notin D_j$ (since $X_i \cap D_j = \phi$), contradicting that D_j is a down-set. Thus $i \leq j$ as required.

For (ii), it is clear that the family $D_1, \dots, D_{k-1}$ is nested. To show that D_i is a down-set, we take $b \in D_i$ and $a \in X$ with $a \prec b$, and show that $a \in D_i$. Let r and s be such that $a \in X_r$ and $b \in X_s$. Thus $s \leq i$ since $b \in X_s \cap D_i$. Also $r \leq s$ since $a \prec b$ and $X_1, \dots, X_k$ is ordered. Thus $r \leq i$ and so $a \in X_r \subseteq D_i$ as required. □

Order-preserving functions give an alternative way of viewing ordered partitions. Let $P = (X, \prec)$ be a poset. For a natural number k, a function $f : X \longrightarrow [k]$ is called order preserving if, whenever $a \prec b$, we have $f(a) \leq f(b)$. For $i \in [k]$, writing $f^{-1}(i)$ to denote the set of elements of X mapped by f to i, we have the following proposition.

Proposition 1.2.2 *Let $P = (X, \prec)$ be a poset and k a positive integer.*

(i) If $f : X \longrightarrow [k]$ is an order-preserving function, then $f^{-1}(1), \dots, f^{-1}(k)$ is an ordered partition of P.

(ii) Conversely, if $X_1, \dots, X_k$ is an ordered partition of P, then the function $f : X \longrightarrow [k]$ defined by $f(a) = i$ if and only if $a \in X_i$ for all $a \in X$ and all $i \in [k]$ is order preserving.

Proof For (i), suppose $a \in f^{-1}(i)$ and $b \in f^{-1}(j)$ with $a \prec b$. Thus $i = f(a) \leq f(b) = j$ as required.

For (ii), suppose $f(a) = i$, $f(b) = j$, and $a \prec b$. Then we have that $a \in X_i$, $b \in X_j$, and since $a \prec b$ with $X_1, \dots, X_k$ an ordered partition, we have $i \leq j$ as required. □

For a poset $P = (X, \prec)$ with $|X| = n$, an order-preserving bijection $\lambda : X \longrightarrow [n]$ is called a *linear extension* of P. It is thus a total ordering of the elements of X, $\lambda^{-1}(1), \lambda^{-1}(2), \dots, \lambda^{-1}(n)$, that respects $\prec$. Therefore, an alternative equivalent way to think of a linear extension of P is as a chain $L = (X, \prec^*)$, where $a \prec b$ implies $a \prec^* b$ for all $a, b \in X$. Linear extensions are studied in Chapter 4.

We conclude this subsection by giving some fundamental results for posets, which we shall refer back to. The theorem below is a well-known and easy result; see, for instance, Trotter [62].

Theorem 1.2.3 *If the height of a poset $P = (X, \prec)$ is h, then there exists a partition of X into h antichains. Moreover, the partition is* ordered.

We give some terminology and notation before proving the theorem. Given a poset $P = (X, \prec)$ and $A \subseteq X$, $x \in A$ is called a *maximal element of A (in P)* if there is no $x' \in A$ such that $x' \succ x$. Similarly $x \in A$ is called a *minimal element of A (in P)* if there is no $x' \in A$ such that $x' \prec x$. The set of maximal (resp. minimal) elements of A is denoted by $\max_P(A)$ (resp. $\min_P(A)$), where we sometimes drop the subscript.

Proof (of Theorem 1.2.3) The proof is by induction on h. The theorem holds when $h = 1$, that is, when P is an antichain (the ordered partition being the trivial partition with one part).

Given $P = (X, \prec)$ with $ht(P) = h \geq 2$, define $X' = X \setminus \max_P(X)$ and let $P' = (X', \prec)$. The height of P' is $h - 1$ (since the maximal elements of the longest chains in P have been removed). By the induction hypothesis, let $X_1, \ldots, X_{h-1}$ be an ordered partition of P' into antichains. Note that $\max_P(X)$ is an antichain of P (since two maximal elements cannot be comparable). Then setting $X_h = \max_P(X)$, we have that $X_1, \ldots, X_h$ is a partition of X into antichains of P, and we claim that it is an ordered partition. Indeed, let $a \in X_i$, $b \in X_j$ with $i < j$. If $j \leq h - 1$ then we know that either $a \prec b$ or $a \parallel b$ since $X_1, \ldots X_{h-1}$ is ordered. If $j = h$, then again we know that either $a \prec b$ or $a \parallel b$ since X_h consists of maximal elements of P. □

We now state the dual theorem to the one above. This theorem, proved originally by Dilworth [12] in 1950, turns out to be a more substantial result.

Theorem 1.2.4 (Dilworth's Theorem) *If a poset $P = (X, \prec)$ has width w, then there exists a partition of X into w chains.*

A generalisation of this result was given in a proof of Greene and Kleitman [28], where, rather than studying the size of the largest antichain of a poset, one studies the size of the largest union of k antichains for k fixed. We shall require an algorithmic version of this result in Chapter 3. We therefore follow Frank's algorithmic treatment [21] of the Greene-Kleitman Theorem.

Let $P = (X, \prec)$ be a poset; let $C_1, \ldots, C_r$ be r chains of P; and let $A_1, \ldots, A_s$ be s antichains of P. We refer to $\mathscr{C} = \{C_1, \ldots, C_r\}$ as a *chain family* if the chains are pairwise disjoint, and likewise, we refer to $\mathscr{A} =$

$\{A_1, \ldots, A_r\}$ as an *antichain family* if the antichains are pairwise disjoint. The *order* of $\mathscr{C}$ (resp. $\mathscr{A}$) is the number of chains (resp. antichains) in $\mathscr{C}$ (resp. $\mathscr{A}$). The *size* of $\mathscr{C}$ (resp. $\mathscr{A}$) is $|\cup_{i=1}^{r} C_i|$ (resp. $|\cup_{i=1}^{s} A_i|$).

Given an antichain family $\mathscr{A} = \{A_1, \ldots, A_s\}$ and a chain family $\mathscr{C} = \{C_1, \ldots, C_r\}$, observe that $|C_i \cap A_j| \leq 1$ for $1 \leq i \leq r$, $1 \leq j \leq s$. Thus, any antichain family of order s can have size at most

$$rs + |X \backslash (\cup_{i=1}^{r} C_i)|.$$

This bound is achieved if $\mathscr{A}$ and $\mathscr{C}$ are *orthogonal*, that is, if we have

(a) $\cup_{i=1}^{r} C_i \bigcup \cup_{i=1}^{s} A_i = X$ and

(b) $C_i \cap A_j \neq \phi$ for $1 \leq i \leq r$, $1 \leq j \leq s$.

The following result is due to Frank [21].

Theorem 1.2.5 *Let $P = (X, \prec)$ be a poset. Given a positive integer $s \leq ht(P)$, there exists an antichain family $\mathscr{A}$ of order s and a chain family $\mathscr{C}$ that are orthogonal. Moreover, $\mathscr{A}$ and $\mathscr{C}$ can be found in time polynomial in $|X|$.*

It immediately follows from this theorem that $\mathscr{A}$ has maximum possible size amongst all antichain families of order s. We state this as a corollary for future reference.

Corollary 1.2.6 *Given a poset $P = (X, \prec)$ and a positive integer s, a maximum union of s antichains can be found in time polynomial in $|X|$.*

1.3 Szemerédi Regularity Lemma

In 1975, Szemerédi [57] proved a deep Ramsey-type theorem for the integers, which had been conjectured by Erdős and Turán [18]: he proved that every set of integers with a positive upper density has arbitrarily long arithmetic progressions. A key step in his proof turned out to be an innocuous looking lemma whose importance has come to be realised more and more over recent decades. The lemma has come to be known as the *Szemerédi Regularity Lemma* [58] or just the *Regularity Lemma* and has found many applications in discrete mathematics, especially graph theory. For a survey on the applications of the Regularity Lemma to graph theory, see [38] and references therein.

Informally, the Regularity Lemma says that all large, dense graphs can in some sense be approximated by random graphs; results for random graphs can then in a way be carried over to all graphs. Going further, the Regularity Lemma says that every graph can be partitioned into a bounded number of parts such that for almost all pairs of parts, the edges between the parts are distributed uniformly, much as they would be if they were generated at random.

For the remainder of this section, we describe the Regularity Lemma for graphs, and in so doing, give the standard notational setup. Readers familiar with the Regularity Lemma will find this routine.

There are various slightly different versions of the Regularity Lemma, even for graphs. Here, we give one that is less common (Theorem 1.8 [38]), but better suited for our adaptation to posets in Chapter 5. We begin with some notation.

Let $G = (V, E)$ be a graph. For $A, B \subseteq V$, we define the *density* between A and B in G to be $d_G(A, B)$, where

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

Again, subscripts may be dropped. For $\epsilon, \delta \in (0,1]$, the pair (A, B) is called $(\epsilon, \delta)_G$*-regular* if, whenever $A' \subset A$ and $B' \subset B$ with $|A'| > \delta|A|$ and $|B'| > \delta|B|$, we have

$$|d_G(A', B') - d_G(A, B)| \leq \epsilon.$$

Thus the condition of $(\epsilon, \delta)_G$-regularity says that the edges in G between A and B are distributed approximately uniformly. The smaller ϵ and δ are, the more uniform is the distribution of edges between A and B. Note also that if $d(A, B)$ is small, say $d(A, B) < \delta^2\epsilon$, then (A, B) is automatically (ϵ, δ)-regular. Indeed, if $A' \subset A$ and $B' \subset B$ with $|A'| > \delta|A|$ and $|B'| > \delta|B|$, then

$$0 \leq d(A', B') = \frac{e(A', B')}{|A'||B'|} < \frac{\epsilon\delta^2|A||B|}{\delta|A|\delta|B|} = \epsilon.$$

Thus

$$|d(A', B') - d(A, B)| \leq \epsilon.$$

Similarly (by considering the complementary graph), (A, B) is (ϵ, δ)-regular if $d(A, B) > 1 - \delta^2\epsilon$.

Recall that an equipartition of V into k parts is a partition $V_1, \ldots, V_k$ of V, where the parts are as equal in size as possible, that is, $\big||V_i| - |V_j|\big| \leq 1$ for all $1 \leq i, j \leq k$.

Let $\epsilon, \delta, \gamma \in (0, 1]$ and let $G = (V, E)$ be a graph. An equipartition $V_1, \ldots, V_k$ of V is called $(\epsilon, \delta, \gamma)_G$-regular if the number of (unordered) pairs, (V_i, V_j), that are not $(\epsilon, \delta)_G$-regular is at most $\gamma\binom{k}{2}$. An equipartition that is $(\epsilon, \epsilon, \epsilon)_G$-regular will be referred to as $(\epsilon)_G$-regular or just ϵ-regular. Note that a $(\epsilon_0, \delta_0, \gamma_0)_G$-regular equipartition is $(\epsilon)_G$-regular for $\epsilon > \max(\epsilon_0, \delta_0, \gamma_0)$.

The Regularity Lemma says that, for every graph, there exists an ϵ-regular equipartition, where the number of parts is bounded by a function of ϵ only. Having a bound on the number of parts that is independent of the graph is a key feature of the Regularity Lemma. Indeed, if we do not insist on such a condition, we can simply partition our graph into vertices, which trivially gives an ϵ-regular partition for every $\epsilon \in (0, 1]$. Here is the formal statement of the Regularity Lemma.

Theorem 1.3.1 (Szemerédi Regularity Lemma) *For every $\epsilon \in (0, 1]$ and every $m \in \mathbb{N}$, there exists a natural number $M = M(\epsilon, m)$ with the following property. For every graph $G = (V, E)$, there exists an $(\epsilon)_G$-regular equipartition of V into k parts, where $m \leq k \leq M$.*

The purpose of m in the above statement is a technical one. When applying the Regularity Lemma, one may wish to have a sufficiently large number of parts in order to ensure that the number of edges within parts is small.

Note that if G has a sufficiently small number of edges (as a function of ϵ), then in any partition of its vertices, the density between most parts will be small; thus, all such pairs of parts will automatically be $(\epsilon, \epsilon)_G$-regular (as previously explained). In this case, all partitions of G are ϵ-regular, and the Regularity Lemma gives us no useful information. For such graphs, there exist sparse versions of the Regularity Lemma that are sometimes useful (see for example [37]), but we shall not be concerned with these.

The function $M(\epsilon, m)$ is an Ackermann-type function: for $m = 2$, it is a tower of powers $2^{2^{2^{\cdots}}}$ of height $\Omega(\epsilon^{-5})$. Thus our regular partition has a very large, but bounded, number of parts. Gowers [27] showed that any upper bound on $M(\epsilon, m)$ necessarily has such tower-type growth as ϵ tends to zero.

1.4 Extremal Hypergraph Theory

A simple way to generalise the notion of a graph is the following: rather than taking edges to be pairs of vertices, we take edges to be subsets of vertices of size k for some fixed k. This results in what is known as a k-uniform hypergraph. Many problems in graph theory have corresponding generalisations for k-uniform hypergraphs, but often turn out to be much harder to solve when $k \geq 3$.

Formally, a *k-uniform hypergraph* (or *k-graph*) H is a pair (V_H, E_H), where V_H is a set of vertices, E_H is a set of edges, and an edge is a k-element subset of V_H. We write $V_H^{(k)}$ for the set of all k-element subsets of V_H; thus $E_H \subseteq V_H^{(k)}$. In particular, a 2-graph is simply a graph.

Given two k-graphs $F = (V_F, E_F)$ and $H = (V_H, E_H)$, we say that H *contains* F or F is a *subgraph* of H if there exists an injective function $g : V_F \to V_H$ such that

$$\{g(e) : e \in E_F\} \subseteq E_H,$$

where we define $g(S) = \{g(v) : v \in S\}$ for all $S \subseteq V_F$. If, in addition, we have

$$\{g(e) : e \in E_F\} = E_H \cap \left(g(V_F)\right)^{(k)},$$

then we say that F is an *induced subgraph* of H and that the set of vertices $g(V_F) \subseteq V_H$ *induces* a copy of F in H.

We say H is *F-free* if H does not contain F. A very natural question asked by Turán first for graphs and then for hypergraphs is the following: given a k-graph F, what is the largest number of edges that an n-vertex F-free k-graph can have? This number is denoted by $ex(n, F)$ and is called the *Turán number* of F. Any n-vertex F-free k-graph having the maximum number $ex(n, F)$ of edges is called an *extremal* F-free hypergraph. The *density* of the extremal F-free hypergraphs is given by

$$d(n, F) = \frac{ex(n, F)}{\binom{n}{k}}.$$

We show that we can take the limit of $d(n, F)$ as $n \to \infty$. We do this by showing that $d(n, F)$ is a decreasing sequence using a standard averaging argument of Katona, Nemetz, and Simonovits [34].

Proposition 1.4.1 *Let F be a fixed k-graph. The limit*

$$\pi(F) := \lim_{n \to \infty} d(n, F)$$

always exists.

Proof We show that $d(n, F)$ is a decreasing sequence. Since it is bounded below by zero, we conclude that it has a limit. We use a simple averaging argument to show the sequence is decreasing. An extremal graph for F on $n+1$ vertices can contain at most $ex(n, F)$ edges in any subset of n vertices. Summing the edges over all such n-vertex subsets, and therefore counting each edge $n - k$ times, we have

$$(n - k)ex(n + 1, F) \leq (n + 1)ex(n, F).$$

Thus

$$d(n + 1, F) = \frac{ex(n + 1, F)}{\binom{n+1}{k}} \leq \frac{(n + 1)ex(n, F)}{(n - k)\binom{n+1}{k}} = d(n, F),$$

as required. □

For F a k-graph, $\pi(F)$ is called the *Turán density* of F. More generally, we can extend these definitions to families of hypergraphs. If $\mathcal{F}$ is a family of k-graphs, H is said to be $\mathcal{F}$-free if H does not contain any member of $\mathcal{F}$ as a subgraph. We let $ex(n, \mathcal{F})$ denote the maximum number of edges in an n-vertex $\mathcal{F}$-free k-graph. We define

$$d(n, \mathcal{F}) = \frac{ex(n, \mathcal{F})}{\binom{n}{k}},$$

and we let $\pi(\mathcal{F}) = \lim_{n \to \infty} d(n, \mathcal{F})$.

Determining these Turán densities is perhaps one of the most fundamental problems in extremal hypergraph theory. The problem has been solved for $k = 2$, that is, the Turán densities of all families of graphs have been determined. For $k \geq 3$, the problem remains open for most hypergraphs.

Let us discuss the extremal results for graphs. Recall that K_r denotes the complete graph on r vertices. Let $Tr(n,l)$ denote the graph on n vertices equipartitioned into l classes, $V_1, \ldots, V_l$ (recall that this means $||V_i| - |V_j|| \leq 1$ for all $i, j \in [l]$), where the edge set is given by

$$\{ab : a \in V_i,\ b \in V_j,\ i \neq j\}.$$

These graphs are called *Turán graphs*, and we write $tr(n,l)$ for the number of edges in $Tr(n,l)$. In 1941, Turán proved the following theorem [63], not only finding the Turán densities for complete graphs, but also giving the unique extremal K_r-free graphs. There have subsequently been many different proofs given for the theorem; see, for instance, [1].

Theorem 1.4.2 (Turán) *We have for $l \geq 1$, that*

$$ex(n, K_{l+1}) = tr(n,l),$$

and moreover, $Tr(n,l)$ is the unique extremal K_{l+1}-free graph.

Counting the number of edges in $Tr(n,l)$, we have

$$tr(n,l) = (1+o(1))\binom{l}{2}\left(\frac{n}{l}\right)^2 = (1+o(1))\left(1 - \frac{1}{l}\right)\binom{n}{2},$$

and thus $\pi(K_{l+1}) = (1 - \frac{1}{l})$.

More generally, the Erdős-Stone Theorem, proved by Erdős and Stone [17], gives as a corollary (discovered by Erdős and Simonovits [16]) the Turán densities of all graphs. The Turán density of a graph turns out to be a function of its chromatic number, which we now define.

A graph $G = (V, E)$ is said to be *k-colourable* if there exists a *k-colouring* of the graph, that is, a function $f : V \to [k]$ such that, whenever $ab \in E$, we have $f(a) \neq f(b)$. In words, the function f simply assigns colours $1, \ldots, k$ to the vertices of G so that no two vertices that have an edge between them are assigned the same colour. Note that a k-colouring of a graph is a partitioning of its vertices into k *independent* sets (an independent set is a set of vertices containing no edges). The *chromatic number* of G, denoted by $\chi(G)$, is the least k for which G is k-colourable. We can now state the theorem.

Theorem 1.4.3 (Erdős-Stone-Simonovits) *If G is a graph for which*

$\chi(G) \geq 2$ *(that is, G has at least one edge), then*

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

Extremal results often come together with what are known as *stability* results. Such results tell us roughly what the structure of a graph or hypergraph looks like if we relax the conditions on the extremal result. The result below for example, due to Erdős [15] and Simonovits [55], gives us information about F-free graphs having close to the maximum number of edges.

Theorem 1.4.4 (Erdős-Simonovits Stability) *Let F be a graph with chromatic number p. Given $\epsilon > 0$, there exists $\delta = \delta(\epsilon)$ and $n_0 = n_0(\epsilon)$ such that if G is an F-free graph with $n > n_0$ vertices and at least $ex(n, F) - \delta n^2$ edges, then G can be obtained from $Tr(n, p - 1)$ by adding or deleting at most ϵn^2 edges.*

Such stability results, as well as being interesting in their own right, can sometimes be used to derive exact results.

There is another type of stability result. Rather than relaxing the condition on the number of edges of an F-free graph, we can relax the F-freeness itself. We give below the first result of this type, where F is a triangle. The result was proved by Ruzsa and Szemerédi [52] and has since been generalised in many ways [3], [50], some of which we discuss in Chapter 6.

Theorem 1.4.5 (Ruzsa-Szemerédi) *Given $\epsilon > 0$, there exists $\delta = \delta(\epsilon)$ and $n_0 = n_0(\epsilon)$ such that if G is a graph on $n > n_0$ vertices containing at most δn^3 triangles, then G can be made triangle free by deleting at most ϵn^2 edges.*

This theorem is a relatively straightforward consequence of the Regularity Lemma, and in fact, all known proofs of the result are based on the Regularity Lemma. Such results play an important role in the area of property testing.

Turning now to hypergraphs, whereas the Turán densities of all graphs are known, the Turán densities of only a few hypergraphs are known. Let us give a brief history of this area of combinatorics.

An obvious generalisation of Turán's theorem would be to determine the Turán density of K_l^k, the k-graph on l vertices with all edges present. The value of $\pi(K_l^k)$ is unknown for all $l > k \geq 3$. Turán, in his seminal paper [63], conjectured that

$$\pi(K_4^3) = \frac{5}{9}.$$

He conjectured that one possible extremal n-vertex K_4^3-free 3-graph is that in which the n vertices are equipartitioned into three classes, V_1, V_2, V_3, and where the edges are those triples that intersect all the V_i's or contain two vertices from V_i and one from $V_{i+1 (\text{mod } 3)}$ for $i = 1, 2, 3$. This famous problem remains open, although there have been some partial results. The best upper bound known for $\pi(K_4^3)$ is $(3 + \sqrt{17})/12 < 0.593$, which is due to Chung and Lu [7]. The best lower bound remains 5/9, although many different extremal constructions have been found that have the same number of edges as Turán's original construction. In particular, Kostochka [39] gave 2^{n-2} different non-isomorphic extremal constructions on $3n$ vertices for each $n \in \mathbb{N}$. Thus, if the conjecture is true, having so many different extremal structures gives an indication as to why the problem is hard. Bounds on $\pi(K_l^k)$ for other values of k and l are given in the recent paper by Lu and Zhao [41].

Another well-studied problem, which we shall discuss in more detail in Chapter 6, is that of determining the Turán density of K_4^-, the complete 3-uniform hypergraph on four vertices minus an edge. Again, this problem is open: the best known bounds are

$$\frac{2}{7} \leq \pi(K_4^-) \leq \frac{1}{3} - \frac{1}{280}.$$

The upper bound is due to Talbot [59], while the lower bound is due to a construction of Frankl and Füredi [22].

We conclude this subsection by listing a few exact results that have recently been proven. For a survey on some of the earlier work done in this field see [23].

There was a period of little progress in extremal hypergraph theory during the 90's; however, an encouraging breakthrough came when de Caen and Füredi [9] proved that the Turán density of the Fano plane (Figure 1.4) is $\frac{3}{4}$, thus proving a 30 year old conjecture of Sós [56] with a surprisingly simple proof. Keevash and Sudakov [36] later proved, using a stability method, that the unique extremal 3-graph free of the Fano plane is the balanced complete bipartite 3-graph, also conjectured by Sós [56]. The result of de Caen and

Figure 1.1: Fano Plane (3-graph on seven vertices and seven edges)

Füredi drew renewed attention to this field of combinatorics.

Since then, Mubayi and Rödl [44] obtained the Turán densities of about ten different 3-graphs. Füredi, Pikhurko, and Simonovits [24, 25] determined the Turán density and the Turán numbers of the 3-graph with vertex set [5] and edge set $\{123, 124, 125, 345\}$. Keevash and Sudakov [35] determined $\pi(C_3^{(2r)})$, where $C_3^{(2r)}$ is the $2r$-graph obtained by letting P_1, P_2, P_3 be pairwise disjoint sets of size r, and taking as edges the three sets $P_i \cup P_j$ for $i \neq j$. Mubayi [43] gave for each $r \geq 3$, an infinite family of r-graphs together with the Turán density of each r-graph in the family. Pikhurko [48] determined the Turán numbers of each of these r-graphs.

The results we have mentioned here are by no means an exhaustive list of Turán-type results: they simply reflect the author's tastes.

1.5 Summary

We conclude with a brief summary of each chapter.

In **Chapter 2**, we consider the following problem: given two graphs on the same vertex set, can we find a partition of the vertices such that the sizes of the cuts induced in both graphs are large? Let $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ be two graphs, where $|V| = n$ and $|E_i| = m_i$ for $i = 1, 2$. We give a simple algorithm that partitions V into sets A and B such that $e_{G_1}(A, B) \geq m_1/2$ and $e_{G_2}(A, B) \geq m_2/2 - \Delta(G_2)/2$, where $\Delta(G)$ denotes the maximum degree of G. We also show, using probabilistic methods, that if G_1 and G_2 belong to certain classes of graphs (for instance, if G_1 and G_2 both have at least $\frac{1}{3}(n+1)^2$ edges, or if G_1 and G_2 are both regular of degree at most $(n/16) - 8$), then we can find a partition of V into sets A

and B such that $e_{G_i}(A, B) \geq m_i/2$ for $i = 1, 2$. This chapter is based on published work [46].

In **Chapter 3**, we consider natural analogues of graph cut problems for posets. We show that for every poset $P = (X, \prec)$ and every integer $k \geq 2$, there exists an ordered k-partition of P such that the total number of comparable pairs within the parts is at most $(m-1)/k$, where $m \geq 1$ is the total number of comparable pairs in P. We show that this bound is best possible for $k = 2$, but we give an improved bound, $m/k - c(k)\sqrt{m}$, for $k \geq 3$, where $c(k)$ is a constant depending only on k. We also show that, given a poset $P = (X, \prec)$ and an integer $2 \leq k \leq |X|$, we can find an ordered partition of P into k parts that minimises the total number of comparable pairs within parts in time polynomial in the size of P. We prove more general, weighted versions of these results. This chapter is based on published work [47].

Chapter 4 is a digression: in it, we describe the well studied graph linear arrangement problem and consider a natural analogue of this problem for posets. Let $P = (X, \prec)$ be a poset that is not an antichain, and let $\lambda : X \to [n]$ be an order preserving bijection, that is, a linear extension of P. For any relation $a \prec b$ of P, the distance between a and b in λ is $\lambda(b) - \lambda(a)$. The average relational distance of λ, denoted $\text{dist}_P(\lambda)$, is the average of these distances over all relations in P. We show that we can find a linear extension of P that maximises $\text{dist}_P(\lambda)$ in polynomial time. Furthermore, we show that this maximum is at least $\frac{1}{3}(|X| + 1)$, and that this bound is best possible. This chapter is based on joint work with Graham Brightwell.

In **Chapter 5**, we formulate and prove a version of the Szemerédi Regularity Lemma tailored for posets. Going into more detail, we prove that for every poset $P = (X, \prec)$, there exists an ordered partition of X into a bounded number of parts that is regular for both the comparability graph and the covering graph of P (note that the usual Regularity Lemma applied to $\text{Com}(P)$ and $\text{Cov}(P)$ would yield the same result except that the partition would not necessarily be ordered). We prove some interesting properties of ordered regular partitions; in particular, we are able to give some structural properties of P within parts of the regular ordered partition. In light of this, and the fact that the usual Regularity Lemma says nothing about what happens inside the parts of a regular partition, we hope that our version of the

Regularity Lemma for posets will lend itself to novel applications.

In **Chapter 6**, we give a new proof of a result of Frankl and Füredi. The result is a characterisation of all 3-graphs in which every set of four vertices spans either zero or two edges. There are two types of graphs in this characterisation: ones that are 2-colourable and ones that are 3-colourable. (A hypergraph is k-colourable if its vertices can be partitioned into k sets so that no part contains an edge.) We use the ideas in our new proof to give a stability result for the 2-colourable case. More precisely, let $q_i(H)$ denote the number of 4-vertex subsets of V_H that span exactly i edges of H. We prove that if $H = (V_H, E_H)$ is an n-vertex 2-colourable 3-graph in which $q_i(H) < \epsilon n^4$ for $i = 1, 3, 4$, then there exists an n-vertex 2-colourable 3-graph $H' = (V_{H'}, E_{H'})$, where $q_i(H') = 0$ for $i = 1, 3, 4$ and $E_H \triangle E_{H'} \leq 1620\epsilon^{\frac{1}{32}}$.

Chapter 2

Cutting Two Graphs Simultaneously

2.1 Introduction

Throughout this chapter, we work with with finite simple graphs. Recall that for a graph $G = (V, E)$ with A and B disjoint subsets of V, $E_G(A, B)$ denotes the set of edges of G that have one end in A and one end in B, and $e_G(A, B) = |E_G(A, B)|$. Let $A^c = V \backslash A$ so that A, A^c is a partition of V. Then $E_G(A, A^c)$ is a cut of G, and we shall sometimes refer to it as the *cut of G generated by A*. The degree of a vertex x in G is $e_G(x, V)$, shortened to $e_G(x)$, and the maximum degree of G is

$$\Delta(G) := \max_{x \in V} e_G(x).$$

We showed in Section 1.1 that for any graph G with m edges, there exists a cut of size at least $\frac{1}{2}m$ and that the constant $\frac{1}{2}$ cannot be improved. Now consider two graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ on the same vertex set V, where $|V| = n$ and $|E_i| = m_i$ for $i = 1, 2$. We investigate the problem of finding $A \subseteq V$ that generates a large cut both in G_1 and in G_2. This is a problem posed originally by Bollobás and Scott [6]. More precisely, their problem was the following:

Problem 2.1.1 (Bollobás, Scott [6]) *Given $m \in \mathbb{N}$, find the largest integer $f(m)$ such that for every pair of graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$, each with m edges, there exists $A \subseteq V$ with $e_{G_i}(A, A^c) \geq f(m)$ for $i = 1, 2$.*

For further details on this problem and other related problems, see Scott [54].

Bollobás and Scott suggested that perhaps $f(m) = (1-o(1))m/2$. Kühn and Osthus [40] proved this: they showed, using probabilistic methods, that if $G_i = (V, E_i)$ with $|E_i| = m_i$ for $i = 1, 2$, then there exists $A \subseteq V$ such that

$$e_{G_i}(A, A^c) \geq \frac{m_i}{2} - \sqrt{m_i} = (1 - o(1))\frac{m_i}{2}$$

for $i = 1, 2$.

In Section 2.2, we prove the following theorem, which is based on a simple algorithm.

Theorem 2.1.2 *Let $G_i = (V, E_i)$ with $|E_i| = m_i$ for $i = 1, 2$. Then there exists $A \subseteq V$, with $||A| - |A^c|| \leq 1$, such that*

$$e_{G_1}(A, A^c) \geq \frac{m_1}{2} \quad \text{and} \quad e_{G_2}(A, A^c) \geq \frac{m_2}{2} - \frac{\Delta(G_2)}{2}.$$

A simple modification of the algorithm in Theorem 2.1.2 yields the following theorem, which again proves that $f(m) = (1 + o(1))m/2$.

Theorem 2.1.3 *Let G_1 and G_2 be graphs as in Theorem 2.1.2. Then there exists $A \subseteq V$, with $||A| - |A^c|| \leq 2$, such that*

$$e_{G_1}(A, A^c) \geq \frac{m_1}{2} \quad \text{and} \quad e_{G_2}(A, A^c) \geq \frac{m_2}{2} - \sqrt{m_2}.$$

Theorem 2.1.2 also proves the following conjecture of Rautenbach and Szigeti [49].

Conjecture 2.1.4 (Rautenbach, Szigeti [49]) *Let $G_i = (V, E_i)$, where $|E_i| = m_i$ for $i = 1, 2$. If both graphs have maximum degree at most Δ then there exists $A \subseteq V$ such that $e_{G_i}(A, A^c) \geq \frac{1}{2}(m_i - \Delta)$ for $i = 1, 2$.*

The following conjecture, which is implicit in [40] and [49] but not formally stated, is a natural extension of Conjecture 2.1.4.

Conjecture 2.1.5 *Let $G_i = (V, E_i)$ with $|E_i| = m_i$ for $i = 1, 2$. Then there exists $A \subseteq V$ such that $e_{G_i}(A, A^c) \geq \lfloor \frac{1}{2} m_i \rfloor$ for $i = 1, 2$.*

Note that the above conjecture is false if we replace $\lfloor \frac{1}{2} m_i \rfloor$ with $\frac{1}{2} m_i$. Indeed, let G_1 be a 5-cycle on 5 vertices and let G_2 be its complementary graph (also a 5-cycle). Taking any partition of the vertices generates a cut in at least one of the two graphs having at most $\lfloor 5/2 \rfloor$ edges. This example was given by Rautenbach and Szigeti [49], and is the only such example that we know of.

In Section 2.3 we show, using probabilistic methods similar to those used in [40], that Conjecture 2.1.5 holds for certain classes of graphs. More precisely, we prove the following two theorems.

Theorem 2.1.6 *Let $G_i = (V, E_i)$ with $|V| = n$ and $|E_i| = m_i \geq \frac{1}{3}(n+1)^2$ for $i = 1, 2$. Then there exists $A \subseteq V$, with $||A| - |A^c|| \leq 1$, satisfying*

$$e_{G_i}(A, A^c) \geq \frac{m_i}{2}$$

for $i = 1, 2$.

Theorem 2.1.7 *Let $G_i = (V, E_i)$ with $|V| = n$ and $|E_i| = m_i$ for $i = 1, 2$. If $r_i := \Delta(G_i) < \sqrt{m_i/8} - 2$ for $i = 1, 2$, then there exists $A \subseteq V$, with $||A| - |A^c|| \leq 1$, satisfying*

$$e_{G_i}(A, A^c) \geq \frac{m_i}{2}$$

for $i = 1, 2$.

In particular, the conditions of Theorem 2.1.7 are met if G_i is r_i-regular, with $1 \leq r_i < \frac{n}{16} - 8$ for $i = 1, 2$. Let us carry out the simple calculation to see that this is true. Assuming that $1 \leq r_i < \frac{n}{16} - 8$ and using $m_i = (r_i n)/2$, we have

$$\begin{aligned} (r_i + 2)^2 &\leq r_i(r_i + 8) && \text{provided } 4r_i \geq 4 \text{ i.e. } r_i \geq 1 \\ &< r_i(n/16) = m_i/8. \end{aligned}$$

Rearranging the above by taking square roots and subtracting 2 gives the condition of Theorem 2.1.7.

2.2 A Simple Algorithm

In this section, we present the proofs of Theorems 2.1.2 and 2.1.3. The proof of Theorem 2.1.2 is based on a simple algorithm, which we later adapt to give a slightly sharper result for the case when $\Delta(G_2)$ is large. Before we proceed to the proof of Theorem 2.1.2, recall that for $G = (V, E)$ a graph, we write $e_G(x, A)$ to denote the number of neighbours of a vertex x in a set $A \subseteq V$.

Proof (of Theorem 2.1.2) We are given $G_i = (V, E_i)$ for $i = 1, 2$, where $|V| = n$ and $|E_i| = m_i$. We assume that n is even. (If n is odd then we add a vertex to V that is isolated in G_1 and G_2 and apply the theorem for n even.)

For $j = 0, \ldots, n/2$, we inductively construct disjoint subsets A_j and B_j of V as follows. Let $A_0 = B_0 = \phi$, and assume that we have constructed $A_{j-1} = \{a_1, \ldots, a_{j-1}\}$ and $B_{j-1} = \{b_1, \ldots, b_{j-1}\}$.

For each $v \in V$, let

$$e_1^j(v) = e_{G_1}(v, B_{j-1}) - e_{G_1}(v, A_{j-1}).$$

Choose a_j to be any vertex in $V \backslash (A_{j-1} \cup B_{j-1})$ that maximises e_1^j, and set $A_j = \{a_1, \ldots, a_j\}$. For each $v \in V$, let

$$e_2^j(v) = e_{G_2}(v, A_j) - e_{G_2}(v, B_{j-1}).$$

Choose b_j to be any vertex in $V \backslash (A_j \cup B_{j-1})$ that maximises e_2^j, and set $B_j = \{b_1, \ldots, b_j\}$. Notice, by our choices of a_j and b_j, that for each j, we have

$$e_1^j(a_j) \geq e_1^j(b_j) \quad \text{and} \quad e_2^j(b_j) \geq e_2^j(a_{j+1}) \geq e_2^{j+1}(a_{j+1}).$$

(Note that we have the strict inequality $e_2^j(a_{j+1}) > e_2^{j+1}(a_{j+1})$ if and only if there is an edge of G_2 between a_{j+1} and b_j.) We shall use these inequalities at the end.

After $n/2$ iterations, we obtain $A_{n/2}$ and $B_{n/2}$, sets of equal sizes that partition V. Let $A = A_{n/2}$, so that $A^c = B_{n/2}$. We claim that

$$e_{G_1}(A, A^c) \geq \frac{m_1}{2} \quad \text{and} \quad e_{G_2}(A, A^c) \geq \frac{m_2}{2} - \frac{\Delta(G_2)}{2}.$$

To see this, observe that for $i = 1, 2$, we have

$$m_i = \sum_{j=1}^{n/2} [e_{G_i}(a_j, A_{j-1}) + e_{G_i}(a_j, B_{j-1}) + e_{G_i}(b_j, A_j) + e_{G_i}(b_j, B_{j-1})]$$

and

$$e_{G_i}(A, A^c) = \sum_{j=1}^{n/2} [e_{G_i}(a_j, B_{j-1}) + e_{G_i}(b_j, A_j)].$$

Subtracting 1/2 of the first equation from the second yields

$$e_{G_i}(A, A^c) - \frac{m_i}{2} = \frac{1}{2}\sum_{j=1}^{n/2} \Big([e_{G_i}(a_j, B_{j-1}) - e_{G_i}(a_j, A_{j-1})] + [e_{G_i}(b_j, A_j) - e_{G_i}(b_j, B_{j-1})] \Big)$$

for $i = 1, 2$.

By comparing the terms in square brackets with $e_i^j(a_j)$ and $e_i^j(b_j)$ respectively, and noting for any vertex v that $e_{G_i}(v, A_j) \geq e_{G_i}(v, A_{j-1})$, we obtain that

$$e_{G_i}(A, A^c) - \frac{m_i}{2} \geq \begin{cases} \frac{1}{2}\sum_{j=1}^{n/2} (e_1^j(a_j) - e_1^j(b_j)) & \text{if } i = 1; \\ \frac{1}{2}\sum_{j=1}^{n/2} (e_2^j(b_j) - e_2^j(a_j)) & \text{if } i = 2. \end{cases}$$

Using that $e_1^j(a_j) \geq e_1^j(b_j)$ for each j, we see that the first sum is non-negative. Using that $e_2^j(b_j) \geq e_2^{j+1}(a_{j+1})$ for each j , we see that the second sum is at least $-e_2^1(a_1) + e_2^{n/2}(b_{n/2}) \geq -\Delta(G_2)$ as $e_2^1(a_1) = 0$. This completes the proof. □

Examining the proof of Theorem 2.1.2, we see that it is the last vertex placed that determines the size of $e_{G_2}(A, A^c) - (m_2/2)$. In particular, we can improve on Theorem 2.1.2 if we can ensure that the degree of $b_{n/2}$ in G_2 is small.

Proof (of Theorem 2.1.3) Let $v_1, \ldots, v_n$ be an ordering of the vertices of V satisfying $e_{G_2}(v_i) \geq e_{G_2}(v_{i+1})$ for all $i = 1, \ldots, n-1$. Let $V^* = \{v_1, \ldots, v_t\}$, where t is an integer to be specified later. For convenience, we ensure that both $|V^*|$ and $|V|$ are even by adding isolated vertices to V^* and/or $(V^*)^c = V \backslash (V^*)$ if necessary. After the addition of these isolated vertices, let $t' = |V^*|$ and $n' = |V|$. We give a modified version of the algorithm in the proof of Theorem 2.1.2. The only difference is that initially, we restrict our attention to V^*, however we describe the algorithm in full for notational convenience.

Let $V_j = V^*$ for $j \leq t'/2$ and $V_j = (V^*)^c$ for $j > t'/2$. For $j = 0, \ldots, n'/2$, we inductively construct disjoint subsets A_j and B_j of V as follows. Let $A_0 = B_0 = \phi$, and assume that we have constructed $A_{j-1} = \{a_1, \ldots, a_{j-1}\}$ and $B_{j-1} = \{b_1, \ldots, b_{j-1}\}$.

For each $v \in V_j$, let

$$e_1^j(v) = e_{G_1}(v, B_{j-1}) - e_{G_1}(v, A_{j-1}).$$

Choose a_j to be any vertex in $V_j \backslash (A_{j-1} \cup B_{j-1})$ that maximises e_1^j, and set $A_j = \{a_1, \ldots, a_j\}$.

For each $v \in V_j$, let

$$e_2^j(v) = e_{G_2}(v, A_j) - e_{G_2}(v, B_{j-1}).$$

Choose b_j to be any vertex in $V_j \backslash (A_j \cup B_{j-1})$ that maximises e_2^j, and set $B_j = \{b_1, \ldots, b_j\}$.

We iterate $n'/2$ times to obtain sets $A_{n'/2}$ and $B_{n'/2}$. We remove from $A_{n'/2}$ and $B_{n'/2}$ any isolated vertices that we may have added at the beginning to obtain sets A and $B = A^c$ that partition V. Note that $||A| - |A^c|| \leq 2$. This completes the description of the modified algorithm.

Note, by our choices of a_j and b_j, that for each j, we have $e_1^j(a_j) \geq e_1^j(b_j)$, and for each j except $j = t'/2$, we have $e_2^j(b_j) \geq e_2^j(a_{j+1}) \geq e_2^{j+1}(a_{j+1})$.

Mimicking the analysis of the algorithm in Theorem 2.1.2 and noting that $e_{G_i}(A, A^c) = e_{G_i}(A_{n'/2}, B_{n'/2})$, we find that

$$e_{G_1}(A, A^c) - \frac{m_1}{2} \geq \frac{1}{2} \sum_{j=1}^{n'/2} (e_1^j(a_j) - e_1^j(b_j)) \geq 0,$$

and

$$\begin{aligned} e_{G_2}(A, A^c) - \frac{m_2}{2} &\geq \frac{1}{2} \sum_{j=1}^{n'/2} (e_2^j(b_j) - e_2^j(a_j)) \\ &\geq \frac{1}{2}(-e_2^1(a_1) + e_2^{t'/2}(b_{t'/2}) - e_2^{(t'/2)+1}(a_{(t'/2)+1}) + e_2^{n'/2}(b_{n'/2})) \\ &\geq \frac{1}{2}\left(0 - \left\lfloor \frac{t}{2} \right\rfloor - \left\lceil \frac{t}{2} \right\rceil - e_{G_2}(v_{t+1})\right) \\ &= -\frac{1}{2}(t + e_{G_2}(v_{t+1})), \end{aligned}$$

where the third inequality follows because $e_2^1(a_1) = 0$,

$$e_2^{t'/2}(b_{t'/2}) \geq -e_{G_2}(b_{t'/2}, B_{(t'/2)-1}) \geq -\left\lfloor \frac{t}{2} \right\rfloor,$$

and

$$-e_2^{(t'/2)+1}(a_{(t'/2)+1}) \geq -e_{G_2}(a_{(t'/2)+1}, B_{t'/2}) \geq -\left\lceil \frac{t}{2} \right\rceil.$$

Since we are free to choose t as we please, we have that

$$e_{G_2}(A, A^c) - \frac{m_2}{2} \geq -\frac{1}{2} \min_t [t + e_{G_2}(v_{t+1})],$$

where we minimise over $t = 0, \ldots, n-1$. We claim that

$$\min_t [t + e_{G_2}(v_{t+1})] \leq \lfloor 2\sqrt{m_2} \rfloor,$$

which proves the theorem. We prove the claim by contradiction. Suppose that $t + e_{G_2}(v_{t+1}) \geq \lceil 2\sqrt{m_2} \rceil$ for all $t = 0, \ldots, n-1$. Then

$$\begin{aligned}
\sum_{t=0}^{n-1} e_{G_2}(v_{t+1}) &\geq \sum_{t=0}^{n-1} \max[(\lceil 2\sqrt{m_2} \rceil - t), 0] \\
&= \sum_{t=0}^{\lceil 2\sqrt{m_2} \rceil} t \\
&= \frac{1}{2} \lceil 2\sqrt{m_2} \rceil (\lceil 2\sqrt{m_2} \rceil + 1) \\
&> 2m_2,
\end{aligned}$$

which is a contradiction, proving the claim. □

2.3 Good Simultaneous Cuts for Special Classes of Graphs

In this section, we turn to the problem of finding pairs of graphs, $G_i = (V, E_i)$ with $|E_i| = m_i$ for $i = 1, 2$, for which we can ensure the existence of $A \subseteq V$ such that

$$e_{G_i}(A, A^c) \geq \lfloor m_i/2 \rfloor$$

for $i = 1, 2$. As conjectured earlier, we believe that the above is true for all pairs of graphs. The proofs in this section are of a probabilistic nature.

We first prove that the above is true for graphs of high density, that is, those graphs that give the poorest bounds in Theorem 2.1.2 and Theorem 2.1.3. We start with a general lemma.

Lemma 2.3.1 *Let X be a random variable taking values in $\{0, \dots, n\}$ for some $n \in \mathbb{N}$. Let μ be its mean and σ^2 its variance. For $p > 0$, let $r(X,p)$ be the largest integer such that*

$$\mathbb{P}(X \leq r(X,p)) < p.$$

Then

$$r(X,p) + 1 \geq \mu - \sqrt{\frac{1-p}{p}}\sigma.$$

Proof Let Y be the two point random variable taking the value $y_0 := \mathbb{E}(X|X \leq r(X,p)+1)$ with probability $p_0 := \mathbb{P}(X \leq r(X,p)+1) > p$, and taking the value $y_1 := \mathbb{E}(X|X > r(X,p)+1)$ with probability $p_1 := 1 - p_0$. Note that $Y = \mathbb{E}(X|I)$, where I is the indicator function of the event that $X \leq r(X,P)+1$.

We carry out an easy calculation to show that

$$y_0 = \mathbb{E}(Y) - \sqrt{\frac{1-p_0}{p_0}Var(Y)}\,ar(Y). \tag{2.1}$$

Indeed, we have that $\mathbb{E}(Y) = p_0 y_0 + (1-p_0) y_1$ and that $Var(Y) = p_0(\mathbb{E}(Y) - y_0)^2 + (1-p_0)(y_1 - \mathbb{E}(Y))^2$. We eliminate y_1 between the equations as follows:

$$\begin{aligned}
\frac{1-p_0}{p_0}Var(Y) &= \frac{1-p_0}{p_0}\left(p_0(\mathbb{E}(Y) - y_0)^2 + (1-p_0)(y_1 - \mathbb{E}(Y))^2\right) \\
&= \frac{1-p_0}{p_0}\left(p_0(1-p_0)(y_1 - y_0)^2 + (1-p_0)p_0(y_1 - y_0)^2\right) \\
&= [(1-p_0)(y_1 - y_0)]^2 \\
&= (\mathbb{E}(Y) - y_0)^2.
\end{aligned}$$

Rearranging gives (2.1). We note that $\mathbb{E}(Y) = \mu$, and we claim that

$Var(Y) \leq \sigma^2$; this is deduced from the convexity of x^2 as follows:

$$\begin{aligned}
Var(Y) &= \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 \\
&= (p_0 y_0^2 + p_1 y_1^2) - \mathbb{E}(X)^2 \\
&= p_0 \Bigg(\sum_{i=0}^{r(X,p)+1} \frac{\mathbb{P}(X=i)}{p_0} i \Bigg)^2 + p_1 \Bigg(\sum_{i=r(X,p)+2}^{n} \frac{\mathbb{P}(X=i)}{p_1} i \Bigg)^2 - \mathbb{E}(X)^2 \\
&\leq p_0 \sum_{i=0}^{r(X,p)+1} \frac{\mathbb{P}(X=i)}{p_0} i^2 + p_1 \sum_{i=r(X,p)+2}^{n} \frac{\mathbb{P}(X=i)}{p_1} i^2 - \mathbb{E}(X)^2 \\
&= Var(X) = \sigma^2.
\end{aligned}$$

Using that $y_0 \leq r(X,p)+1$; $p_0 \geq p$; $\mathbb{E}(Y) = \mu$; and $Var(Y) \leq \sigma^2$ together with (2.1), we obtain

$$r(X,p) + 1 \geq y_0 = \mathbb{E}(Y) - \sqrt{\frac{1-p_0}{p_0} Var(Y)} \geq \mu - \sqrt{\frac{1-p}{p}}\sigma,$$

as required. □

The following corollary is the main probabilistic tool used in the proofs of Theorem 2.1.6 and Theorem 2.1.7.

Corollary 2.3.2 *Let X_1 and X_2 be random variables taking values from $\{0,\dots,n\}$, and let X_i have mean μ_i and variance σ_i^2 for $i = 1, 2$. Then*

$$\mathbb{P}(X_1 \geq \mu_1 - \sigma_1,\ X_2 \geq \mu_2 - \sigma_2) > 0.$$

Proof The following easy calculation proves the corollary. We have

$$\begin{aligned}
\mathbb{P}(X_1 \geq \mu_1 - \sigma_1, X_2 \geq \mu_2 - \sigma_2) &= \mathbb{P}(X_1 \geq \lceil \mu_1 - \sigma_1 \rceil,\ X_2 \geq \lceil \mu_2 - \sigma_2 \rceil) \\
&\geq 1 - \mathbb{P}(X_1 \leq \lceil \mu_1 - \sigma_1 \rceil - 1) - \mathbb{P}(X_2 \leq \lceil \mu_2 - \sigma_2 \rceil - 1) \\
&\geq 1 - \mathbb{P}(X_1 \leq r(X_1, 1/2)) - \mathbb{P}(X_2 \leq r(X_2, 1/2)) \\
&> 1 - 1/2 - 1/2 = 0.
\end{aligned}$$

□

The idea of the proof of Theorem 2.1.6 is an extension of the ideas of Kühn and Osthus [40].

Proof (of Theorem 2.1.6) The graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ are given. We may assume that $n = |V|$ is even by adding a vertex to V that is isolated in both G_1 and G_2 if n is odd; consequently, we must work with the weaker premise that $m_i \geq \frac{1}{3}n^2$ rather than the original premise of the theorem that $m_i \geq \frac{1}{3}(n+1)^2$. We also assume $n \geq 4$; the cases for $n \leq 3$ can easily be checked by hand.

Pick a subset A of V of size $n/2$ uniformly at random, and set $X_i = e_{G_i}(A, A^c)$ for $i = 1, 2$. Let μ_i and σ_i^2 be respectively the mean and variance of X_i for $i = 1, 2$. We show that if G_1 and G_2 are sufficiently dense, then

$$\mu_i - \sigma_i \geq \frac{m_i}{2}.$$

Corollary 2.3.2 then gives that

$$\mathbb{P}(X_1 \geq m_1/2, X_2 \geq m_2/2) > 0;$$

hence there exists some subset of V that induces a cut in G_1 (resp. G_2) containing at least $\frac{1}{2}m_1$ (resp. $\frac{1}{2}m_2$) edges.

It remains only to bound $\mu_i - \sigma_i$. We start by computing the expectation and variance of the X_i. Let us focus on X_1. For each $e \in E_1$, define

$$X_e = \begin{cases} 1 \text{ if } e \in E_{G_1}(A, A^c); \\ 0 \text{ otherwise.} \end{cases}$$

Note that $X_1 = \sum_{e \in E_1} X_e$. Using this and the linearity of expectation, we have

$$\mu_1 = \mathbb{E}(X_1) = \sum_{e \in E_1} \mathbb{E}(X_e) = \sum_{e \in E_1} \mathbb{P}(e \in E_{G_1}(A, A^c)) = \frac{1}{2}m_1(1 + \frac{1}{n-1}),$$

where we have used that

$$\begin{aligned}\mathbb{P}(xy = e \in E_{G_1}(A, A^c)) &= \mathbb{P}(x \in A, y \in A^c) + \mathbb{P}(y \in A, x \in A^c) \\ &= \left(\frac{\frac{1}{2}n}{n}\right)\left(\frac{\frac{1}{2}n}{n-1}\right) + \left(\frac{\frac{1}{2}n}{n}\right)\left(\frac{\frac{1}{2}n}{n-1}\right) \\ &= \frac{1}{2}\left(1 + \frac{1}{n-1}\right).\end{aligned}$$

Next we compute $\mathbb{E}(X_1^2)$. Again, writing X_1 as a sum of indicator functions and expanding, we get

$$\begin{aligned}\mathbb{E}(X_1^2) &= \sum_{e\in E_1}\mathbb{E}(X_e\cdot X_e) + \sum_{\substack{e,f\in E_1\\ e\neq f}}\mathbb{E}(X_e\cdot X_f)\\ &= \mathbb{E}(X) + \sum_{\substack{e,f\in E_1\\ e\neq f}}\mathbb{P}(e,f\in E_{G_1}(A,A^c)). \end{aligned}\tag{2.2}$$

For two edges e and f of a graph, with $e \neq f$, we write e inc f if the edges are incident (meet at exactly one vertex), and we write e ind f if they have no common vertices, that is, they are independent. We split the sum above according to whether or not e and f are incident. We claim that

$$\mathbb{P}(e,f\in E_{G_1}(A,A^c)) = \begin{cases}\frac{1}{4}\Big(1+\frac{1}{n-1}\Big)\Big(1+\frac{1}{n-3}\Big), & \text{if } e \text{ ind } f;\\ \frac{1}{4}\Big(1+\frac{1}{n-1}\Big), & \text{if } e \text{ inc } f.\end{cases}\tag{2.3}$$

Indeed, if e ind f with $e = xy$ and $f = x'y'$, then we have $e, f \in E_G(A, A^c)$ if and only if one of the following four equally likely and mutually exclusive events occurs: either $x, x' \in A$ and $y, y' \in A^c$; or $x, y' \in A$ and $y, x' \in A^c$; or $y, x' \in A$ and $x, y' \in A^c$; or $y, y' \in A$ and $x, x' \in A^c$. Thus, in this case we have

$$\begin{aligned}\mathbb{P}(e,f\in E_{G_1}(A,A^c)) &= 4\Big(\frac{\frac{1}{2}n}{n}\Big)\Big(\frac{\frac{1}{2}n-1}{n-1}\Big)\Big(\frac{\frac{1}{2}n}{n-2}\Big)\Big(\frac{\frac{1}{2}n-1}{n-3}\Big)\\ &= \frac{1}{4}\Big(1+\frac{1}{n-1}\Big)\Big(1+\frac{1}{n-3}\Big).\end{aligned}$$

If e inc f with $e = xa$ and $f = xb$, then we have $e, f \in E_G(A, A^c)$ if and only if one of the following two equally likely and mutually exclusive events occurs: either $x \in A$ and $a, b \in A^c$; or $a, b \in A$ and $x \in A^c$. In this case we have

$$\begin{aligned}\mathbb{P}(e,f\in E_{G_1}(A,A^c)) &= 2\Big(\frac{\frac{1}{2}n}{n}\Big)\Big(\frac{\frac{1}{2}n}{n-1}\Big)\Big(\frac{\frac{1}{2}n-1}{n-2}\Big)\\ &= \frac{1}{4}\Big(1+\frac{1}{n-1}\Big).\end{aligned}$$

Using (2.2) and (2.3), we obtain

$$
\begin{aligned}
\mathbb{E}(X_1^2) &= \mathbb{E}(X) + \sum_{\substack{e,f\in E_1\\ e \text{ ind } f}} \mathbb{P}(e,f \in E_{G_1}(A,A^c)) + \sum_{\substack{e,f\in E_1\\ e \text{ inc } f}} \mathbb{P}(e,f \in E_{G_1}(A,A^c)) \\
&= \frac{1}{2}m_1(1+\frac{1}{n-1}) + [m_1(m_1-1) - P_2(G_1)]\frac{1}{4}\Big(1+\frac{1}{n-1}\Big)\Big(1+\frac{1}{n-3}\Big) \\
&\quad + P_2(G_1)\frac{1}{4}\Big(1+\frac{1}{n-1}\Big) \\
&= \frac{1}{4}\Big(1+\frac{1}{n-1}\Big)\Big(2m_1 + \Big(1+\frac{1}{n-3}\Big)m_1(m_1-1) - \frac{P_2(G_1)}{n-3}\Big),
\end{aligned}
$$

where $P_2(G_1)$ denotes the number of (ordered) pairs of incident edges in G_1. Alternatively, $P_2(G_1)$ is twice the number of paths of length 2 in G_1, and we can bound it as follows. Let $v_1, \dots, v_n$ be the vertices in V, and let d_i be the degree of v_i in G_1. Then

$$
\begin{aligned}
P_2(G_1) &= \sum_{i=1}^{n} d_i(d_i-1) \\
&= \sum_{i=1}^{n} d_i^2 - 2m_1 \\
&\geq n\Big(\frac{1}{n}\sum_{i=1}^{n} d_i\Big)^2 - 2m_1 \text{ (Cauchy-Schwarz inequality)} \\
&= \frac{4m_1^2}{n} - 2m_1.
\end{aligned}
$$

Using this bound, together with the expression for $\mathbb{E}(X_1^2)$, we find that

$$
\begin{aligned}
\mathbb{E}(X_1^2) &\leq \frac{1}{4}\Big(1+\frac{1}{n-1}\Big)\Big(m_1 + \Big(1+\frac{1}{n-3}\Big)m_1(m_1-1) \\
&\qquad - \frac{4m_1^2}{n(n-3)} + \frac{2m_1}{n-3}\Big) \\
&= \frac{1}{4}\Big(1+\frac{1}{n-1}\Big)\Big(m_1 + \Big(1+\frac{1}{n-3}\Big)m_1^2 - \frac{4m_1^2}{n(n-3)} + \frac{m_1}{n-3}\Big).
\end{aligned}
$$

Using our expression for $\mathbb{E}(X_1)$, we obtain

$$\begin{aligned}
\sigma_1^2 &= Var(X_1) = \mathbb{E}(X_1^2) - \mathbb{E}(X_1)^2 \\
&\leq \frac{1}{4}\Big(1+\frac{1}{n-1}\Big)\Big(m_1 + \Big(1+\frac{1}{n-3}\Big)m_1^2 - \frac{4m_1^2}{n(n-3)} + \frac{m_1}{n-3}\Big) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad -\frac{1}{4}\Big(1+\frac{1}{n-1}\Big)^2 m_i^2 \\
&= \frac{1}{4}\Big(1+\frac{1}{n-1}\Big)\Big(m_1 + \frac{2m_1^2}{(n-1)(n-3)} - \frac{4m_1^2}{n(n-3)} + \frac{m_1}{n-3}\Big) \\
&= \frac{1}{4}\Big(\frac{n}{n-1}\Big)\Big(\Big(\frac{n-2}{n-3}\Big)m_1 - \Big(\frac{2n-4}{n(n-1)(n-3)}\Big)m_1^2\Big),
\end{aligned}$$

and similarly for σ_2^2. Recall: we wish to show that when G_1 and G_2 are sufficiently dense, we have

$$\mu_i - \sigma_i \geq \frac{m_i}{2} \quad \text{for } i = 1, 2, \tag{2.4}$$

or equivalently

$$\sigma_i^2 \leq \Big(\mu_i - \frac{m_i}{2}\Big)^2 \quad \text{for } i = 1, 2.$$

Substituting the expression for μ_i and the bound for σ_i^2, we find it is sufficient to show that

$$\frac{1}{4}\Big(\frac{n}{n-1}\Big)\Big(\Big(\frac{n-2}{n-3}\Big)m_i - \Big(\frac{2n-4}{n(n-1)(n-3)}\Big)m_i^2\Big) \leq \frac{m_i^2}{4(n-1)^2}.$$

Since we have assumed that $n \geq 4$, we can divide both sides by $\frac{1}{4}\frac{n}{n-1}$. Then collecting the m_i^2 terms on the right hand side, we find that the previous inequality is equivalent to

$$\begin{aligned}
\Big(\frac{n-2}{n-3}\Big)m_i &\leq \Big(\frac{1}{n(n-1)} + \Big(\frac{2n-4}{n(n-1)(n-3)}\Big)\Big)m_i^2 \\
&= \Big(\frac{3n-7}{n(n-1)(n-3)}\Big)m_i^2.
\end{aligned}$$

Thus, since $n > 3$ and $m_i > 0$, we can rearrange the above to deduce that (2.4) holds if

$$m_i \geq \frac{n(n-1)(n-2)}{3n-7},$$

which holds if $m_i \geq \frac{1}{3}n^2$ (assuming $n \geq 3$) for $i = 1, 2$. □

Next we prove a theorem showing that pairs of graphs with small maximum degree (relative to the number of edges in the graphs) also satisfy Conjecture 2.1.4. The proof of the theorem broadly follows that of the previous theorem; the only difference is the way in which the random cut is constructed.

Going into more detail, the random cut is constructed as follows. We first deterministically pair up the vertices of our vertex set V so that a large proportion of the pairs form edges of our graphs. We then partition V randomly, ensuring that vertices of each pair are in different parts.

This motivates the following lemma and its corollary.

Lemma 2.3.3 *For graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$, let $A_i \subseteq E_i$ be sets of independent edges for $i = 1, 2$. Then there exists a set $A \subseteq A_1 \cup A_2$ of independent edges such that for $i = 1, 2$ we have*

$$|A \cap A_i| \geq \left\lfloor \frac{1}{2}|A_i| \right\rfloor - 1.$$

Proof Observe that each edge of $A_1 \cap A_2$ is independent of all other edges in $A_1 \cup A_2$. Let $B_i = A_i \backslash (A_1 \cap A_2)$. Then it is sufficient to find a set $B \subseteq B_1 \cup B_2$ of independent edges such that $|B \cap B_i| \geq \lfloor |B_i|/2 \rfloor - 1$ for $i = 1, 2$ (then set $A = B \cup (A_1 \cap A_2)$).

We construct B as follows. Assume, without loss of generality, that $|B_2| = |B_1| + b$, where b is a non-negative integer. Note that $B_1 \cup B_2$ is a disjoint union of paths and cycles in which edges alternate between being in B_1 and being in B_2. Let S be the set of these paths and cycles.

A path in S whose first and last edges are both in B_1 (resp. B_2) is referred to as a 1-*path* (resp. 2-*path*). Let P^1 (resp. P^2) be the set of 1-paths (resp. 2-paths). Any other path in S is necessarily a path with an even number of edges, so we call it an *even path.* Let P^e be the set of even paths in S. Let C be the set of cycles in S (each of which necessarily has an even number of edges).

We have that S is the disjoint union of C, P^e, P^1 and P^2. For $s \in S$,

$|s|$ refers to the number of edges in s. Let

$$\begin{aligned}
C &= \{c_1, c_2, \ldots, c_j\} \text{ with } |c_1| \geq |c_2| \geq \ldots \geq |c_j|, \\
P^e &= \{p_1^e, p_2^e, \ldots, p_k^e\} \text{ with } |p_1^e| \geq |p_2^e| \geq \ldots \geq |p_k^e|, \\
P^1 &= \{p_1^1, p_2^1, \ldots, p_l^1\} \text{ with } |p_1^1| \geq |p_2^1| \geq \ldots \geq |p_l^1|, \\
P^2 &= \{p_1^2, p_2^2, \ldots, p_m^2\} \text{ with } |p_1^2| \geq |p_2^2| \geq \ldots \geq |p_m^2|,
\end{aligned}$$

and note that the number of 2-paths exceeds the number of 1-paths by $b = |B_2| - |B_1|$; hence we have $l + b = m$. We order the elements of S

$$c_1, c_2, \ldots, c_j, p_1^e, p_2^e, \ldots, p_k^e, p_1^2, p_1^1, p_2^2, p_2^1, \ldots, p_l^2, p_l^1, p_{l+1}^2, p_{l+2}^2, \ldots, p_m^2,$$

and we call this ordering O_S. For each $s \in S$, fix an ordering $f_1, \ldots, f_q$ of the edges of s such that f_i and f_{i+1} are incident for $i = 1, \ldots, q-1$, and if s is an even path or cycle, then we choose f_1 to be in B_2. Concatenate these orderings of elements of S according to O_s to give an ordering $e_1, \ldots, e_t$ of the edges of $B_1 \cup B_2$. Note that the edges in our ordering $e_1, \ldots, e_t$ alternate between B_1 and B_2 except at a transition between P_{l+z}^2 and P_{l+z+1}^2 $(z = 1, \ldots, b-1)$, where we have two consecutive edges in B_2. We call such a transition, a *P^2-transition.*

Choose x minimal such that $|\{e_1, \ldots, e_x\} \cap B_1| = \lfloor |B_1|/2 \rfloor - 1$, and let $B_1' = \{e_1, \ldots, e_x\} \cap B_1$. Let $B_2' = \{e_{x+2}, \ldots, e_t\} \cap B_2$, and let $B = B_1' \cup B_2'$. It is not too difficult to see that B is a set of independent edges. (Indeed, B_1' and B_2' are each independent sets of edges. Furthermore the only way that an edge from $\{e_1, \ldots, e_x\}$ can be incident with an edge from $\{e_{x+2}, \ldots e_t\}$ is if e_x lies in a cycle from C. Even then, there is only one incidence, and we have ensured that we avoid the incidence in $B_1' \cup B_2'$ by our choice of having the first edge of every cycle belong to G_2.)

It remains only to show that $|B \cap B_2| = |B_2'| \geq \lfloor |B_2|/2 \rfloor - 1$. Let y be the number of P^2-transitions in $e_1, \ldots, e_x$. We have

$$\begin{aligned}
|B_2'| &= |B_2| - |\{e_1, \ldots, e_x\} \cap B_2\}| - 1 \\
&= |B_2| - (|B_1'| + y) - 1.
\end{aligned}$$

(The -1 is present in the first equality because e_{x+1} is necessarily an edge of B_2. The second equality follows from the alternating correspondence of B_1 and B_2 edges except at P^2-transitions.) Since $p^2_{l+1}, \ldots, p^2_m$ are ordered according to size in O_s, we find that $y \leq b/2 - 1$, otherwise $|B_1'| \geq |B_1|/2$. Using this and that $|B_1'| = \lfloor |B_1|/2 \rfloor - 1$, we get

$$\begin{aligned} |B_2'| &= |B_2| - (|B_1'| + y) - 1 \\ &\geq |B_2| - \frac{1}{2}(|B_1| + b) - 1 \\ &\geq \left\lfloor \frac{1}{2}|B_2| \right\rfloor - 1, \end{aligned}$$

as required. □

Corollary 2.3.4 *Let $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ be graphs with $|V| = n$ and $|E_i| = m_i$. If $\Delta(G_i) = r_i$ for $i = 1, 2$, then there exists a pairing, $P = \{(v_1, w_1), \ldots, (v_{\lfloor n/2 \rfloor}, w_{\lfloor n/2 \rfloor})\}$, of the vertices in V such that for $i = 1, 2$, we have*

$$|P \cap E_i| \geq \left\lfloor \frac{m_i}{2(r_i + 1)} \right\rfloor - 1.$$

The proof of Corollary 2.3.4 requires the following well-known result of Vizing.

Theorem 2.3.5 (Vizing [64]) *Every graph G has an edge colouring (a colouring of the edges of G so that no incident edges receive the same colour) that requires either $\Delta(G)$ or $\Delta(G) + 1$ colours.*

In fact, Vizing proved the above for multigraphs, but we require the result only for graphs. See, for example, [11] for a proof.

Proof (of Corollary 2.3.4) By Vizing's Theorem, we can find an $(r_i + 1)$-colouring of the edges of G_i, and so, in each of the graphs G_i, we can find an independent set of edges of size at least $m_i/(r_i + 1)$. Let A_i be such a set of independent edges for the graph G_i. By Lemma 2.3.3, we know that there exists a set $A \subseteq A_1 \cup A_2$ of independent edges such that $|A \cap A_i| \geq \lfloor |A_i|/2 \rfloor - 1$. This proves the corollary since the edges in A induce a partial pairing of V and we extend this (in any way) to a total pairing P with the desired property. □

We are now ready to prove Theorem 2.1.7.

Proof (of Theorem 2.1.7) Assume n is even (if not, add a vertex to V that is isolated in G_1 and G_2, and apply the theorem for the case when n is even). By Corollary 2.3.4, there exists a pairing $P = \{(v_1, w_1), \ldots, (v_{n/2}, w_{n/2})\}$ of the vertices of V such that $k_i := |P \cap E_i| \geq \lfloor \frac{m_i}{2(r_i+1)} \rfloor - 1$ for $i = 1, 2$. Let A be a random subset of V constructed as follows. For each pair (v_i, w_i) of P, we either choose $v_i \in A$, $w_i \notin A$ or $v_i \notin A$, $w_i \in A$, each with probability 1/2. The choices for each $i = 1, \ldots, n/2$ are made independently of one another. Let $X_i = e_{G_i}(A, A^c)$ and let X_i have mean μ_i and variance σ_i^2. By Corollary 2.3.2, it is sufficient to prove that

$$\mu_i - \sigma_i \geq \frac{1}{2} m_i,$$

for $i = 1, 2$. As before we compute μ_i and σ_i^2.

Let $G_i' = (V, E_i') = (V, E_i \backslash P)$ with $m_i' = |E_i'| = m_i - k_i$, and let $X_i' = e_{G_i'}(A, A^c)$. Thus we have $X_i' = X_i - k_i$, and so X_i' has mean $\mu_i' = \mu_i - k_i$ and variance $\sigma_i'^2 = \sigma_i^2$.

For $e \in E_i'$, we have that

$$\mathbb{P}(e \in E_{G_i'}(A, A^c)) = 1/2,$$

so as in the proof of Theorem 2.1.6, we have

$$\mathbb{E}(X_i') = \sum_{e \in E_i'} \mathbb{P}(e \in E_{G_i'}(A, A^c)) = \frac{m_i'}{2}.$$

Two edges e, f ($e \neq f$) in E_i' are said to be *linked* if there exists $p_1, p_2 \in P$ such that $e \cup f \subseteq p_1 \cup p_2$. For $e, f \in E_i'$, we have

$$\mathbb{P}(e, f \in E_{G_i}(A, A^c)) = \begin{cases} \frac{1}{2} & \text{if } e, f \text{ are linked and not incident;} \\ 0 & \text{if } e, f \text{ are linked and incident;} \\ \frac{1}{4} & \text{otherwise.} \end{cases}$$

To see why this is true, first assume that e and f are linked and that $p_1 = (u_1, v_1)$, $p_2 = (u_2, v_2)$. If e and f are not incident, then we must have (without loss of generality) that $e = v_1 w_2$, $f = v_2 w_1$; both edges belong to $E_{G_i'}(A, A^c)$ if and only if e does, giving the probability of $\frac{1}{2}$. If e and f are incident, then (without loss of generality) $e = u_1 u_2$, $f = u_1 v_2$; under no circumstances can both edges belong to $E_{G_i'}(A, A^c)$. Finally, if e and f are not linked, then the events $e \in E_{G_i'}(A, A^c)$ and $f \in E_{G_i'}(A, A^c)$ are independent, giving the probability of $\frac{1}{4}$.

For any edge $e \in E_i'$, there is at most one edge $f \in E_i'$ that is linked and not incident with e. Hence there are at most m_i' (ordered) pairs of edges of E_i' that are linked and not incident. As in the proof of Theorem 2.1.6, we have

$$\begin{aligned}\mathbb{E}(X_i'^2) &= \mathbb{E}(X_i') + \sum_{\substack{e,f \in E_i' \\ e \neq f}} \mathbb{P}(e, f \in E_{G_i'}(A, A^c)) \\ &\leq \frac{1}{2}m_i' + \frac{1}{4}[m_i'(m_i' - 1) - m_i'] + \frac{1}{2}m_i' \\ &= \frac{1}{4}m_i'^2 + \frac{1}{2}m_i',\end{aligned}$$

and

$$\begin{aligned}\sigma_i'^2 &= \mathbb{E}(X_i'^2) - \mathbb{E}(X_i')^2 \\ &\leq \frac{1}{2}m_i'.\end{aligned}$$

Therefore $\mu_i = \frac{1}{2}(m_i + k_i)$ and $\sigma_i^2 \leq \frac{1}{2}(m_i - k_i)$. We find that $\mu_i - \sigma_i \geq m_i/2$ if $\sigma_i^2 \leq \frac{1}{4}k_i^2$, i.e. if

$$m_i \leq \frac{1}{2}k_i^2 + k_i.$$

Given that $k_i \geq \lfloor \frac{m_i}{2(r_i+1)} \rfloor - 1$, it is easy to check that the above holds if $r_i \leq \sqrt{m_i/8} - 2$. To see this, observe that

$$\begin{aligned}k_i + 2 \geq \frac{m_i}{2(r_i + 1)} &\geq \frac{m_i}{2(\sqrt{m_i/8} - 1)} \\ &= \sqrt{2m_i} + 4 + \frac{8}{2(\sqrt{m_i/8} - 1)}.\end{aligned}$$

Thus $k_i \geq \sqrt{2m_i}$, so that $\frac{1}{2}k_i^2 + k_i \geq m_i$ as required. □

Note that the condition $\Delta(G_i) \leq \sqrt{m_i/8} - 2$ is only used at the end of the proof in order to bound k_i. More generally, any pair of graphs G_i, $i = 1, 2$, satisfying the condition that $m_i \leq \frac{1}{2}k_i^2 + k_i$ will satisfy Conjecture 2.1.5.

Chapter 3

Partitioning Posets

3.1 Introduction

In this chapter, we consider extremal and algorithmic analogues of the graph maxcut problem for posets. We consider the following question: how large can a k-cut of a poset comparability graph be if we restrict attention to those k-cuts arising from partitions that respect the order of the poset? The extremal problem has some similarities to and some differences from the graph case, while the algorithmic problem is quite different: there is a polynomial-time algorithm for finding the optimal ordered partition, contrasting the maxcut problem for graphs, which, as we mentioned in Section 1.1, is NP-hard.

All posets in this chapter are understood to be finite. Recall that an ordered k-partition of a poset $P = (X, \prec)$ is a partition $X_1, \ldots, X_k$ of X for which, whenever $a \in X_i$ and $b \in X_j$ with $a \prec b$, we have $i \leq j$. Recall, also, that the comparability graph of a poset $P = (X, \prec)$ is denoted by $\mathrm{Com}(P)$. For the purposes of this chapter, we simplify notation in the following way. When considering the edges of $\mathrm{Com}(P)$ within a set $A \subseteq X$ or across a partition $A_1, \ldots, A_k$ of X, rather than writing $E_{\mathrm{Com}(P)}(A)$, $e_{\mathrm{Com}(P)}(A)$, $E_{\mathrm{Com}(P)}(A_1, \ldots, A_k)$, and $e_{\mathrm{Com}(P)}(A_1, \ldots, A_k)$, we write instead $E_P(A)$, $e_P(A)$, $E_P(A_1, \ldots, A_k)$, and $e_P(A_1, \ldots, A_k)$ respectively.

In light of the extremal graph cut results in Section 1.1, there are some natural questions we can ask about ordered partitions: given a poset $P = (X, \prec)$, can we find an ordered 2-partition X_1, X_2 of P for which $e_P(X_1, X_2)$ is large (or equivalently for which $e_P(X_1) + e_P(X_2)$ is small)? Is there a corresponding version of Proposition 1.1.1 or the result of Edwards, Theorem 1.1.2, for ordered 2-partitions of posets? We can generalise the problem by asking if there exists an ordered k-partition $X_1, \dots, X_k$ of P for which $e_P(X_1, \dots, X_k)$ is large (or equivalently for which $e_P(X_1)+\cdots+e_p(X_k)$ is small). Again, we might seek results corresponding to Proposition 1.1.3 or the generalisation of Edwards' result, Theorem 1.1.4. It turns out to be useful to take the generalisation one step further, by considering a weighted version of the problem. Let us fix some notation.

Given a poset $P = (X, \prec)$, a positive integer k, and positive real numbers $a_1, \dots, a_k$, define

$$f(P; a_1, \dots, a_k) = \min \Big(\sum_{i=1}^{k} a_i e_P(X_i) \Big),$$

where the minimum is taken over all ordered k-partitions $X_1, \dots, X_k$ of P. Define

$$f(m; a_1, \dots, a_k) = \max(f(P; a_1, \dots, a_k)),$$

where the maximum is taken over all posets $P = (X, \prec)$ for which $e_P(X) = m$.

Of course, the case where $a_1 = \cdots = a_k = 1$ is the one we are most interested in, but we study the more general weighted case because it is crucial to our proof techniques.

In Sections 3.2 and 3.3, we consider the problem of bounding $f(m)$. In Section 3.2, we prove the following theorem.

Theorem 3.1.1 *Let k be a positive integer. For positive real numbers $a_1, \dots, a_k$ and a positive integer m, we have that*

$$f(m; a_1, \dots, a_k) \leq \Big(\sum_{i=1}^{k} a_i^{-1} \Big)^{-1} m.$$

Let us compare a few special cases of this result with analogous ones for graphs. Consider the case when $a_1 = \cdots = a_k = 1$. Then $(\sum_{i=1}^{k} a_i^{-1})^{-1} = \frac{1}{k}$, and Theorem 3.1.1 tells us that every poset $P = (X, \prec)$ has an ordered partition $X_1, \dots, X_k$ such that

$$e_P(X_1) + \cdots + e_P(X_k) \leq \frac{1}{k} e_P(X),$$

or equivalently that $e_P(X_1, \dots, X_k) \geq (1 - \frac{1}{k})e_P(X)$. Thus, this is the analogue of Proposition 1.1.1 for posets. Note, however, that this result cannot be deduced simply by applying Proposition 1.1.1 to Com(P) since the partition that Proposition 1.1.1 gives us is not necessarily ordered.

Theorem 3.1.1 is in fact the analogue for posets of a generalisation of Proposition 1.1.1, which we give below.

Theorem 3.1.2 *Given positive real numbers $a_1, \dots, a_k$ and a graph $G = (V, E)$, there exists a partition of V into parts $V_1, \dots, V_k$ such that*

$$\sum_{i=1}^{k} a_i e_G(V_i) \leq \Big(\sum_{i=1}^{k} a_i^{-1}\Big)^{-1} |E|.$$

The proof of Theorem 3.1.2 is simply an extension of the probabilistic proof of Proposition 1.1.1.

Proof We partition V into sets $V_1, \dots, V_k$ by assigning each vertex independently at random to one of $V_1, \dots, V_k$, where

$$\mathbb{P}(v \in V_j) = \frac{a_j^{-1}}{\sum_{i=1}^{k} a_i^{-1}}.$$

These probabilities are chosen optimally for the argument that follows. Given an edge $e = ab \in E$, we have that $\mathbb{P}(e \in E_G(V_j)) = a_j^{-2}/(\sum_{i=1}^{k} a_i^{-1})^2$, and so

$$\begin{aligned}\mathbb{E}(\sum_{j=1}^{k} a_j e_G(V_j)) &= \sum_{j=1}^{k} a_j \sum_{e \in E} \mathbb{P}(e \in E_G(V_j)) = \sum_{j=1}^{k} a_j \sum_{e \in E} \frac{a_j^{-2}}{(\sum_{i=1}^{k} a_i^{-1})^2} \\ &= \sum_{e \in E} \frac{1}{\sum_{i=1}^{k} a_i^{-1}} = \Big(\sum_{i=1}^{k} a_i^{-1}\Big)^{-1} |E|.\end{aligned}$$

There must be some partition $V_1', \dots, V_k'$ of V for which $\sum_{i=1}^{k} a_i e_G(V_i)$ is at most its expected value, proving the theorem. □

Although Theorem 3.1.2 is a natural bound for graphs, it can be sharpened by an extra term of order $\Theta(\sqrt{m})$ in the same way that Theorem 1.1.4 sharpens Proposition 1.1.3. Thus, we might expect a similar improvement on the bound given in Theorem 3.1.1. There is a surprising dichotomy here: when $k \geq 3$, we can improve the bound in Theorem 3.1.1 by $\Theta(\sqrt{m})$, but when $k = 2$, Theorem 3.1.1 is best possible.

For the case $k = 2$, we show in Section 3.3 that for every positive integer m, we have

$$f(m;1,1) = \left\lfloor \frac{m-1}{2} \right\rfloor.$$

We show further that for m a fixed odd positive integer, there is a unique poset $P_r = (X_r, \prec)$ (Figure 3.1 below) for which $f(P_r;1,1) = f(m;1,1)$ (here $r = (m-1)/2$). Describing P_r in words, we have the ground set $X_r = \{y_1, y_2, x_1, \ldots, x_r\}$, with $\{x_1, \ldots, x_r\}$ an antichain, and $y_1 \prec x_i \prec y_2$ for $i = 1, \ldots, r$ (and the transitive relation $y_1 \prec y_2$).

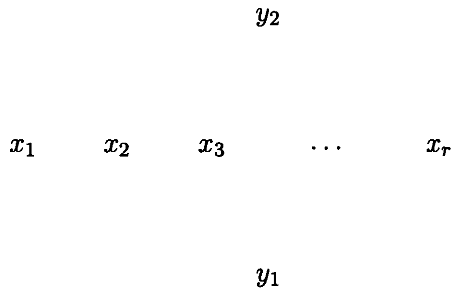

Figure 3.1: Hasse diagram of P_r

We can see immediately that to minimise $e_{P_r}(X_1) + e_{P_r}(X_2)$ (over all ordered partitions X_1, X_2 of P), we must have $y_1 \in X_1$ and $y_2 \in X_2$. Then, no matter how we place each x_i, exactly one of the edges y_1x_i and x_iy_2 lies inside a part. This gives us that $f(P_r;1,1) = r = (m-1)/2$, where m is the number of comparable pairs in P_r. (If we drop the condition that X_1, X_2 should be ordered, then the optimal partition of $\mathrm{Com}(P_r)$ is one with only a single edge inside parts.)

Thus we see that when $k = 2$ with $a_1 = a_2 = 1$, we cannot improve the bound given in Theorem 3.1.1 by more than a constant (independent of m), and we show in Section 3.3.1 that this is the case for general rational a_1, a_2, using examples similar to P_r. However, for $k \geq 3$, we find that we can improve the bound in Theorem 3.1.1 by at least $c\sqrt{m}$, where c is a constant independent of m (but dependent on k). These results are summarised in the next theorem.

Theorem 3.1.3 *We have the following.*

(a) For fixed positive real numbers a_1 and a_2 with a_2/a_1 rational, we have

$$f(m;a_1,a_2) = \left(\frac{1}{a_1} + \frac{1}{a_2}\right)^{-1} m - \Theta(1).$$

(b) *For a fixed integer $k \geq 3$ and fixed positive real numbers $a_1, \ldots, a_k$, we have*

$$f(m; a_1, \ldots, a_k) = \Big(\sum_{i=1}^{k} a_i^{-1}\Big)^{-1} m - \Theta(\sqrt{m}).$$

Note that we do not allow arbitrary real values for a_1 and a_2 in the statement of Theorem 3.1.3(a). In Section 3.3.2, we show, by giving an explicit example, that Theorem 3.1.3(a) does not hold in general for real values of a_1 and a_2. We show that when $a_1 = 1$ and $a_2 = (1 + \sqrt{5})/2$, we have

$$f(m; a_1, a_2) = \Big(\frac{1}{a_1} + \frac{1}{a_2}\Big)^{-1} m - \Omega(\log m).$$

For general real values of a_1 and a_2, we know only that

$$f(m; a_1, a_2) = \Big(\frac{1}{a_1} + \frac{1}{a_2}\Big)^{-1} m - O(\sqrt{m}).$$

The example of the chain, C_n, on n elements gives the bound above. It also shows that the error term in Theorem 3.1.3(b) is of the correct order of magnitude. For $\binom{n}{2} \leq m \leq \binom{n+1}{2}$, an easy calculation given at the end of Section 3.3.3 shows that

$$f(m; a_1, \ldots, a_k) \geq f(C_n; a_1, \ldots, a_k) = \Big(\sum_{i=1}^{k} a_i^{-1}\Big)^{-1} m - \Theta(\sqrt{m}),$$

where $k \geq 2$ and $a_1, \ldots, a_k$ are real numbers.

In Section 3.4, we consider the algorithmic analogue of the graph maxcut problem for posets. Starting with the case $k = 2$, we give a very simple polynomial-time algorithm that, given a poset P, minimises $e_P(X_1)+e_P(X_2)$ (equivalently maximises $e_P(X_1, X_2)$) over all ordered 2-partitions X_1, X_2 of P. Moreover the algorithm finds all ordered 2-partitions that achieve the minimum. This is in contrast to the graph maxcut problem, which is NP-hard. We go on to show that all the various generalisations of this problem are also polynomial-time solvable, but the algorithms we present for some of these generalisations are more complicated.

As we discussed in the previous paragraph, we have a very simple algorithm for determining $f(P; 1, 1)$ and for finding all ordered 2-partitions that achieve the minimum. This algorithm generalises easily to one that determines $f(P; a_1, a_2)$ (for given positive rationals a_1, a_2) and finds all the ordered 2-partitions that achieve the minimum. For $k \geq 3$, the algorithms we present are not so simple.

Given a positive integer k and a poset P, we show that the problem of determining the value of $f(P; 1, \ldots, 1)$ and finding an ordered k-partition that achieves this minimum can be reduced to the problem of finding a largest union of k antichains in a poset $L(P)$ derived from P. The poset $L(P)$ is easily constructed from P and has size that is polynomial in the size of P. Finding a largest union of k antichains in a poset is polynomial-time solvable (Corollary 1.2.6). Thus there exists a polynomial-time algorithm for determining $f(P; 1, \ldots, 1)$ and for finding an ordered k-partition that achieves this minimum.

Finally, the general problem of determining $f(P; a_1, \ldots, a_k)$ (for given positive rationals $a_1, \ldots, a_k$) and finding a corresponding ordered k-partition is shown to be polynomial-time solvable as a relatively straightforward consequence of the fact that a submodular function on a distributive lattice can be minimised in polynomial time. However, since all known algorithms for minimising submodular functions are rather intricate, the method we give for determining $f(P; a_1, \ldots, a_k)$ is by no means a simple one.

3.2 Good Partitions

In this section, we prove Theorem 3.1.1. The key step is to prove the result for the case $k = 2$; the full result then follows by a straightforward induction argument. We begin with some notation.

Let $P = (X, \prec)$ be a poset. For $A \subseteq X$, recall that $\max_P(A)$ (resp. $\min_P(A)$) denotes the set of maximal (resp. minimal) elements of the poset induced by P on A.

For $x \in X$, let

$$U_P(x) = \{y \in X : y \succ x\} \quad \text{with} \quad u_P(x) = |U_P(x)|,$$
$$\text{and } D_P(x) = \{y \in X : y \prec x\} \quad \text{with} \quad d_P(x) = |D_P(x)|.$$

Given positive real numbers a_1 and a_2, define the function $h^P_{a_1,a_2} : X \to \mathbb{R}$ by

$$h^P_{a_1,a_2}(x) = a_2 u_P(x) - a_1 d_P(x).$$

We drop subscripts and/or superscripts when it is clear what these are.

Observe that $h^P_{a_1,a_2}$ is a strictly decreasing function, that is, for $x, y \in X$ with $x \prec y$, we have $h^P_{a_1,a_2}(y) < h^P_{a_1,a_2}(x)$ (since $u_P(x) > u_P(y)$ and $d_P(x) < d_P(y)$).

A partition of X into parts X_1 and X_2 is called an (a_1, a_2)-*good* partition of P if $h^P_{a_1,a_2}(x) \geq 0$ for all $x \in X_1$ and $h^P_{a_1,a_2}(x) \leq 0$ for all $x \in X_2$ (thus an (a_1, a_2)-good partition of P is uniquely defined except that any element x satisfying $h(x) = 0$ can be in either X_1 or X_2). It is clear that every (a_1, a_2)-good partition of P is an ordered partition of P (if $a \in X_1$ and $b \in X_2$, then $h^P_{a_1.a_2}(a) \geq 0 \geq h^P_{a_1.a_2}(b)$ and so either $a \prec b$ or $a \parallel b$).

We have the following lemma, which is the case $k = 2$ of Theorem 3.1.1.

Lemma 3.2.1 *Fix positive real numbers a_1 and a_2. For $P = (X, \prec)$ a poset, let X_1, X_2 be an (a_1, a_2)-good partition of P. Then*

$$a_1 e(X_1) + a_2 e(X_2) \leq \frac{a_1 a_2}{a_1 + a_2} e(X) = \left(\frac{1}{a_1} + \frac{1}{a_2}\right)^{-1} e(X).$$

Proof The proof is by induction on $|X|$. The lemma is trivially true when $|X| = 1$.

Define $r : X \to \mathbb{R}$ by

$$r(x) = \begin{cases} a_1 d_P(x) & \text{if} \quad x \in X_1; \\ a_2 u_P(x) & \text{if} \quad x \in X_2. \end{cases}$$

Choose x^* to be any element of X that maximises r. We see that $x^* \in B$, where $B = \max_P(X_1) \cup \min_P(X_2)$. We assume that $x^* \in \max_P(X_1)$; the case $x^* \in \min_P(X_2)$ can be deduced by applying the argument that follows to the dual poset with the roles of a_1 and a_2 reversed.

Let $X' = X \backslash \{x^*\}$, $P' = (X', \prec)$, $X'_1 = X_1 \backslash \{x^*\}$, and $X'_2 = X_2 \backslash \{x^*\}$. We claim that X'_1, X'_2 is an (a_1, a_2)-good partition of P'. Let us assume that the claim is true and continue with the proof. We have

$$\begin{aligned} e_P(X_1) &= e_{P'}(X'_1) + d_P(x^*), \\ e_P(X_2) &= e_{P'}(X'_2), \\ \text{and} \quad e_P(X) &= e_{P'}(X') + u_P(x^*) + d_P(x^*). \end{aligned}$$

Putting this together, we have

$$\begin{aligned}
\frac{a_1a_2}{a_1+a_2}e_P(X) - a_1e_P(X_1) - a_2e_P(X_2) &= \frac{a_1a_2}{a_1+a_2}e_{P'}(X') - a_1e_{P'}(X_1') \\
&\quad - a_2e_{P'}(X_2') + u_P(x^*)\frac{a_1a_2}{a_1+a_2} + d_P(x^*)\Big(\frac{a_1a_2}{a_1+a_2} - a_1\Big) \\
&= \frac{a_1a_2}{a_1+a_2}e_{P'}(X') - a_1e_{P'}(X_1') - a_2e_{P'}(X_2') + \frac{a_1}{a_1+a_2}h^P_{a_1,a_2}(x^*) \qquad (3.1) \\
&\geq 0,
\end{aligned}$$

where the last inequality follows by induction and the fact that $h^P_{a_1,a_2}(x^*) \geq 0$ (since $x^* \in X_1$).

It remains only to show that X_1', X_2' is an (a_1, a_2)-good partition of P', that is, we must show that

$$h^{P'}(x) \begin{cases} \geq 0 & \forall x \in X_1'; \\ \leq 0 & \forall x \in X_2'. \end{cases}$$

Observe that $h^{P'}(x) = h^P(x)$ if x and x^* are incomparable in P and so the above holds for such elements x.

If $x \prec x^*$, then $x \in X_1$ and

$$\begin{aligned}
h^{P'}(x) &= a_2u_{P'}(x) - a_1d_{P'}(x) \\
&\geq a_2u_P(x^*) - a_1(d_P(x^*) - 1) \\
&= h^P(x^*) + a_1 \\
&\geq 0.
\end{aligned}$$

If $x \succ x^*$ then $x \in X_2$ and

$$\begin{aligned}
h^{P'}(x) &= a_2u_{P'}(x) - a_1d_{P'}(x) \\
&\leq a_2u_P(x) - a_1d_P(x^*) \\
&\leq 0 \quad \text{(by our choice of } x^*\text{)}.
\end{aligned}$$

This completes the proof. □

We make some remarks, which we shall make use of later.

Remark Lemma 3.2.1 says that $f(P; a_1, a_2) \leq (a_1^{-1} + a_2^{-1})^{-1}e_P(X)$ for all posets P. Examining the proof of Lemma 3.2.1, we see that we make a gain on this bound every time we remove a vertex x^* (in the induction) for which $|h(x^*)| > 0$. Hence, one way to construct a poset P for which $f(P; a_1, a_2)$ is close to our bound is to include many vertices x for which $h(x) = 0$. In fact, this is necessary in light of Lemma 3.3.2, which we prove in the next section.

Remark We note also that we have strict inequality in Lemma 3.2.1 if $e_P(X) \geq 1$. This is because, as we inductively remove vertices from our poset P, we will eventually be left with a poset of height 2. For such a poset, Lemma 3.2.1 holds with strict inequality. Then, working backwards through the induction, we find that Lemma 3.2.1 holds with strict inequality for P.

We now prove Theorem 3.1.1 by an easy induction argument, using Lemma 3.2.1 as the induction step.

Proof (of Theorem 3.1.1) Given $a_1, \ldots, a_k$, it is sufficient to show that for every poset $P = (X, \prec)$, there exists an ordered partition of P into sets $X_1, \ldots, X_k$ such that

$$\sum_{i=1}^{k} a_i e(X_i) \leq \Big(\sum_{i=1}^{k} a_i^{-1}\Big)^{-1} e(X).$$

We use induction on k. The above is trivially true for $k = 1$. Assume it is true for $k - 1$.

Let $b_1 = (\sum_{i=1}^{k-1} a_i^{-1})^{-1}$ and $b_2 = a_k$. By Lemma 3.2.1, there exists an ordered partition of P into parts Y_1 and Y_2 such that

$$b_1 e(Y_1) + b_2 e(Y_2) \leq \Big(\frac{1}{b_1} + \frac{1}{b_2}\Big)^{-1} e(X) = \Big(\sum_{i=1}^{k} a_i^{-1}\Big)^{-1} e(X).$$

By the induction hypothesis, there exists an ordered partition of Y_1 into parts $X_1, \ldots, X_{k-1}$ such that

$$\sum_{i=1}^{k-1} a_i e(X_i) \leq \Big(\sum_{i=1}^{k-1} a_i^{-1}\Big)^{-1} e(Y_1) = b_1 e(Y_1).$$

Setting $X_k = Y_2$ gives the desired ordered partition of P. $\square$

3.3 Better Partitions

3.3.1 Rational Weights in Bipartitions

Our first task is to prove Theorem 3.1.3(a), which says that for the case $k = 2$, Theorem 3.1.1 is close to best possible. We do this, as the remark after Lemma 3.2.1 suggests, by constructing posets that include a large number of vertices x for which $h^P_{a_1,a_2}(x) = 0$.

Proof (of Theorem 3.1.3(a)) In light of Lemma 3.2.1, it is sufficient to prove the lower bound

$$f(m; a_1, a_2) \geq \frac{a_1 a_2}{a_1 + a_2} m - \Theta(1).$$

For the moment, we assume that a_1 and a_2 are positive integers. For an integer $t \geq 0$, let $P(t) = P(a_1, a_2, t)$ be the complete three-layer poset with a_1 elements in the top layer A_1, a_2 elements in the bottom layer A_2, and t elements in the middle layer T (so, A_1, T, and A_2 are antichains and every element in T is below every element in A_1 and above every element in A_2).

We let X'_1, X'_2 be an ordered partition that minimises $a_1 e_{P(t)}(X_1) + a_2 e_{P(t)}(X_2)$ over all ordered partitions X_1, X_2 of $P(t)$. It is easy to see that $A_2 \subseteq X'_1$ and $A_1 \subseteq X'_2$. (Indeed, if $A_1 \not\subseteq X'_2$ then there exists some $x \in A_1$ such that $x \in X'_1$; since X'_1, X'_2 is an ordered partition, this implies $T \cup A_2 \subseteq X'_1$. Moving x from X'_1 to X'_2 reduces $a_1 e_{P(t)}(X_1) + a_2 e_{P(t)}(X_2)$ contradicting that X'_1, X'_2 minimises the same. Thus $A_1 \subseteq X'_2$ and by symmetry, $A_2 \subseteq X'_1$.)

Assuming X'_1 contains t_1 elements from T and X'_2 contains the other $t_2 := t - t_1$ elements from T, and noting that $P(t)$ has $m(t) = (a_1 + a_2)t + a_1 a_2$ comparable pairs, we have that

$$\begin{aligned}
\min(a_1 e_{P(t)}(X_1) + a_2 e_{P(t)}(X_2)) &= a_1 e_{P(t)}(X'_1) + a_2 e_{P(t)}(X'_2) \\
&= a_1(a_2 t_1) + a_2(a_1(t - t_1)) \\
&= a_1 a_2 t \\
&= \frac{a_1 a_2}{a_1 + a_2}(m(t) - a_1 a_2) \\
&= \frac{a_1 a_2}{a_1 + a_2} m(t) - d(a_1, a_2)
\end{aligned}$$

for all t, where $d(a_1, a_2)$ is a constant independent of $m(t)$. Given m, we choose t so that $m(t) \leq m < m(t + 1)$, and thus $m - m(t) < a_1 + a_2$. Now we have

$$\begin{aligned}f(m;a_1,a_2) \geq f(m(t);a_1,a_2) \geq f(P(t);a_1,a_2) \\ &= \frac{a_1a_2}{a_1+a_2}m(t) - d(a_1,a_2) \\ &= \frac{a_1a_2}{a_1+a_2}m - \Theta(1).\end{aligned}$$

The above is also true for rational values of a_1 and a_2, and more generally when a_1/a_2 is rational, since

$$f(m;\frac{a_1}{r},\frac{a_2}{r}) = \frac{1}{r}f(m;a_1,a_2),$$

for any real $r > 0$. □

3.3.2 Irrational Weights in Bipartitions

Next, we give an example to show that Theorem 3.1.3(a) does not hold in general for real values of a_1 and a_2. We make use of some elementary results in the theory of continued fractions and Diophantine approximation, all of which can be found in, for example, [31].

Theorem 3.3.1 *For $a_1 = 1$ and $a_2 = (1+\sqrt{5})/2$, we have that*

$$f(m;a_1,a_2) = \left(\frac{1}{a_1}+\frac{1}{a_2}\right)^{-1} m - \Omega(\log m).$$

Proof We start with some preliminaries. We make use of the result that $\phi = (1+\sqrt{5})/2$, the *golden ratio*, has best rational approximation given by ratios of successive *Fibonacci numbers*. Let us go into more detail. The Fibonacci sequence F_n is defined by the recursive relation $F_n = F_{n-1}+F_{n-2}$ with the initial conditions that $F_0 = 0$ and $F_1 = 1$. We have

$$F_n = \frac{1}{\sqrt{5}}(\phi^n - \hat{\phi}^n),$$

where $\hat{\phi} = (1-\sqrt{5})/2$. Note that $\phi+\hat{\phi} = 1$, $\phi-\hat{\phi} = \sqrt{5}$, $\phi\hat{\phi} = -1$, and $|\hat{\phi}| < 1$; we shall use these in later calculations.

A consequence of what is sometimes referred to as the *law of best approximation* (Theorem 182 in [31]) is the following. For any natural numbers

r, s where $s \leq F_n$, we have that

$$|s\phi - r| \geq |F_n\phi - F_{n+1}|.$$

We prove that for every poset $P = (X, \prec)$, where $|X| = n$, there exists an ordered partition X_1, X_2 such that

$$a_1 e_P(X_1) + a_2 e_P(X_2) = (a_1^{-1} + a_2^{-1})^{-1} e_P(X) - \Omega(\log n).$$

This then proves the theorem since $\log n > \log \sqrt{m} = \frac{1}{2}\log m$.

Note that every poset $P = (X, \prec)$ without isolated elements has a unique (a_1, a_2)-good partition for our choice of a_1 and a_2; this is because $h^P_{a_1,a_2}(x)$ is an integral linear function of ϕ, so it is non-zero for all $x \in X$.

Now fix $P = (X, \prec)$ with $|X| = n$. Define the sequence of posets $P_i = (X_i, \prec)$, $i = 0, \dots, n-1$ as follows. Let $P_0 = P$. Given $P_i = (X_i, \prec)$, let X_1^i, X_2^i be the (a_1, a_2)-good partition of P_i. We know from the proof of Lemma 3.2.1 that there exists some $x_i^* \in X_i$ such that, defining $X_{i+1} = X_i \backslash \{x_i^*\}$, $P_{i+1} = (X_{i+1}, \prec)$, $X_1^{i+1} = X_1^i \backslash \{x_i^*\}$, and $X_2^{i+1} = X_2^i \backslash \{x_i^*\}$, we have that X_1^{i+1}, X_2^{i+1} is the (a_1, a_2)-good partition of P_{i+1}. Furthermore, using (3.1) from the proof of Lemma 3.2.1, we have

$$\begin{aligned}&(a_1^{-1} + a_2^{-1})^{-1} e_{P_i}(X_i) - a_1 e_{P_i}(X_1^i) - a_2 e_{P_i}(X_2^i)\\ &= (a_1^{-1} + a_2^{-1})^{-1} e_{P_{i+1}}(X_{i+1}) - a_1 e_{P_{i+1}}(X_1^{i+1}) - a_2 e_{P_{i+1}}(X_2^{i+1}) + c|h^{P_i}(x_i^*)|,\end{aligned}$$

where c is either $a_1^{-1}(a_1^{-1}+a_2^{-1})^{-1}$ or $a_2^{-1}(a_1^{-1}+a_2^{-1})^{-1}$ depending on whether x_i^* is in X_1^i or X_2^i. Thus (since P_{n-1} has a single element and no relations) the above gives us that

$$(a_1^{-1} + a_2^{-1})^{-1} e_P(X) - a_1 e_P(X_1) - a_2 e_P(X_2) = \Theta\Big(\sum_{i=0}^{n-1} |h^{P_i}(x_i^*)|\Big),$$

and it remains for us to show that $\sum_{i=0}^{n-1} |h^{P_i}(x_i^*)| = \Omega(\log n)$.

Note that P_{n-i} has i elements. Let $k(i)$ be the smallest integer so that $F_{k(i)} \geq i$. Observe that $k(i) = \Theta(\log i)$. Then, using the law of best approximation, we have that

$$\begin{aligned}|h^{P_{n-i}}(x_{n-i}^*)| &= |\phi u_{P_{n-i}}(x_{n-i}^*) - d_{P_{n-i}}(x_{n-i}^*)| \geq |F_{k(i)}\phi - F_{k(i)+1}| \\ &= \frac{1}{\sqrt{5}}\left|(\phi^{k(i)} - \hat{\phi}^{k(i)})\phi - (\phi^{k(i)+1} - \hat{\phi}^{k(i)+1})\right| = |\hat{\phi}^{k(i)}|.\end{aligned}$$

Now, using the above and the fact that $k(i) = j$ for $F_j - F_{j-1}$ values of i, we have that

$$\sum_{i=0}^{n-1} |h^{P_i}(x_i^*)| = \sum_{i=1}^{n} |h^{P_{n-i}}(x_{n-i}^*)| \geq \sum_{i=1}^{n} |\hat{\phi}^{k(i)}| \geq \sum_{j=1}^{k(n)-1} |\hat{\phi}^j|(F_j - F_{j-1}).$$

Also

$$\begin{aligned}|\hat{\phi}^j|(F_j - F_{j-1}) = |\hat{\phi}^j|F_{j-2} &= \frac{1}{\sqrt{5}}|\hat{\phi}^j(\phi^{j-2} - \hat{\phi}^{j-2})| \\ &= \frac{1}{\sqrt{5}}|\hat{\phi}^2(-1)^{j-2} - \hat{\phi}^{2j-2}| \geq \frac{1}{\sqrt{5}}(|\hat{\phi}^2| - |\hat{\phi}^{2j-2}|).\end{aligned}$$

Finally, we have that

$$\begin{aligned}\sum_{i=0}^{n-1} |h^{P_i}(x_i^*)| &\geq \frac{1}{\sqrt{5}}\left(\sum_{j=1}^{k(n)-1} |\hat{\phi}^2| - \sum_{j=1}^{k(n)-1} |\hat{\phi}^{2j-2}|\right) \\ &= \Theta(k(n)) = \Theta(\log n),\end{aligned}$$

as required. □

It is unclear if the bound in Theorem 3.3.1 gives the correct asymptotic value for $f(m; 1, (1+\sqrt{5})/2)$. More generally, it seems that the growth of $(a_1^{-1} + a_2^{-1})^{-1}m - f(m; a_1, a_2)$ depends on how well we can approximate a_2/a_1 by rationals.

3.3.3 Weighted k-partitions

Our next lemma will be the key step in proving the upper bound for Theorem 3.1.3(b). It will also enable us to prove the uniqueness of certain extremal posets in the next section. First we introduce the notion of a *balanced* element.

Fix positive real numbers a_1 and a_2, and let $P = (X, \prec)$ be a poset. For

t a positive real number, define

$$\mathrm{Bal}^P_{a_1,a_2}(t) = \{x \in X : |h^P_{a_1,a_2}(x)| \le t\},$$

and let $\mathrm{bal}^P_{a_1,a_2}(t) = |\mathrm{Bal}^P_{a_1,a_2}(t)|$. We refer to elements in $\mathrm{Bal}^P_{a_1,a_2}(t)$ as *balanced* elements. Once again, subscripts and superscripts may be dropped.

Here is the aforementioned lemma, which gives us an upper bound on $f(P; a_1, a_2)$ that takes into account the number of balanced elements in P.

Lemma 3.3.2 *Fix positive real numbers a_1 and a_2, and let $P = (X, \prec)$ be a poset with X_1, X_2 an (a_1, a_2)-good partition of P. For $0 \le t < \frac{1}{2}\min(a_1, a_2)$, we have*

$$a_1 e_P(X_1) + a_2 e_P(X_2) \le \frac{a_1 a_2}{a_1 + a_2} e_P(X) - t\frac{\min(a_1, a_2)}{2(a_1 + a_2)}(|X| - \mathrm{bal}^P_{a_1,a_2}(t)).$$

Furthermore, if $a_1 = a_2$, then the above inequality holds under the weaker condition $0 \le t < \min(a_1, a_2)$.

We concentrate on the proof of the first part of the lemma. The second part has almost the same proof; we make remarks where the proofs differ.

Proof The proof is again by induction on $|X|$. The lemma is true for $|X| = 1$ since an isolated element is balanced. Assume it is true for all posets with fewer than $|X|$ elements.

Let x^* be as in the proof of Lemma 3.2.1, and as before, let $X' = X \backslash \{x^*\}$, $P' = (X', \prec)$, $X_1' = X_1 \backslash \{x^*\}$, and $X_2' = X_2 \backslash \{x^*\}$. We know from the proof of Lemma 3.2.1 that X_1', X_2' is an (a_1, a_2)-good partition of P'. We shall assume that $x^* \in \max_P(X_1)$ so that $h^P(x^*) \ge 0$; the same argument holds if $x^* \in \min_P(X_2)$ by considering the dual poset with the roles of a_1 and a_2 reversed. We must consider the two cases $|h^P(x^*)| \le t$ and $|h^P(x^*)| > t$ separately.

Suppose $h^P(x^*) \le t$. (Regarding the case $a_1 = a_2$, if $h^P(x^*) \le t < \min(a_1, a_2)$, then $h^P(x^*) = 0$.) We claim that $\mathrm{bal}^{P'}(t) \le \mathrm{bal}^P(t) - 1$. We prove this by showing that the removal of x^* from P does not create any new balanced elements in P'. Fix $x \in X$ with $x \notin \mathrm{Bal}^P(t)$.

If x is incomparable to x^*, then $h^{P'}(x) = h^P(x)$; hence $|h^{P'}(x)| > t$ and $x \notin \mathrm{Bal}^{P'}(t)$.

If $x \prec x^*$ then

$$\begin{aligned} h^{P'}(x) &= a_2 u_{P'}(x) - a_1 d_{P'}(x) \\ &\geq a_2 u_P(x^*) - a_1(d_P(x^*) - 1) \\ &= h^P(x^*) + a_1 > t, \end{aligned}$$

so $x \notin \mathrm{Bal}^{P'}(t)$.

If $x \succ x^*$ then

$$\begin{aligned} h^{P'}(x) &= a_2 u_{P'}(x) - a_1 d_{P'}(x) \\ &\leq a_2(u_P(x^*) - 1) - a_1(d_P(x^*)) \\ &= h^P(x^*) - a_2 < -t, \end{aligned}$$

so $x \notin \mathrm{Bal}^{P'}(t)$.

We have proved our claim and we have, as in Lemma 3.2.1, that

$$\begin{aligned} &\frac{a_1 a_2}{a_1 + a_2} e_P(X) - a_1 e_P(X_1) - a_2 e_P(X_2) \\ &\quad = \frac{a_1 a_2}{a_1 + a_2} e_{P'}(X') - a_1 e_{P'}(X_1') - a_2 e_{P'}(X_2') + \frac{a_1}{a_1 + a_2} h^P(x^*) \\ &\quad \geq t \frac{\min(a_1, a_2)}{2(a_1 + a_2)} (|X'| - \mathrm{bal}^{P'}(t)) \\ &\quad \geq t \frac{\min(a_1, a_2)}{2(a_1 + a_2)} (|X| - \mathrm{bal}^{P}(t)), \end{aligned}$$

where the first inequality follows from the induction hypothesis and the fact that $h^P(x^*) \geq 0$, and second inequality follows because $|X'| = |X| - 1$ and $\mathrm{bal}^{P'}(t) \leq \mathrm{bal}^P(t) - 1$.

Next, suppose $h^P(x^*) > t$. We claim that

$$h^P(x^*) \geq \begin{cases} t(\mathrm{bal}^{P'}(t) - \mathrm{bal}^P(t)) & \text{if } \mathrm{bal}^{P'}(t) > \mathrm{bal}^P(t); \\ t & \text{if } \mathrm{bal}^{P'}(t) \leq \mathrm{bal}^P(t), \end{cases}$$

the second case being trivial. Again, we consider which elements in X change from being unbalanced to balanced when x^* is removed from X. Fix $x \in X'$ with $x \notin \mathrm{Bal}^P(t)$.

If x and x^* are incomparable, then as before $h^P(x) = h^{P'}(x)$, and $x \notin \mathrm{Bal}^{P'}(t)$.

If $x \prec x^*$, then as before

$$h^{P'}(x) \geq h^P(x^*) + a_1 > t,$$

and $x \notin \mathrm{Bal}^{P'}(t)$.

If $x \succ x^*$, then x may become balanced upon removal of x^*, so we must consider all such elements. Let $x_1, \ldots, x_r$ be the elements of X such that for each i, $x_i \succ x^*$ and x_i is balanced in P' but not in P. We first show that $x_1, \ldots, x_r$ is an antichain in P'. If $x_j \succ x_i$ for some $1 \leq i, j \leq r$, then we have $h^{P'}(x_i) \leq 0$ (since $x_i \in X_2'$ and X_1', X_2' is an (a_1, a_2)-good partition of P'). We also have $h^{P'}(x_j) \leq h^{P'}(x_i) - (a_1 + a_2)$, giving that $h^{P'}(x_j) < -t$ (by our choice of t), contradicting that x_j is balanced in P'. Thus, $x_1, \ldots, x_r$ must be an antichain.

Now, we have

$$\begin{aligned}
h^P(x^*) &= a_2 u_P(x^*) - a_1 d_P(x^*) \\
&= a_2\left(r + \left|\bigcup_{i=1}^{r} U_{P'}(x_i)\right|\right) - a_1(d_{P'}(x_r)) \\
&\geq a_2 r + h^{P'}(x_r) \\
&\geq tr,
\end{aligned}$$

where the last inequality follows since $h^{P'}(x_r) > -t$ and $a_2 > 2t$ (by our choice of t). (Regarding the case $a_1 = a_2$, $h^{P'}(x_r) = 0$ and $a_2 > t$ (by our choice of t).) This proves the claim since $r = \mathrm{bal}^{P'}(t) - \mathrm{bal}^P(t)$. Combining the two cases of the claim, we have the following (weaker) inequality,

$$h^P(x^*) \geq \frac{1}{2}t\big(1 + \mathrm{bal}^{P'}(t) - \mathrm{bal}^P(t)\big).$$

Finally, to complete the induction, we have

$$\begin{aligned}
&\frac{a_1 a_2}{a_1 + a_2} e_P(X) - a_1 e_P(X_1) - a_2 e_P(X_2) \\
&\qquad = \frac{a_1 a_2}{a_1 + a_2} e_{P'}(X') - a_1 e_{P'}(X_1') - a_2 e_{P'}(X_2') + \frac{a_1}{a_1 + a_2} h^P(x^*) \\
&\qquad \geq t\frac{\min(a_1, a_2)}{2(a_1 + a_2)}(|X'| - \mathrm{bal}^{P'}(t)) + \frac{\min(a_1, a_2)}{a_1 + a_2} h^P(x^*) \\
&\qquad \geq t\frac{\min(a_1, a_2)}{2(a_1 + a_2)}(|X| - \mathrm{bal}^P(t)),
\end{aligned}$$

where the first inequality follows by induction, and the last inequality follows from our bound on $h^P(x^*)$. $\square$

Next, we use Lemma 3.3.2 as an induction step to prove the following theorem, which is the upper bound in Theorem 3.1.3(b).

Theorem 3.3.3 *Fix an integer $k \geq 3$ and positive real numbers $a_1, \dots, a_k$. There exists a constant $c = c(k, a_1, \dots, a_k)$ such that*

$$f(m; a_1, \dots, a_k) \leq \Big(\sum_{i=1}^{k} a_i^{-1}\Big)^{-1} m - c\sqrt{m}.$$

Proof For $r = 1, \dots, k-1$, let $b_r = (\sum_{i=1}^{r} a_i^{-1})^{-1}$ and $b'_r = (\sum_{i=r+1}^{k} a_i^{-1})^{-1}$. For $r < s$, let

$$v_{r,s} = \left|\frac{b'_r b_r^{-1} - b'_s b_s^{-1}}{b_r^{-1} + b_s^{-1}}\right| > 0 \quad \text{and} \quad w_{r,s} = \left|\frac{b_r b_r'^{-1} - b_s b_s'^{-1}}{b_r'^{-1} + b_s'^{-1}}\right| > 0.$$

Let $t_{r,s} = \frac{1}{2}\min(v_{r,s}, w_{r,s}, b_r, b'_r, b_s, b'_s) > 0$ and let $t = \min_{r<s} t_{r,s} > 0$. Finally, let

$$c = c(k, a_1, \dots, a_k) = \Big(\min_{i=1,\dots,k-1} \frac{\min(b_i, b'_i)}{2(b_i + b'_i)}\Big) t \Big(1 - \frac{1}{k-1}\Big).$$

Note that $c > 0$ since $k \geq 3$.

Let $P = (X, \prec)$ be a poset, which, we may assume, has no isolated elements. It is sufficient to show that there exists an ordered partition $X_1, \dots, X_k$ of P such that

$$\sum_{i=1}^{k} a_i e_P(X_i) \leq \Big(\sum_{i=1}^{k} a_i^{-1}\Big)^{-1} e_P(X) - c\sqrt{m}.$$

We show first that our choice of t ensures that for $x \in X$, there can be at most one value of r for which $x \in \mathrm{Bal}^P_{b_r, b'_r}(t)$. Suppose not. Then there exists $1 \leq r < s \leq k-1$ such that

$$\begin{aligned} |h_{b_r,b'_r}(x)| &= |b_r u(x) - b'_r d(x)| \leq t \\ \text{and} \quad |h_{b_s,b'_s}(x)| &= |b_s u(x) - b'_s d(x)| \leq t. \end{aligned}$$

Dividing the first equation by b_r and the second by b_s and subtracting the resulting equations, we obtain

$$|b'_r b_r^{-1} - b'_s b_s^{-1}| d(x) \leq \Big| |u(x) - b'_r b_r^{-1} d(x)| - |u(x) - b'_s b_s^{-1} d(x)| \Big| \leq (b_r^{-1} + b_s^{-1}) t.$$

Thus we have that $v_{r,s} d(x) \leq t$, and by a similar argument, we have that $w_{r,s} u(x) \leq t$. Since x is not isolated, either $d(x) \geq 1$ or $u(x) \geq 1$, whence we have that $t \geq \min(v_{r,s}, w_{r,s})$. But this contradicts our choice of t.

Hence, $x \in \mathrm{Bal}^P_{b_r, b'_r}(t)$ for at most one value of $r = 1, \ldots, k-1$. Let R be the value of r that minimises $\mathrm{bal}^P_{b_r, b'_r}(t)$. Then we have that

$$\mathrm{bal}^P_{b_R, b'_R}(t) \leq \frac{|X|}{k-1}.$$

By Lemma 3.3.2 there exists an ordered partition Y_1, Y_2 of P such that

$$\begin{aligned} b_R e_P(Y_1) &+ b'_R e_P(Y_2) \\ &\leq \frac{b_R b'_R}{b_R + b'_R} e_P(X) - t \frac{\min(b_R, b'_R)}{2(b_R + b'_R)} (|X| - \mathrm{bal}^P_{b_R, b'_R}(t)) \\ &\leq \Big(\sum_{i=1}^{k} a_i^{-1} \Big)^{-1} e_P(X) - t \frac{\min(b_R, b'_R)}{2(b_R + b'_R)} \Big(1 - \frac{1}{k-1} \Big) |X| \\ &\leq \Big(\sum_{i=1}^{k} a_i^{-1} \Big)^{-1} e_P(X) - c\sqrt{m}. \end{aligned}$$

(Note that our choice of t is consistent with the condition in Lemma 3.3.2.) By Lemma 3.2.1, we can find an ordered partition of Y_1 into sets $X_1, \ldots, X_R$ and of Y_2 into sets $X_{R+1}, \ldots, X_k$ such that

$$\sum_{i=1}^{R} a_i e_P(X_i) \leq b_R e_P(Y_1)$$

$$\text{and} \sum_{i=R+1}^{k} a_i e_P(X_i) \leq b'_R e_P(Y_2).$$

Then $X_1, \ldots, X_k$ is the desired ordered partition of P. □

We end this subsection by showing that the error term in the bound given by Theorem 3.3.3 is of the correct order of magnitude, that is, we complete the proof of Theorem 3.1.3(b).

Proof (of Theorem 3.1.3(b)) We have shown the upper bound in Theorem 3.3.3. For the lower bound, assuming that $\binom{n}{2} \leq m < \binom{n+1}{2}$ and recalling that C_n is the chain on n elements, we have

$$f(m; a_1, \ldots, a_k) \geq f\left(\binom{n}{2}, a_1, \ldots, a_k\right) \geq f(C_n; a_1, \ldots, a_k)$$

Let $X_1, \ldots, X_k$ be the ordered partition of C_n in which $|X_i| = x_i$ for each i and $x_1 + \cdots + x_k = n$. Then we have that

$$\sum_{i=1}^{k} a_i e_{C_n}(X_i) = \sum_{i=1}^{k} a_i \binom{x_i}{2} \geq \sum_{i=1}^{k} \frac{1}{2} a_i (x_i - 1)^2.$$

For $\mathbf{x} = (x_1, \ldots, x_k) \in \mathbb{R}^k$, define $g(\mathbf{x}) = \sum_{i=1}^{k} \frac{1}{2} a_i (x_i - 1)^2$, so that

$$f(C_n; a_1, \ldots, a_k) \geq \min_{x_1 + \cdots + x_k = n} g(\mathbf{x}).$$

We find that the minimum occurs at $\mathbf{y} = (y_1, \ldots, y_k)$ where

$$y_i - 1 = \left(a_i \sum_{j=1}^{k} a_j^{-1}\right)^{-1} (n - k).$$

Indeed, since g is a convex function, any point that is stationary in all directions along the constrained domain is a global minimum (on the constrained domain). We see that $\mathbf{y}$ is certainly a point on the constrained domain since

$$\sum_{i=1}^{k} y_i = n.$$

Then checking that

$$g'(\mathbf{y}) = \big(a_i(y_i - 1)\big)_{i=1}^{k} = \Big(\big(\sum_{j=1}^{k} a_j^{-1}\big)^{-1} (n - k)\Big)_{i=1}^{k}$$

is parallel to $(1, 1, \ldots, 1)$, that is, $g'(\mathbf{y})$ is normal to the plane $x_1 + \cdots + x_k = 0$, we see that $\mathbf{y}$ is indeed a point that is stationary in all directions along the constrained domain, and hence $\mathbf{y}$ minimises g. Therefore

$$f(m; a_1, \dots, a_k) \geq f(C_n; a_1, \dots, a_k) \geq g(\mathbf{y}) = \Big(\sum_{i=1}^{k} a_i^{-1}\Big)^{-1} \frac{1}{2}(n-k)^2$$
$$= \Big(\sum_{i=1}^{k} a_i^{-1}\Big)^{-1} m - \Theta(\sqrt{m}).$$

□

We remark that the calculation above, when $k = 2$, shows that for arbitrary positive real values of a_1 and a_2, we have

$$f(m; a_1, a_2) = \Big(\frac{1}{a_1} + \frac{1}{a_2}\Big)^{-1} m - O(\sqrt{m}).$$

3.3.4 Extremal Results

For the special case when $k = 2$ and $a_1 = a_2 = 1$, we can give the exact values of $f(m; 1, 1)$. We shall make use of the remarks after Lemma 3.2.1.

Theorem 3.3.4 *For m a positive integer, we have that*

$$f(m; 1, 1) = \Big\lfloor \frac{m-1}{2} \Big\rfloor.$$

Proof By Lemma 3.2.1, we have that $f(P; 1, 1) \leq \frac{1}{2} e_P(X)$ for every poset $P = (X, \prec)$, and furthermore, the inequality is strict if $e_P(X) > 0$. Thus for $m > 0$, we have $f(m; 1, 1) < \frac{m}{2}$ or equivalently $f(m; 1, 1) \leq \lfloor (m-1)/2 \rfloor$.

Recall the poset $P_r = P(1, 1, r)$, defined in the introduction of this chapter. We saw that for m odd, we have $f(m; 1, 1) \geq f(P_{(m-1)/2}; 1, 1) = (m-1)/2$. For m even, taking disjoint copies of P_0 and $P_{(m-2)/2}$, which we denote by $P_0 \sqcup P_{(m-2)/2}$, we have that $f(m; 1, 1) \geq f(P_0 \sqcup P_{(m-2)/2}; 1, 1) = \lfloor (m-1)/2 \rfloor$. This proves that the values of $f(m; 1, 1)$ are as stated. □

We conclude this section by showing how we can use Lemma 3.3.2 in proving the uniqueness of $P_{(m-1)/2}$ as the extremal poset corresponding to $f(m; 1, 1)$ when m is odd.

Theorem 3.3.5 *Fix m an odd positive integer. If P is a poset with m comparable pairs and no isolated elements, and*

$$f(P; 1, 1) = f(m; 1, 1),$$

then $P = P_{(m-1)/2}$.

Proof Let P be a poset as in the premise of the theorem, so that $f(P;1,1) = f(m;1,1) = (m-1)/2$. We apply Lemma 3.3.2 to the poset P, where $a_1 = a_2 = 1$ and t is any fixed number in the range $2/3 < t < 1$. Thus, we have

$$\frac{m-1}{2} = f(P;1,1) \leq \frac{1}{2}m - \frac{1}{4}t(|X| - \mathrm{bal}^P_{1,1}(t)).$$

Therefore, we must have that $|X| - \mathrm{bal}^P_{1,1}(t) \leq 2/t$, whence there are at most two elements of P that are not in $\mathrm{Bal}^P_{1,1}(t)$ (since $t > 2/3$). But maximal and minimal elements of P are not in $\mathrm{Bal}^P_{1,1}(t)$ (since $t < 1$); hence there are exactly two elements, which we call y_1 and y_2, that are not in $\mathrm{Bal}^P_{1,1}(t)$.

Observe that, since $h_{1,1}$ is an integer function, if $x \in \mathrm{Bal}^P_{1,1}(t)$ for $t < 1$, then $h_{1,1}(x) = 0$. Also, since $h_{1,1}$ is a strictly decreasing function, then $\mathrm{Bal}^P_{1,1}(t)$ must be an antichain. Since the elements in $\mathrm{Bal}^P_{1,1}(t)$ are not isolated, they must each be (without loss of generality) above y_1 and below y_2 (in order to ensure that $h_{1,1}(x) = 0$ for each $x \in \mathrm{Bal}^P_{1,1}(t)$). Thus we have that $P = P_r$ for some r, and since P has m comparable pairs, we must have that $r = (m-1)/2$. □

We end this subsection with the following conjecture about the exact value of f when $k \geq 3$.

Conjecture 3.3.6 *Let $k \geq 3$ be a fixed integer, and let $a_1 = \cdots = a_k = 1$. For $m = \binom{n}{2}$, we have*

$$f(m; a_1, \ldots, a_k) = f(C_n; a_1, \ldots, a_k).$$

Examples like P_r fail to be extremal when $k \geq 3$ because of the increased freedom we have when partitioning into three or more parts. Informally, it seems that this increased freedom, together with transitivity in posets, allows us to create partitions where a large number of comparable pairs go across parts. Thus, in order to construct an extremal example, we also require a large number of comparable pairs within parts. Chains seem the most likely candidates to satisfy this.

3.4 Best Partitions

In this section, we give an algorithm that finds an optimal ordered partition for any given poset P. More precisely, we have the following theorem.

Theorem 3.4.1 *There exists a strongly polynomial-time algorithm that, for an input* $(P, k, a_1, \ldots, a_k)$, *where* $P = (X, \prec)$ *is a poset,* k *is a positive integer, and* $a_1, \ldots, a_k$ *are positive rationals, outputs an ordered partition* $X_1, \ldots, X_k$ *of* P *for which*

$$\sum_{i=1}^{k} a_i e(X_i) = f(P; a_1, \ldots, a_k).$$

We give a brief non-technical description of what is meant by *strong polynomial time.* Strong polynomial time is only of relevance for problems that include numbers in the input. Following [30], the main difference between a polynomial-time algorithm and a strongly polynomial-time algorithm is that a polynomial-time algorithm performs a number of operations which is bounded by a polynomial in the input size, whereas a strongly polynomial-time algorithm performs a number of elementary arithmetic operations (addition, subtraction, multiplication, division, comparison) which is bounded by a polynomial in the number of input numbers. If the input has some non-numerical part (e.g. a graph or a name), we assume it is encoded as a $\{0, 1\}$ sequence, and each entry of this sequence is considered a number. Correspondingly, we consider nonnumeric steps in the algorithm (like setting or removing a label, deleting an edge from a graph) as arithmetic operations. Thus a strongly polynomial-time algorithm has run time that is independent of the numerical data size and depends only on the inherent dimensions of the problem. See [30] for more details.

Thus, in Theorem 3.4.1, the algorithm whose existence we claim performs a number of arithmetic operations that is polynomial in $|X|$ and k.

Before we give the algorithm in general, we give simpler algorithms for two special cases. We start by giving a particularly simple algorithm for the case $k = 2$, where in fact, we are able to find all optimal ordered partitions in polynomial time.

Theorem 3.4.2 *Let* a_1 *and* a_2 *be positive real numbers. For a poset* $P =$

$(X, \prec)$ *together with an ordered partition* X_1, X_2 *of* P*, we have that*

$$a_1 e(X_1) + a_2 e(X_2) = f(P; a_1, a_2)$$

if and only if X_1, X_2 *is an* (a_1, a_2)*-good partition of* P*.*

We note that if a_1 and a_2 are rationals, then all (a_1, a_2)-good partitions can be found in strongly polynomial time. Indeed, letting $n = |X|$, computing $u(x)$ and $d(x)$ for a given $x \in X$ requires $O(n)$ arithmetic operations, so computing $h_{a_1,a_2}(x)$ also requires $O(n)$ arithmetic operations. Knowing whether $h(x)$ is positive, negative, or zero determines whether $x \in X_1$, $x \in X_2$, or if x can be chosen to be in either. Doing this for all $x \in X$ thus requires $O(n^2)$ arithmetic operations.

Proof Given an ordered partition X_1, X_2, observe that

$$\begin{aligned} a_1 e(X_1) + a_2 e(X_2) &= a_2 \sum_{x \in X_2} u(x) + a_1 \sum_{x \in X_1} d(x) \\ &= a_2 \sum_{x \in X} u(x) + \sum_{x \in X_1} (a_1 d(x) - a_2 u(x)) \\ &= a_2 e(X) - \sum_{x \in X_1} h_{a_1,a_2}(x). \end{aligned}$$

This is minimised if and only if $h_{a_1,a_2}(x) \geq 0$ for all $x \in X_1$ and $h_{a_1,a_2}(x) \leq 0$ for all $x \in X_2$, that is, X_1, X_2 is an (a_1, a_2)-good partition. □

Given that the above theorem tells us exactly which ordered partitions are optimal, one might expect that we can use this information directly to bound $f(P; a_1, a_2)$ rather than using the inductive proof of Lemma 3.2.1. Such an argument has eluded us.

For the case $k \geq 3$, one might expect that we can apply Theorem 3.4.2 repeatedly to give an optimal ordered partition into k parts. The obstruction to this is that, by performing our optimisation sequentially, our choice of partition at one stage affects the poset we are required to partition at subsequent stages; thus, sacrificing optimality at an earlier stage can leave us with posets with larger ordered cuts for later stages.

Let us consider the case for general k, but where $a_1 = \cdots = a_k = 1$. On this occasion it will be more convenient for us to maximise $e_P(X_1, \ldots, X_k)$ rather than to minimise $e_P(X_1) + \cdots + e_P(X_k)$. We show that this problem can be reduced to one of finding a maximum sized union of $k-1$ antichains in a certain poset constructed from P. The problem of finding this maximum sized union of $k-1$ antichains is known to be solvable in polynomial time, by Corollary 1.2.6. We thank Omid Amini and Stéphan Thomassé for the argument that follows.

We begin by noting that, given a poset $P = (X, \prec)$, we have $A \subseteq X$ is a union of $k-1$ antichains if and only if P induces a poset of height at most $k-1$ on A. (Indeed, if $P_A = (A, \prec)$ has height at most $k-1$, then by Theorem 1.2.3, A can be decomposed into $k-1$ antichains (where we permit the empty antichain). Conversely, if P_A is the union of $k-1$ antichains, then no set of k elements can form a chain; otherwise at least two of the elements would be part of the same antichain. Therefore P_A has height at most $k-1$.)

We say that $Y \subseteq X$ is a *maximal* union of $k-1$ antichains if there is no $Z \supsetneq Y$ that is also a union of $k-1$ antichains.

We define the *line poset* of P to be $L(P) = (E^*_P(X), \prec_{L(P)})$, where

$$E^*_P(X) = \{(a,b) : a \prec b\},$$

and where $(a_1,b_1) \prec_{L(P)} (a_2,b_2)$ if and only if $b_1 \preceq a_2$. (We have that $\prec_{L(P)}$ is irreflexive since if $(a_1,b_1) \prec_{L(P)} (a_1,b_1)$, then $b_1 \preceq a_1$, contradicting that $(a_1,b_1) \in E^*_P(X)$. Also $\prec_{L(P)}$ is transitive since if $(a_1,b_1) \prec_{L(P)} (a_2,b_2)$ and $(a_2,b_2) \prec_{L(P)} (a_3,b_3)$, then $b_1 \preceq a_2 \preceq b_2 \preceq a_3$, and so $(a_1,b_1) \prec_{L(P)} (a_3,b_3)$.)

For an ordered k-partition $X_1, \ldots, X_k$, we define

$$E^*_P(X_1, \ldots, X_k) = \{(a,b) \in E^*_P(X) : a \in X_i, b \in X_j, i < j\}.$$

Clearly $|E^*_P(X_1, \ldots, X_k)| = e_P(X_1, \ldots, X_k)$ and $|E^*_P(X)| = e_p(X)$.

We have the following lemma.

Lemma 3.4.3 *Given a positive integer k and a poset $P = (X, \prec)$, we have that $Y \subseteq E^*_P(X)$ is a union of at most $k-1$ antichains in the line poset $L(P)$ if and only if $Y \subseteq E^*_P(X_1, \ldots, X_k)$ for some ordered partition $X_1, \ldots, X_k$ of P.*

Therefore, finding an ordered k-partition that maximises $e_P(X_1, \ldots, X_k)$ is equivalent to finding a maximum sized union of $k-1$ antichains in $L(P)$, where the latter can be found in time polynomial in $|P|$ (since $|L(P)|$ is polynomially bounded in $|P|$).

Proof Let $X_1, \ldots, X_k$ be an ordered partition of P. We have that if $Y \subseteq E_P^*(X_1, \ldots, X_k)$, then $L(P)$ induces a poset of height at most $k-1$ on Y. (Indeed, define $\omega : X \to [k]$ by setting $\omega(x) = i$ if and only if $x \in X_i$. If we have a chain $(a_1, b_1) \prec_{L(P)} (a_2, b_2) \prec_{L(P)} \cdots \prec_{L(P)} (a_r, b_r)$ of $L(P)$ in Y, then we have $\omega(a_1) < \omega(b_1) \leq \omega(a_2) < \omega(b_2) \leq \cdots \leq \omega(a_r) < \omega(b_r)$. But then we must have $r \leq k-1$ as required.) Therefore, Y is the union of at most $k-1$ antichains in $L(P)$.

Conversely, let $Y \subseteq E_P^*(X)$ be a maximal union of $k-1$ antichains in $L(P)$, that is, $L(P)$ induces a poset of height $k-1$ on Y. We prove, by induction on k, that there exists an ordered partition $X_1, \ldots, X_k$ such that $Y = E_P^*(X_1, \ldots, X_k)$. This then proves the lemma.

When $k = 1$, Y has height zero, so is empty and corresponds to the ordered partition with one part, namely the whole of X.

For general k, let $Y_1 = \min_{L(P)}(Y)$ and let $X_1 = \{a : \exists(a, b) \in Y_1\}$. Let $Y' = Y \backslash Y_1$ and $X' = X \backslash X_1$, and let $P' = (X', \prec)$. We claim that

(a) X_1 is a down-set of P, and

(b) Y' is a maximal union of $k-2$ antichains in $L(P')$.

Let us continue with the proof assuming the claim is true. By induction, there exists an ordered partition $X_2, \ldots, X_k$ of P' such that $Y' = E_{P'}^*(X_2, \ldots, X_k)$. Since X_1 is a down-set of P, we have that $X_1, \ldots, X_k$ forms an ordered partition of P and that $Y = Y_1 \cup Y' \subseteq E_P^*(X_1, X') \cup E_P^*(X_2, \ldots, X_k) = E_P^*(X_1, \ldots, X_k)$. Since Y is maximal, we must have $Y = E_P^*(X_1, \ldots, X_k)$.

For part (a) of the claim, let $a \in X_1$ with $b \in X$ and $b \prec a$. Since $a \in X_1$, there exists some $a' \in X$ with $a \prec a'$ such that $(a, a') \in Y_1 = \min_{L(P)}(Y)$. We show below that (b, a') is incomparable to every element in $Y_1 = \min_{L(P)}(Y)$, hence $(b, a') \in Y_1$ (by the maximality of Y), thus $b \in X_1$, proving part (a) of the claim.

In order to show that (b, a') is incomparable to every element in Y_1, suppose $(a_1, a_2) \in Y_1$. Then we cannot have $(a_1, a_2) \prec_{L(P)} (b, a')$, otherwise we have $(a_1, a_2) \prec_{L(P)} (a, a')$ contradicting that $(a, a') \in Y_1$. We cannot have $(b, a') \prec_{L(P)} (a_1, a_2)$, otherwise $(a, a') \prec_{L(P)} (a_1, a_2)$, contradicting that $(a_1, a_2) \in Y_1$.

For part (b) of the claim, it is clear that $Y' \subseteq E^*_{P'}(X')$. It is also clear that $Y' = Y \setminus \min_{L(P)}(Y)$ is a maximal union of $k-2$ antichains in $L(P')$ (by noting that $ht(Y') = ht(Y) - 1$ and that Y' must be maximal in $L(P')$ since Y is maximal in $L(P)$). This completes the proof. □

We now turn to the proof for the general weighted case, Theorem 3.4.1, which relies on the fact that a submodular function on a lattice family can be minimised in strongly polynomial time. This result was originally due to Grötschel, Lovász, and Schrijver [29, 30] and was refined most notably by Iwata, Fleischer, and Fujishige [33] and Schrijver [53]. We begin with some preliminaries.

Given a set V, a set L of subsets of V (with the inclusion order) is called a *lattice family* if, whenever $A, B \in L$, we have $A \cap B \in L$ and $A \cup B \in L$. For example, the set of down-sets of a poset P on V, which we denote by $D(P)$, forms a lattice family (since the union and intersection of any two down-sets is also a down-set).

A function $g : L \to \mathbb{R}$ is called *submodular* if

$$g(A) + g(B) \geq g(A \cap B) + g(A \cup B)$$

for all $A, B \in L$.

We have the following special case of a result of Schrijver [53].

Theorem 3.4.4 *Let $D(P)$ be the set of down-sets of some partial order $P = (V, \prec)$. Let g be a submodular function on $D(P)$. Given a value-giving oracle for g, a set $U \in D(P)$ that minimises g can be found by an algorithm using a number of calls to the oracle and a number of arithmetic steps that are both polynomial in $|V|$.*

The value-giving oracle is able to access values of g in polynomial time. It is required because we would like an algorithm whose running time is polynomial in $|V|$; we do not wish to input all values of g since this would require $|D(P)|$ operations, where $|D(P)|$ is potentially exponential in $|V|$.

Fix a poset $P = (X, \prec)$ and a positive integer k. We shall utilise order-preserving functions in place of ordered partitions via the natural correspondence of Proposition 1.2.2. Recall that for $X_1, \ldots, X_k$ an ordered k-partition of P, the function $\omega : X \to [k]$ defined by setting $\omega(x) = i$ if and only if $x \in X_i$ is the order-preserving function corresponding to $X_1, \ldots, X_k$ in

Proposition 1.2.2. Let $\Omega_k(P)$ be the set of all order-preserving functions corresponding to the ordered k-partitions of P. It is well known that $\Omega_k(P)$ has a natural lattice structure, but we shall show it explicitly in Lemma 3.4.5.

First we define the poset $R_{k-1}(P)$, which is derived from our poset $P = (X, \prec)$. Let $X = \{x_1, \ldots, x_n\}$. Define $R_{k-1}(P) = (Y, \prec^*)$, where $Y = Y_1 \cup \cdots \cup Y_{k-1}$, $Y_i = \{y_{i1}, \ldots y_{in}\}$ for $i = 1, \ldots, k-1$, and $y_{ir} \prec^* y_{js}$ if and only if $i \geq j$ and $x_r \preceq x_s$ (assuming y_{ir} and y_{js} are distinct). Then $\prec^*$ is an order since it is irreflexive by definition and it inherits the transitivity of $\preceq$ and $\geq$. We can view $R_{k-1}(P)$ as simply $k-1$ copies of P stacked on top of each other (Y_1 at the top, Y_{k-1} at the bottom), where any given element of P in a higher layer is greater than (in $\prec^*$) the same element in a lower layer.

Recall that $D(R_{k-1}(P))$ is the set of down-sets of $R_{k-1}(P)$. For $\omega \in \Omega_k(P)$, define $d(\omega)$ to be the subset of Y such that $y_{ir} \in D(\omega)$ if and only if $i \geq \omega(x_r)$. It is easy to see that $D(\omega)$ is a down-set of $R_{k-1}(P)$ (indeed, if $y_{js} \in d(\omega)$ and $y_{ir} \prec^* y_{js}$, then $i \geq j \geq \omega(x_s)$ and $x_s \succeq x_r$; hence $i \geq j \geq \omega(x_s) \geq \omega(x_r)$ and so $y_{ir} \in D(R_{k-1}(P))$).

We now show the natural correspondence between $\Omega_k(P)$ and the lattice family $D(R_{k-1}(P))$.

Lemma 3.4.5 *Let $P = (X, \prec)$ be given. The function $d : \Omega_k(P) \rightarrow D(R_{k-1}(P))$, where $\omega \mapsto d(\omega)$ is as defined above, is a bijection. Furthermore, if $\omega_1 \mapsto D_1$ and $\omega_2 \mapsto D_2$, then $\min(\omega_1, \omega_2) \mapsto D_1 \cup D_2$ and $\max(\omega_1, \omega_2) \mapsto D_1 \cap D_2$, where $\min(\omega_1, \omega_2)$ (resp. $\max(\omega_1, \omega_2)$) is the pointwise minimum (resp. maximum) of ω_1 and ω_2.*

Proof Recall that $R_{k-1}(P) = (Y, \prec^*)$ consists of layers $Y_1, \ldots, Y_{k-1}$, where each Y_i is a copy of P. Thus any down-set D of $R_{k-1}(P)$ induces a down-set D_i of P on Y_i. Furthermore the Y_i are stacked on top of each other (Y_1 at the top, Y_{k-1} at the bottom), where any given element of P in a higher layer is greater than (in $\prec^*$) the same element in a lower layer. This implies that $D_1 \subseteq \ldots \subseteq D_{k-1}$. Clearly the above correspondence between down-sets of $R_{k-1}(P)$ and sequences of $k-1$ down-sets in P is injective. Thus, since nested sequences of $k-1$ down-sets are in bijective correspondence with ordered k-partitions by Proposition 1.2.1, we have that $|D(R_{k-1}(P)| \leq |\Omega_k(P)|$. Now, in order to show that d is bijective, it is sufficient to show that it is injective. Indeed. if $\omega_1, \omega_2 \in \Omega_k(P)$ are distinct, then there exists

$x \in X$ such that $i := \omega_1(x) < \omega_2(x) =: j$, and so the element corresponding to x in Y_i is in $d(\omega_1)$ but not in $d(\omega_2)$.

We prove the second part of the lemma. Suppose $\omega_1, \omega_2 \in \Omega_k(P)$.

We have that $y_{ir} \in d(\min(\omega_1, \omega_2))$ if and only if $i \geq \min(\omega_1, \omega_2)(x_r)$, or equivalently $i \geq \omega_1(x_r)$ or $i \geq \omega_2(x_r)$, that is precisely if $y_{ir} \in d(\omega_1) \cup d(\omega_2)$. Thus $d(\min(\omega_1, \omega_2)) = d(\omega_1) \cup d(\omega_2)$.

We have that $y_{ir} \in d(\max(\omega_1, \omega_2)$ if and only if $i \geq \max(\omega_1, \omega_2)(x_r)$, or equivalently $i \geq \omega_1(x_r)$ and $i \geq \omega_2(x_r)$, that is precisely if $y_{ir} \in d(\omega_1) \cap d(\omega_2)$. Thus $d(\max(\omega_1, \omega_2)) = d(\omega_1) \cap d(\omega_2)$. □

We are now ready to prove Theorem 3.4.1.

Proof (of Theorem 3.4.1) A poset $P = (X, \prec)$, a positive integer k, and positive real numbers $a_1, \ldots, a_k$ are given. We define $g : \Omega_k(P) \to \mathbb{R}$ by setting $g(\omega) = \sum_{i=1}^k a_i e_P(X_i)$, where ω is the order-preserving function corresponding to $X_1, \ldots, X_k$. We note that we can compute $g(\omega)$ using a number of arithmetic operations polynomial in n and k.

Invoking Theorem 3.4.4, we see that if g is submodular, then we can minimise g over $\Omega_k(P)$ using a number of arithmetic operations that is polynomial in $|R_{k-1}(P)| = (k-1)n \leq n^2$. That is, we can minimise g in strongly polynomial time, thus proving the theorem.

It remains only to show that g is submodular, that is, we wish to show that if $\omega, \phi \in \Omega_k(P)$, then

$$g(\omega) + g(\phi) \geq g(\max(\omega, \phi)) + g(\min(\omega, \phi)).$$

Noting that the sum of submodular functions is submodular, we write g as a sum of indicator functions and show that each indicator function is submodular.

Define $I_{x,y,i} : \Omega_k(P) \to \{0, 1\}$, where

$$I_{x,y,i}(\omega) = \begin{cases} 1 \text{ if } \omega(x) = \omega(y) = i; \\ 0 \text{ otherwise.} \end{cases}$$

Then we have that

$$g = \sum_{i=1}^{k} a_i \sum_{x \prec y} I_{x,y,i}.$$

We now carry out an easy case analysis to show that if $x \prec y$, then $I_{x,y,i}$ is submodular. We wish to show, for every pair $\omega, \phi \in \Omega_k(P)$, that

$$I_{x,y,i}(\omega) + I_{x,y,i}(\phi) \geq I_{x,y,i}(\max(\omega, \phi)) + I_{x,y,i}(\min(\omega, \phi)). \tag{3.2}$$

Since $x \prec y$, we have that $\omega(x) \leq \omega(y)$ and $\phi(x) \leq \phi(y)$. Henceforth, we drop the subscripts on I.

Suppose $I(\max(\omega, \phi)) + I(\min(\omega, \phi)) = 2$. Then we have

$$\begin{aligned} & \max(\omega, \phi)(x) = \min(\omega, \phi)(x) = \max(\omega, \phi)(y) = \min(\omega, \phi)(y) = i, \\ \text{so} \quad & \omega(x) = \omega(y) = \phi(x) = \phi(y) = i, \\ \text{hence} \quad & I(\omega) + I(\phi) = 2. \end{aligned}$$

Suppose $I(\max(\omega, \phi)) + I(\min(\omega, \phi)) = 1$. Then without loss of generality, we have $\max(\omega, \phi)(x) = \max(\omega, \phi)(y) = i$. Without loss of generality, $\phi(x) \leq \omega(x) = i$. Now we have one of the following two possibilities:

$$\begin{aligned} & \text{(a) } \phi(y) \leq \omega(y) = i, \\ \text{or} \quad & \text{(b) } \omega(y) \leq \phi(y) = i. \end{aligned}$$

For case (a), we have that $\omega(x) = \omega(y) = i$, so that $I(\omega) + I(\phi) \geq 1$. For case (b) we have $\omega(y) \leq \omega(x) = i$, but we know that $\omega(x) \leq \omega(y)$ (since $x \prec y$). Hence $\omega(x) = \omega(y) = i$ and $I(\omega) + I(\phi) \geq 1$.

If $I(\max(\omega, \phi)) + I(\min(\omega, \phi)) = 0$ then (3.2) trivially holds. Thus (3.2) holds in all cases and the proof is complete. □

Chapter 4

A Linear Arrangement Problem for Posets

This chapter is based on joint work (in equal part) with Graham Brightwell.

4.1 Introduction

In graph theory, the *linear arrangement problem* or *optimal arrangement problem* or *wire-length problem* is the following: given a graph $G = (V, E)$, where $|V| = n$ and $|E| = m$, find a function amongst all bijective functions $f : V \to [n]$ that minimises

$$\frac{1}{m} \sum_{ab \in E} |f(a) - f(b)|.$$

Note that the factor $1/m$ makes no difference to this problem and is generally omitted. Also, since

$$\sum_{ab \in V^{(2)}} |f(a) - f(b)| = \sum_{1 \leq i < j \leq n} (j - i) = \frac{1}{6}(n+1)n(n-1),$$

we see that the maximisation problem for a given graph G is equivalent to the minimisation problem for its complement.

The linear arrangement problem is known to be NP-hard (see [26]), and furthermore, there are few classes of graphs for which this problem is known to be polynomial-time solvable. The problem, which is fairly well studied, falls inside a more general class of problems called *graph layout problems.*

These ask for an ordering of graph vertices so as to optimise some objective function of edge lengths. For a survey of such problems, see [10].

We formulate a natural analogue of the linear arrangement problem for posets. Given a poset $P = (X, \prec)$ with $|X| = n$, recall that a linear extension λ of P is a bijection, $\lambda : P \to [n]$, which satisfies the condition that $\lambda(a) < \lambda(b)$ whenever $a \prec b$ for every pair of elements $a, b \in X$. We write Λ_P for the set of all linear extensions of P.

Given a linear extension λ of $P = (X, \prec)$ and $a, b \in X$ with $a \prec b$, we define the *distance from a to b in* λ to be $\text{dist}(a, b; \lambda) = \lambda(b) - \lambda(a)$. The *average relational distance in* λ, $\text{dist}_P(\lambda)$, is given by

$$\text{dist}_P(\lambda) = \frac{1}{m} \sum_{(a,b):a \prec b} \text{dist}(a, b; \lambda) = \frac{1}{m} \sum_{(a,b):a \prec b} (\lambda(b) - \lambda(a)),$$

where m is the number of comparable pairs in P. For this to make sense, we require that $m > 0$, and so we shall assume throughout this chapter that P is not an antichain.

Clearly dist_P is a natural function to consider on Λ_P, and in this chapter, we give some of its properties. In contrast to the linear arrangement problem for graphs, we show that an element of Λ_P maximising dist_P can be found in polynomial time. The algorithm is very simple and makes use of the polynomial-time algorithm for the poset maxcut problem discussed in the previous chapter.

Remark The problem we consider is not simply a restriction of the graph linear arrangement problem to comparability graphs of posets: we are maximising over *linear extensions* of P rather than arbitrary bijections.

Indeed, consider the following example, which shows that we can obtain a higher average relational distance if we maximise over arbitrary bijections rather than linear extensions. Let $P_r^* = (X, \prec^*)$ be the poset where X consists of the elements $x, x_1, \ldots, x_r, y, y_1, \ldots, y_r$, and where $x \prec^* x_i$ for $i = 1, \ldots, r$ and $y \prec^* y_i$ for $i = 1, \ldots, r$. Thus P_r^* has $2r + 2$ elements and $2r$ relations.

We note that all linear extensions λ of P_r^* in which x and y are the first two elements, that is $\{\lambda(x), \lambda(y)\} = \{1, 2\}$, have the same average relational distance, and moreover these linear extensions turn out to maximise the average relational distance. For the ordering $y, x, x_1, \ldots, x_r, y_1, \ldots, y_r$, this distance is

$$\frac{1}{2r}\Big(\sum_{i=1}^{r} i + \sum_{i=1}^{r}(r+1+i)\Big) = \frac{1}{2r}\Big(r(r+1) + 2\sum_{i=1}^{r} i\Big) = \frac{2r(r+1)}{2r} = r+1.$$

However, if we permit arbitrary bijections from X to $[2r+2]$, then the average relational distance is maximised by the bijection that orders the elements $x, y_1, \dots, y_r, x_1, \dots, x_r, y$, and its average relational distance is

$$\frac{1}{2r}\Big(2\sum_{i=1}^{r}(r+i)\Big) = \frac{2r^2 + r(r+1)}{2r} > r+1 \quad \text{for } r \geq 2.$$

Remark For posets, the maximisation and minimisation problems are not equivalent in the sense they are for graphs. Indeed, whereas we show that the maximisation problem for posets is polynomial-time solvable, we believe that the minimisation problem, that is, the problem of minimising $\mathrm{dist}_P(\lambda)$ over all linear extensions λ of P, is NP-hard.

Remark The maximisation problem we have described for posets is equivalent to the following minimisation problem: given a poset P, minimise over all linear extensions λ of P, the function

$$\sum_{a \| b} |\lambda(a) - \lambda(b)|,$$

where $a \parallel b$ denotes that a and b are incomparable in P. This problem is related to the *linear discrepancy* of a poset P, denoted by $ld(P)$, and defined as the minimum over all linear extensions λ of P, of the function

$$\max_{a \| b} |\lambda(a) - \lambda(b)|.$$

This problem has been studied by Fishburn, Tanenbaum, and Trenk [60, 19], and is in turn related to the bandwidth problem for graphs, another graph layout problem. Fishburn, Tanenbaum, and Trenk [19] showed that the linear discrepancy of a poset P is always equal to the bandwidth of $\mathrm{Inc}(P)$, the incomparability graph of P.

In the final section of the chapter, we turn to an extremal problem. We prove the following extremal bound on dist_P: for any poset P on n elements that is not an antichain, we have

$$\max_{\lambda \in \Lambda_P} (\mathrm{dist}_P(\lambda)) \geq \frac{1}{3}(n+1).$$

Note that equality holds in the above bound for $P = C_n$, the chain on n elements. Exactly the same bound holds for the corresponding graph problem, and it is trivial to prove. Given a graph $G = (V, E)$, take a random bijection $f : V \to [n]$. We see that

$$\begin{aligned}
\mathbb{E}\Big(\frac{1}{m}\sum_{ab\in E}|f(a)-f(b)|\Big) &= \frac{1}{m}\sum_{ij\in[n]^{(2)}}|j-i|\mathbb{P}(\exists ab\in E : f(a)f(b)=ij) \\
&= \frac{1}{m}\Big(\frac{1}{6}(n-1)n(n+1)\Big)\frac{m}{\binom{n}{2}} \\
&= \frac{1}{3}(n+1).
\end{aligned}$$

Now the existence of the desired bijection is ensured.

Computing an expectation for our problem is not quite so simple since we are averaging over linear extensions of P rather than arbitrary bijections. Instead, we shall bound the expectation.

4.2 Maximisation of dist_P

We begin by recalling some notation from the previous chapter. Given a poset $P = (X, \prec)$ and $x \in X$, we define

$$\begin{aligned}
u(x) &= |\{y \in X : y \succ x\}| \text{ and} \\
d(x) &= |\{y \in X : y \prec x\}|.
\end{aligned}$$

For $A, B \subseteq X$, we define

$$\begin{aligned}
e_P(A) &= |\{(a,b) : a \prec b \text{ and } a, b \in A\}| \text{ and} \\
e_P(A,B) &= |\{(a,b) : a \prec b \text{ and } a \in A, b \in B\}|.
\end{aligned}$$

In practice, A will generally be a down-set and B an up-set of P with A and B disjoint.

We are now ready to prove our theorem.

Theorem 4.2.1 *Given a poset $P = (X, \prec)$, an element of Λ_P that maximises* dist_P *can be found in time polynomial in $|X|$.*

Proof Let $h : X \to \mathbb{Z}$ be the function defined by $h(x) = d(x) - u(x)$ for each $x \in X$. Observe that h is a strictly increasing function on P, that is, whenever $a, b \in X$ with $a \prec b$, we have $h(a) < h(b)$.

Now h imposes a partial order $P_h = (X, \prec_h)$ on X defined as follows. For $a, b \in X$, we have that $a \prec_h b$ if and only if $h(a) < h(b)$. (Note that P_h is a linear ordering if and only if h is injective.) Since h is a strictly increasing function with respect to P, we see that any linear extension of P_h is also a linear extension of P. We claim that the linear extensions of P that maximise dist_P are precisely the linear extensions of P_h.

Assuming the claim is true, the values of h can be found in time polynomial in $|X|$, and we can sort the elements of X according to their h-values in time polynomial in $|X|$ to give a linear extension of P_h as desired.

It remains only to prove the claim. Fix a linear extension λ of P. Let A_i be the set of the first $i-1$ elements of P in λ, and let B_i be the remaining elements of P, that is,

$$A_i = \{\lambda^{-1}(j) : j < i\} \text{ and}$$
$$B_i = \{\lambda^{-1}(j) : j \geq i\}.$$

The number of comparable pairs of P from A_i to B_i is denoted by $e_i = e_P(A_i, B_i)$. Given a comparable pair (a, b) of P, where $a \prec b$, we note that (a, b) is counted in $e_P(A_i, B_i)$ for values of i satisfying $\lambda(a) < i \leq \lambda(b)$. Therefore the comparable pair (a, b) is counted precisely $\lambda(b) - \lambda(a) = \text{dist}(a, b; \lambda)$ times in $\sum_{i=1}^{n} e_i$. Hence

$$\frac{1}{m}\sum_{i=1}^{n} e_i = \frac{1}{m}\sum_{(a,b):a\prec b} \text{dist}(a, b; \lambda) = \text{dist}_P(\lambda).$$

We now evaluate e_i in terms of h using the same argument as in the proof of Theorem 3.4.2. For each i, A_i is a down-set of P disjoint from B_i, which is an up-set of P. Therefore, we have

$$\begin{aligned} e_i &= e_P(X) - e_P(A_i) - e_P(B_i) \\ &= \sum_{x \in X} d(x) - \sum_{x \in A_i} d(x) - \sum_{x \in B_i} u(x) \\ &= \sum_{x \in B_i} d(x) - \sum_{x \in B_i} u(x) \\ &= \sum_{x \in B_i} h(x) = \sum_{j=i}^{n} h(\lambda^{-1}(j)). \end{aligned}$$

Now we have that

$$\mathrm{dist}_P(\lambda) = \frac{1}{m} \sum_{i=1}^{n} e_i = \frac{1}{m} \sum_{i=1}^{n} \sum_{j=i}^{n} h(\lambda^{-1}(j)) = \frac{1}{m} \sum_{i=1}^{n} i h(\lambda^{-1}(i)). \qquad (4.1)$$

We now see from the formula above that a linear extension λ of P maximises dist_P if and only if $h(\lambda^{-1}(i))$ is an increasing function of i, that is, if and only if λ is a linear extension of P_h. This proves our claim and completes the proof. □

Alternatively, one can prove that maximising dist_P is polynomial-time solvable by repeatedly performing local optimisations: given a linear extension λ of P, if we can switch a consecutive pair of elements in λ to obtain a linear extension for which dist_P is larger, then we make the switch. We iterate this process until no more switches can be made. It is easy to prove that what remains is an optimal linear extension. The proof above gives the explicit formula (4.1), which might prove to be useful elsewhere.

4.3 An Extremal Bound for dist_P

In this section, we prove the following theorem.

Theorem 4.3.1 *For every poset P that is not an antichain, there exists a linear extension λ^* such that*

$$\mathrm{dist}_P(\lambda^*) \geq \frac{1}{3}(n+1).$$

Proof Pick a linear extension μ of P uniformly at random. We shall prove in Lemma 4.3.3 that

$$\mathbb{E}(\mathrm{dist}_P(\mu)) \geq \frac{1}{3}(n+1).$$

This then ensures the existence of the desired linear extension. □

We give some notation. Fix a poset $P = (X, \prec)$, where $|X| = n$. For $i, j \in [n]$ with $i < j$, write

$$N_P(i,j) = \{\mu : \mu \text{ is a linear extension of } P,\ \mu^{-1}(i) \prec \mu^{-1}(j)\},$$

and let $n_P(i,j) = |N_P(i,j)|$, the number of linear extensions of P in which the element in the i^{th} position is below (in P) that in the j^{th} position. We have the following lemma.

Lemma 4.3.2 *Let $P = (X, \prec)$ be a poset and let i, j, i', j' be elements of $[n]$ such that $i \leq i' < j' \leq j$. Then*

$$n_P(i,j) \geq n_P(i',j').$$

Proof It is sufficient to prove that

$$n_P(a, b+1) \geq n_P(a,b)$$

for all applicable $a, b \in [n]$ with $a < b$. Indeed, we can then conclude by induction that $n_P(a, b+k_1) \geq n_P(a,b)$ for all $k_1 \in [n-b]$, and furthermore by symmetry, we have $n_P(a,b) \geq n_P(a+k_2, b)$ for all $k_2 \in [b-a]$. Thus, we have

$$n_P(a, b+k_1) \geq n_P(a+k_2, b),$$

and setting $a = i$, $b = j'$, $k_1 = j - j'$, and $k_2 = i' - i$ gives the desired inequality.

Fix $a, b \in [n]$ with $a < b$. In order to show that $n_P(a, b+1) \geq n_P(a,b)$, we give an injection from $N_P(a,b)$ to $N_P(a, b+1)$. We define $\theta : N_P(a,b) \to N_P(a, b+1)$ as follows.

Suppose $\lambda \in N_P(a,b)$, so that $\lambda^{-1}(a) \prec \lambda^{-1}(b)$. If $\lambda^{-1}(b) \prec \lambda^{-1}(b+1)$, then $\lambda^{-1}(a) \prec \lambda^{-1}(b+1)$ by transitivity, so that $\lambda \in N_P(a, b+1)$, and in this case, we set $\theta(\lambda) = \lambda$.

If $\lambda^{-1}(b)$ is incomparable to $\lambda^{-1}(b+1)$ in P, then let μ be the same linear extension as λ with the order of $\lambda^{-1}(b)$ and $\lambda^{-1}(b+1)$ reversed. More precisely,

$$\mu(x) = \begin{cases} x & \text{if } x \in X\backslash\{\lambda^{-1}(b), \lambda^{-1}(b+1)\}; \\ b & \text{if } x = \lambda^{-1}(b+1); \\ b+1 & \text{if } x = \lambda^{-1}(b). \end{cases}$$

Clearly, μ is a linear extension of P. Furthermore, we have that $\mu^{-1}(a) \prec \mu^{-1}(b+1)$, so that $\mu \in N_P(a, b+1)$. In this case, we set $\theta(\lambda) = \mu$.

By breaking $N_P(a,b)$ into those linear extensions in which $\lambda^{-1}(b) \prec \lambda^{-1}(b+1)$ and those in which $\lambda^{-1}(b) \parallel \lambda^{-1}(b+1)$, we see that θ cannot map two different linear extensions in $N_P(a,b)$ onto the same linear extension in $N_P(a, b+1)$. Hence θ is injective, and this completes the proof. □

We are now ready to prove Lemma 4.3.3, which then completes the proof of Theorem 4.3.1.

Lemma 4.3.3 *Let $P = (X, \prec)$ be a poset that is not an antichain, and let λ be a linear extension of P chosen uniformly at random from Λ_P. Then*

$$\mathbb{E}(\mathrm{dist}_P(\lambda)) \geq \frac{1}{3}(n+1).$$

Proof Observe first that

$$\mathrm{dist}_P(\lambda) = \frac{1}{m} \sum_{i,j \in [n]: i<j} (j-i) I_{ij},$$

where I_{ij} is the indicator function of the event that $\lambda^{-1}(i) \prec \lambda^{-1}(j)$. Taking expectations of both sides, we have that

$$\mathbb{E}(\mathrm{dist}_P(\lambda)) = \frac{1}{m} \sum_{i,j \in [n]: i<j} (j-i) \mathbb{P}(\lambda^{-1}(i) \prec \lambda^{-1}(j)).$$

Since, for any fixed linear extension λ of P, we have $\lambda^{-1}(i) \prec \lambda^{-1}(j)$ for exactly m pairs (i,j), then

$$\frac{1}{m} \sum_{i,j \in [n]: i<j} \mathbb{P}(\lambda^{-1}(i) \prec \lambda^{-1}(j)) = 1.$$

(These probabilities are not necessarily equal as they were in the graph version of the problem.) Let $\mathscr{I}$ denote the set of intervals of the form

$[i,j]$, where $i,j \in [n]$ and $i < j$. Let $p_{[i,j]} := \frac{1}{m}\mathbb{P}(\lambda^{-1}(i) \prec \lambda^{-1}(j))$ be the components of a vector $\mathbf{p} \in [0,1]^{\mathscr{I}}$. Now we have that

$$\mathbb{E}(\mathrm{dist}_P(\lambda)) = \frac{1}{m} \sum_{[i,j]\in\mathscr{I}} (j-i)p_{[i,j]} =: \phi(\mathbf{p}).$$

Then $\mathbf{p}$ satisfies the following:

$$p_{[i,j]} \geq 0 \quad \text{for all } [i,j] \in \mathscr{I}, \tag{4.2}$$

$$\sum_{[i,j]\in\mathscr{I}} p_{[i,j]} = 1, \tag{4.3}$$

$$\text{and} \qquad p_{[i,j]} \geq p_{[i',j']} \quad \text{whenever } [i',j'] \subseteq [i,j]. \tag{4.4}$$

The set of inequalities (4.4) is a consequence of Lemma 4.3.2. Let S be the set of vectors in $[0,1]^{\mathscr{I}}$ that satisfy (4.2), (4.3) and (4.4). Then we have that

$$\mathbb{E}(\mathrm{dist}_P(\lambda)) \geq \min_{\mathbf{p}\in S} \phi(\mathbf{p}).$$

Note that ϕ has a minimum in S since ϕ is continuous and S is closed and bounded. Let $\mathbf{p}^* \in S$ be the vector with all its components equal (to $\binom{n}{2}^{-1}$). We make the following claim.

Claim 1 We have

$$\min_{\mathbf{p}\in S} \phi(\mathbf{p}) = \phi(\mathbf{p}^*).$$

Proving this claim proves the lemma, since $\phi(\mathbf{p}^*) = \frac{1}{3}(n+1)$.

Proof Suppose $\mathbf{p} \in S$ and the components of $\mathbf{p}$ are not all equal. We prove the claim by showing that either $\phi(\mathbf{p}) = \phi(\mathbf{p}^*)$ or $\mathbf{p}$ does not minimise ϕ.

Consider the inclusion order $Q = (\mathscr{I}, \subset)$ on $\mathscr{I}$. Thinking of $\mathbf{p}$ as a function from $\mathscr{I}$ to $[0,1]$, we see that $\mathbf{p} \in S$ implies that $\mathbf{p}$ is an increasing function on Q. Consider the vector $\mathbf{p}'$, which is obtained from $\mathbf{p}$ as follows. For $i,j \in [n]$ with $i < j$, let

$$p'_{[i,j]} = \frac{1}{|\mathscr{I}_{j-i}|} \sum_{I\in\mathscr{I}_{j-i}} p_I, \tag{4.5}$$

where $\mathscr{I}_k \subset \mathscr{I}$ is the set of intervals in $\mathscr{I}$ of length k. Thus p'_I is the average of all the components of $\mathbf{p}$ corresponding to intervals of the same length as I. Therefore, we have $\phi(\mathbf{p}') = \phi(\mathbf{p})$.

Next, we show that $\mathbf{p}' \in S$. Clearly $\mathbf{p}'$ satisfies (4.2). Since the components of $\mathbf{p}'$ are averages of components of $\mathbf{p}$, we see that $\mathbf{p}'$ satisfies the inequality (4.3). In order to show that $\mathbf{p}'$ satisfies (4.4), it is sufficient to show that $\mathbf{p}'$ is an increasing function on $\mathscr{I}$, that is, for each $k \in [n-1]$, we must show that

$$\frac{1}{|\mathscr{I}_k|} \sum_{I \in \mathscr{I}_k} p_I \leq \frac{1}{|\mathscr{I}_{k+1}|} \sum_{I \in \mathscr{I}_{k+1}} p_I.$$

Let $[a, a+k]$ be the interval that minimises $p_{[i,j]}$ amongst all intervals $[i, j] \in \mathscr{I}_k$. Consider the following bijection $g : \mathscr{I}_{k+1} \to \mathscr{I}_k \backslash \{[a, a+k]\}$. Let

$$g([i, i+(k+1)]) = \begin{cases} [i, i+k] & \text{if } i < a; \\ [i+1, i+k+1] & \text{if } i \geq a. \end{cases}$$

Now, for each $I \in \mathscr{I}_{k+1}$, we have $g(I) \subset I$ and hence $p_{g(I)} \leq p_I$. Therefore

$$\begin{aligned} \frac{1}{|\mathscr{I}_{k+1}|} \sum_{I \in \mathscr{I}_{k+1}} p_I \geq \frac{1}{|\mathscr{I}_{k+1}|} \sum_{I \in \mathscr{I}_{k+1}} p_{g(I)} &= \frac{1}{|\mathscr{I}_{k+1}|} \sum_{I \in \mathscr{I}_k \backslash \{[a,a+k]\}} p_I \\ &\geq \frac{1}{|\mathscr{I}_k|} \sum_{I \in \mathscr{I}_k} p_I, \end{aligned}$$

where the last inequality follows by our choice of $[a, a+k]$. Thus we have shown that $\mathbf{p}' \in S$.

Now, if all the components of $\mathbf{p}'$ are equal, then $\phi(\mathbf{p}) = \phi(\mathbf{p}') = \phi(\mathbf{p}^*)$. If not, then there is some covering pair of Q, $I_1 \subset I_2$, for which $p'_{I_1} < p'_{I_2}$. Suppose $p'_{I_2} = p'_{I_1} + \epsilon$, where $\epsilon > 0$. Then changing $\mathbf{p}'$ by increasing p'_{I_1} by $\epsilon/2$ and decreasing p'_{I_2} by $\epsilon/2$ to give a vector $\mathbf{p}''$, it is easy to see that $\mathbf{p}'' \in S$ (the sum of the components is unchanged, and $\mathbf{p}''$ is still an increasing function). Furthermore, we have that $\phi(\mathbf{p}'') = \phi(\mathbf{p}') - \frac{\epsilon}{2} < \phi(\mathbf{p}') = \phi(\mathbf{p})$, and thus $\mathbf{p}$ does not minimise ϕ. This completes the proof of Claim 1 and the lemma. $\square$

Chapter 5

Poset Regularity

5.1 Introduction

5.1.1 Regularity for Com(P) and Cov(P)

In Section 1.3, we gave a brief account of the Szemerédi Regularity Lemma. In this chapter, we present a version of the Regularity Lemma tailored for posets. Throughout the chapter, all posets are finite and all graphs are finite, simple, and undirected. Let us begin by recalling the statement of the Regularity Lemma from Section 1.3.

Szemerédi Regularity Lemma For every $\epsilon \in (0,1]$ and every $m \in \mathbb{N}$, there exists a natural number $M = M(\epsilon, m)$ with the following property. For every graph $G = (V, E)$, there exists an $(\epsilon)_G$-regular equipartition of V into k parts, where $m \leq k \leq M$.

When studying a poset P, there are three natural graphs associated with P to which one may wish to apply the Regularity Lemma: these are Com(P), the comparability graph; Inc(P), the incomparability graph; and Cov(P), the covering graph, all introduced in Section 1.2. Since Inc(P) is the complementary graph of Com(P), any regular partition for Com(P) is also a regular partition for Inc(P). We can of course apply the Regularity Lemma for graphs directly to Com(P) or Cov(P), but since these graphs have special properties, we ought to be able to find a regular partition with extra features. This is indeed the case.

Our first two theorems ensure the existence of ordered partitions of P (with a bounded number of parts) that are regular for Com(P) and Cov(P).

One would expect, if nothing else, that *ordered* regular partitions of $\mathrm{Com}(P)$ and $\mathrm{Cov}(P)$ are more convenient to work with than just regular partitions of these graphs.

Before we state the theorems, we need to adapt and extend the definition of ordered partitions in the following way to avoid clumsy exposition. Let $P = (X, \prec)$ a poset, and let $\mathcal{P}$ be a partition of $Y \subseteq X$. We say that $\mathcal{P}$ is a *P-ordered* partition of Y if the parts of $\mathcal{P}$ can be ordered $Y_1, \ldots, Y_k$ such that, whenever $a \in Y_i$ and $b \in Y_j$ with $a \prec b$, we have $i \leq j$. This is the same definition as given in Section 1.2 except for three minor differences: first, we have extended the definition to partitions where no explicit ordering is given on the parts; second, we have extended the definition to include partitions of subsets of X; third, we use the term *P-ordered* rather than *ordered partition of P*.

We can now state our first two theorems.

Theorem 5.1.1 *For every $\epsilon \in (0, 1]$ and every $m \in \mathbb{N}$, there exists a natural number $M_1 = M_1(\epsilon, m)$ with the following property. For every poset $P = (X, \prec)$, there exists an equipartition of X into k parts that is P-ordered and $(\epsilon)_{\mathrm{Com}(P)}$-regular, where $m \leq k \leq M_1$.*

Theorem 5.1.2 *For every $\epsilon \in (0, 1]$ and every $m \in \mathbb{N}$, there exists a natural number $M_2 = M_2(\epsilon, m)$ with the following property. For every poset $P = (X, \prec)$, there exists an equipartition of X into k parts that is P-ordered and $(\epsilon)_{\mathrm{Cov}(P)}$-regular, where $m \leq k \leq M_2$.*

Having ordered partitions that are regular enables us to derive certain additional properties about these partitions. We postpone this discussion for Section 5.1.3. Next, we discuss a generalisation of the above two theorems.

5.1.2 Regular Graph Partitions Respecting an Order

Theorems 5.1.1 and 5.1.2 are important special cases of a more general theorem that we shall prove. The theorem requires a rather intricate definition for its statement. The definition describes precisely how a graph $G = (X, E)$ and a poset $P = (X, \prec)$ should be related in order to guarantee the existence of a P-ordered regular equipartition using our method.

Let $G = (V, E)$ be a graph, and let $A, B \subseteq V$ be disjoint, where (A, B) is not $(\epsilon, \delta)_G$-regular. Thus, there exists $A' \subseteq A$ and $B' \subseteq B$ with $|A'| > \delta|A|$

and $|B'| > \delta|B|$, where

$$|d_G(A,B) - d_G(A',B')| > \epsilon.$$

We call (A', B') a *witness* to the $(\epsilon, \delta)_G$-irregularity of (A, B).

Let $P = (X, \prec)$ be a poset with $Y \subseteq X$, and let $\mathcal{P}$ be a P-ordered partition of Y. For $Y' \subseteq Y$, we say Y' is *$\mathcal{P}$-unifiable* if Y' is the union of some of the parts of $\mathcal{P}$.

Let $G = (X, E)$ be a graph and let $P = (X, \prec)$ be a poset. The definition below describes a sufficient condition for the existence of a P-ordered regular partition of X into a bounded number of parts. Roughly speaking, this condition says the following: given Y_1, Y_2, any P-ordered equipartition of $Y \subseteq X$, if (Y_1, Y_2) is irregular (for G), then Y_1 and Y_2 have P-ordered partitions into a *small* number of parts such that the union of some of the parts form witnesses to some *weaker* irregularity.

We now state the key definition precisely along with the main theorem.

Definition 5.1.3 *Let $G = (X, E)$ be a graph; let $P = (X, \prec)$ a poset; and fix constants $l \in \mathbb{N}$ and $\epsilon, \epsilon' \in (0, 1]$ with $\epsilon' \leq \epsilon$. We say that G is* $(P, \epsilon, \epsilon', l)$-good *if the following holds. For every $Y \subseteq X$ and every P-ordered equipartition Y_1, Y_2 of Y, if (Y_1, Y_2) is not $(\epsilon, \epsilon)_G$-regular, then there exist P-ordered partitions $\mathcal{P}_1$ of Y_1 and $\mathcal{P}_2$ of Y_2 such that*

1. *there exists a $\mathcal{P}_1$-unifiable set $\bar{Y_1} \subseteq Y_1$ and a $\mathcal{P}_2$-unifiable $\bar{Y_2} \subseteq Y_2$ such that $(\bar{Y_1}, \bar{Y_2})$ is witness to the $(\epsilon', \epsilon')_G$-irregularity of (Y_1, Y_2), and*

2. *the partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ each have at most l parts.*

We discuss the implications of this definition after giving our generalised regularity result below.

Theorem 5.1.4 *For every $\epsilon, \epsilon' \in (0, 1]$ satisfying $\epsilon' \leq \epsilon$ and every $l, m \in \mathbb{N}$, there exists a natural number $M = M(\epsilon, \epsilon', l, m)$ with the following property. For every graph $G = (X, E)$ and every poset $P = (X, \prec)$, if G is $(P, \epsilon, \epsilon', l)$-good, then there exists an equipartition of X into k parts that is P-ordered and $(\epsilon)_G$-regular, where $m \leq k \leq M$.*

Let us examine when Theorem 5.1.4 is of value. In almost all circumstances, the usual Regularity Lemma is applied not to a single graph, but to an infinite class of graphs. Its power lies in giving us an ϵ-regular equipartition with a bounded number of parts, where the bound depends on ϵ only. Similarly, in practice, we would expect to apply Theorem 5.1.4 to an infinite class of pairs (G, P), where $G = (X, E)$ is a graph and $P = (X, \prec)$ is a poset. Furthermore, the P-ordered ϵ-regular partitions that we seek should have a bounded number of parts, where the bound depends on ϵ only. Notice that the number of parts in the partition given by Theorem 5.1.4 is bounded by M, which depends on ϵ' and l as well as ϵ. Thus, we seek classes of pairs (G, P) for which G is $(P, \epsilon, \epsilon', l)$-good and where ϵ' and l are functions of ϵ. More precisely, we have the following.

Let

$$\mathcal{D} = \{(G, P) : G = (X, E) \text{ a graph, } P = (X, \prec) \text{ a poset}\},$$

and let $\mathcal{C} \subseteq \mathcal{D}$.

Definition 5.1.5 *We say that $\mathcal{C}$ is a* good *class if there exist positive-valued functions ϵ' and l of $\epsilon \in (0,1]$ with the following property. For every $\epsilon \in (0,1]$, we have $\epsilon'(\epsilon) \leq \epsilon$, and for every $(G, P) \in \mathcal{C}$, we have that G is $(P, \epsilon, \epsilon'(\epsilon), l(\epsilon))$-good.*

Let $\mathcal{C}$ be a good class as described above. By Theorem 5.1.4, for every pair $(G, P) \in \mathcal{C}$ and every every $\epsilon \in (0,1]$, there exists a P-ordered $(\epsilon)_G$-regular equipartition of X, where the number of parts is bounded by a function of ϵ only.

Let us give a few examples of good classes.

Examples

1. Let $\mathcal{C}_0 = \{(G, P) : G = (X, E),\ P = (X, \prec) \text{ an antichain}\}$. It is easy to see that for every $\epsilon \in (0,1]$ and every $(G, P) \in \mathcal{C}_0$, we have that G is $(P, \epsilon, \epsilon, 2)$-good. Thus $\mathcal{C}_0$ is good, and applying Theorem 5.1.4 to $\mathcal{C}_0$ yields the usual Regularity Lemma for graphs.

2. Let $\mathcal{C}_1 = \{(\mathrm{Com}(P), P) : P \text{ a poset}\}$. We shall prove that for every $\epsilon \in (0,1]$ and every poset P, $\mathrm{Com}(P)$ is $(P, \epsilon, \epsilon, 2)$-good. Thus, $\mathcal{C}_1$ is good, and applying Theorem 5.1.4 to $\mathcal{C}_1$ yields Theorem 5.1.1.

3. Let $\mathcal{C}_2 = \{(\mathrm{Cov}(P), P) : P \text{ a poset}\}$. We shall prove that for every $\epsilon \in (0,1]$ and every poset P, $\mathrm{Cov}(P)$ is $(P, \epsilon, \frac{\epsilon^2}{2}, \lceil 8\epsilon^{-3} + 1 \rceil)$-good. Thus, $\mathcal{C}_2$ is good, and applying Theorem 5.1.4 to $\mathcal{C}_2$ yields Theorem 5.1.2.

Here is an example of a class that is not good.

Example For n an even positive integer, consider the graph $G_n = (A_n \cup B_n, E)$, where $A_n = \{a_1, \dots, a_n\}$, $B_n = \{b_1, \dots, b_n\}$, and

$$E = \{a_i b_j : i, j \text{ both even}\}.$$

Let $L_n = (A_n \cup B_n, \prec)$, where $\prec$ is the total order given by $a_1 \prec a_2 \prec \cdots \prec a_n \prec b_1 \prec b_2 \prec \cdots \prec b_n$. We claim that the class

$$\mathcal{C}_3 = \{(G_n, P_n) : n \in \mathbb{N}\}$$

is not a good class. We show that if G_n is $(L_n, \frac{1}{2}, \epsilon', l)$-good for some $0 < \epsilon' < \frac{1}{2}$, then $l > \epsilon'^2 n$. This then shows that $\mathcal{C}_3$ is not a good class since l is forced to depend on the graph under consideration.

Suppose G_n is $(L_n, \frac{1}{2}, \epsilon', l)$-good. We apply Definition 5.1.3 to this example, where we set $Y = A_n \cup B_n$, $Y_1 = A_n$, and $Y_2 = B_n$. Observe that Y_1, Y_2 is an L_n-ordered equipartition of G_n and that $d_{G_n}(Y_1, Y_2) = \frac{1}{4}$. Furthermore, the pair (Y_1, Y_2) is not $(\frac{1}{2}, \frac{1}{2})_{G_n}$-regular; indeed, (A_e, B_e) is a witness to this, where A_e (resp. B_e) is the set of even vertices of $A_n = Y_1$ (resp. $B_n = Y_2$).

It remains for us to show that if $(\bar{Y_1}, \bar{Y_2})$ is a witness to the (ϵ', ϵ')-irregularity of (Y_1, Y_2) and $\bar{Y_1}$ (resp. $\bar{Y_2}$) is $\mathcal{P}_1$ (resp. $\mathcal{P}_2$)-unifiable, where $\mathcal{P}_1$ (resp. $\mathcal{P}_2$) is an L_n-ordered partition of Y_1 (resp. Y_2), then one of $\mathcal{P}_1, \mathcal{P}_2$ has at least $\epsilon'^2 n$ parts.

Informally, the reason for this is that if $d(\bar{Y_1}, \bar{Y_2}) > \frac{1}{4} + \epsilon'$, then $\bar{Y_1}$ and/or $\bar{Y_2}$ must contain significantly more even vertices than odd vertices, and if $d(\bar{Y_1}, \bar{Y_2}) < \frac{1}{4} - \epsilon'$, then $\bar{Y_1}$ and/or $\bar{Y_2}$ must contain significantly more odd vertices than even vertices. However, in any part of $\mathcal{P}_1$ or $\mathcal{P}_2$, the number of even and odd vertices differ by at most 1, hence for $\bar{Y_1}$ (resp. $\bar{Y_2}$) to be $\mathcal{P}_1$ (resp. $\mathcal{P}_2$)-unifiable, $\mathcal{P}_1$ and $\mathcal{P}_2$ must contain a large number of parts.

Let $|\bar{Y_1} \cap A_e| = \delta_1 |\bar{Y_1}|$ and $|\bar{Y_2} \cap B_e| = \delta_2 |\bar{Y_2}|$. Thus, the number of even vertices of $\bar{Y_1}$ differs from the number of odd vertices of $\bar{Y_1}$ by $|1 - 2\delta_1||\bar{Y_1}|$.

Therefore, $|\mathcal{P}_1| \geq |1-2\delta_1||\bar{Y_1}| > |1-2\delta_1|\epsilon' n$ and similarly for $|\mathcal{P}_2|$. Since

$$\epsilon' < \left|d(\bar{Y_1},\bar{Y_2}) - \frac{1}{4}\right| = \left|\delta_1\delta_2 - \frac{1}{4}\right| = \frac{1}{4}|(2\delta_1)(2\delta_2) - 2\delta_1 + 2\delta_1 - 1|$$
$$\leq \frac{1}{4}\Big(2\delta_1|2\delta_2 - 1| + |2\delta_1 - 1|\Big),$$

we have that either $|2\delta_1 - 1| > \epsilon'$ or $|2\delta_2 - 1| > \epsilon'$. Thus either $|\mathcal{P}_1| > \epsilon'^2 n$ or $|\mathcal{P}_2| > \epsilon'^2 n$.

5.1.3 Simultaneous Ordered Regular Partitions and Further Properties

Given the proof of the usual Regularity Lemma for graphs, it is not a big step to apply the result to several graphs simultaneously. Similarly Theorem 5.1.6, given below, says that we can apply Theorems 5.1.1 and 5.1.2 simultaneously to a given poset.

Theorem 5.1.6 *For every $\epsilon \in (0,1]$ and every $m \in \mathbb{N}$, there exists a natural number $M_3 = M_3(\epsilon, m)$ with the following property. For every poset $P = (X, \prec)$, there exists an equipartition of X into k parts that is P-ordered, $(\epsilon)_{\mathrm{Com}(P)}$-regular, and $(\epsilon)_{\mathrm{Cov}(P)}$-regular where $m \leq k \leq M_3$.*

Once we have established the existence of ordered equipartitions that are regular both for the comparability graph and the covering graph of a poset, we can then derive some straightforward properties of such partitions. These are summarised in the following theorem.

Theorem 5.1.7 *Given a poset $P = (X, \prec)$, let $X_1, \ldots, X_k$ be a P-ordered equipartition of X that is $(\epsilon)_{\mathrm{Com}(P)}$-regular and $(\epsilon)_{\mathrm{Cov}(P)}$-regular. Then we have the following.*

(i) If $1 \leq r < s < t \leq k$ with (X_r, X_s) and (X_s, X_t) both $(\epsilon,\epsilon)_{\mathrm{Com}(P)}$-regular pairs, and $d_{\mathrm{Com}(P)}(X_r, X_s), d_{\mathrm{Com}(P)}(X_s, X_t) > 2\epsilon$, then we have $d_{\mathrm{Com}(P)}(X_r, X_t) > 1 - 2\epsilon$.

(ii) If $1 \leq r < s \leq k$ with (X_r, X_s) an $(\epsilon,\epsilon)_{\mathrm{Com}(P)}$-regular pair, and $\epsilon < d_{\mathrm{Com}(P)}(X_r, X_s) < 1 - 2\epsilon$, then we have $e_{\mathrm{Com}(P)}(X_r) < 2\epsilon|X_r|^2$ and $e_{\mathrm{Com}(P)}(X_s) < 2\epsilon|X_s|^2$.

(iii) If $1 \leq r < s \leq k$ *with* (X_r, X_s) *being both* $(\epsilon, \epsilon)_{\mathrm{Com}(P)}$*-regular and* $(\epsilon, \epsilon)_{\mathrm{Cov}(P)}$*-regular, and* $d_{\mathrm{Cov}(P)}(X_r, X_s) > 3\epsilon$*, then there exist antichains,* $A_r \subseteq X_r$ *and* $A_s \subseteq X_s$*, such that* $|A_r| \geq (1 - 3\epsilon)|X_r|$ *and* $|A_s| \geq (1 - 3\epsilon)|X_s|$.

The first property in the above theorem is a simple consequence of transitivity. The second and third properties give us information about what happens inside some of the parts of our regular ordered equipartition. In applications of the usual Regularity Lemma, we have no control over what happens inside the parts of our regular partition; this lack of control contributes to a small inevitable error term. However, we hope that properties (ii) and (iii) of Theorem 5.1.7 might enable us to reduce or even eliminate such errors, and thereby give a novel application of the Regularity Lemma to posets.

Unfortunately, we have been unable to find a genuine application of this result, that is, an application that relies on our result but that cannot easily be proved from the usual Regularity Lemma. Let us give an example to illustrate one reason why finding an application has proven to be difficult.

Example Let $Q_n = ([n]^2, \prec)$ be the poset on the $n \times n$ lattice of positive integers, where $(x, y) \prec (a, b)$ if and only if $x < a$ and $y < b$. (It is clear that this is a poset since $<$ is irreflexive and transitive.) Let k be a positive integer (we assume k divides n for convenience). We partition $[n]^2$ into k^2 smaller lattices: for $i, j \in [k]$, let $X_{ij} = ((i-1)k, ik] \times ((j-1)k, jk]$ (here $(r, s]$ denotes the set of integers strictly larger than r and at most s).

The X_{ij} form a Q_n-ordered partition of $[n]^2$. Indeed, if $a \in X_{ij}$, $b \in X_{pq}$ and $a \prec b$, then we must have $i \leq p$ and $j \leq q$. Then, ordering the X_{ij} lexicographically, that is, $X_{11}, \ldots, X_{1k}, X_{21}, \ldots, X_{2k}, \ldots, X_{k1}, \ldots, X_{kk}$, demonstrates that the X_{ij} form an Q_n-ordered partition.

Observe that the density in $\mathrm{Com}(Q_n)$ and $\mathrm{Cov}(Q_n)$ between most pairs of parts in our partition is either 0 or 1, and hence, such pairs are regular. We have

$$d_{\mathrm{Com}(Q_n)}(X_{ij}, X_{pq}) = \begin{cases} 1 & \text{if} \quad i < p, j < q \quad \text{or} \quad i > p, j > q; \\ 0 & \text{if} \quad i < p, j > q \quad \text{or} \quad i > p, j < q. \end{cases}$$

Thus, as long as $i \neq p$ and $j \neq q$, the pair (X_{ij}, X_{pq}) is $(\epsilon, \epsilon)_{\mathrm{Com}(Q_n)}$-regular for all $\epsilon > 0$. Therefore, all but at most $k^2 \cdot 2(k-1)$ of the $\binom{k^2}{2}$ pairs of parts are not regular. Thus, our partition is $(\epsilon)_{\mathrm{Com}(Q_n)}$-regular as long as we choose k such that

$$\frac{2k^2(k-1)}{\binom{k^2}{2}} = \frac{4}{k+1} < \epsilon.$$

Similarly for $\mathrm{Cov}(Q_n)$, if $p \neq i-1, i, i+1$ and $q \neq j-1, j, j+1$ then $d_{\mathrm{Cov}(Q_n)}(X_{ij}, X_{pq}) = 0$, and so, such a pair (X_{ij}, X_{pq}) is $(\epsilon, \epsilon)_{\mathrm{Cov}(Q_n)}$-regular for all $\epsilon > 0$. Therefore, all but at most $k^2 \cdot (9k-10)$ of $\binom{k^2}{2}$ pairs of parts are not regular in $\mathrm{Cov}(P)$. Thus, our partition is $(\epsilon)_{\mathrm{Cov}(Q_n)}$-regular as long as we choose k such that

$$\frac{k^2(9k-10)}{\binom{k^2}{2}} < \frac{9k^2(k-1)}{\binom{k^2}{2}} = \frac{18}{k+1} < \epsilon.$$

Here we have produced regular ordered partitions of Q_n, but where each part induces a copy of $Q_{n/k}$. Far from being sparse, the poset induced by each part is essentially a copy of the whole poset. Thus, in this instance, parts (ii) and (iii) of Theorem 5.1.7 give little information. The reason for this is that the densities between regular parts are not bounded away from 0 and 1. Such a problem occurs in many other examples.

The rest of the chapter is arranged as follows. In Section 5.2 we prove Theorem 5.1.4. In Section 5.3, we prove that the classes $\mathcal{C}_1$ and $\mathcal{C}_2$ are good, and from this we deduce Theorems 5.1.1 and 5.1.2. In Section 5.4, we prove Theorems 5.1.6 and 5.1.7.

5.2 General Ordered Regular Partitions

In this section, we prove our general regularity result, Theorem 5.1.4. We give a self-contained and detailed treatment of the proof for those readers unfamiliar with the Regularity Lemma. As one would expect, the proof of Theorem 5.1.4 is similar to the proof of the usual Regularity Lemma; we try to highlight where the proofs differ. We follow [61] in our treatment of the standard aspects in our proof.

As with the standard proof of the Regularity Lemma, we require some preliminary results. We begin with some definitions.

Let $G = (V, E)$ be a graph and let $\mathcal{P}$ be a partition of V into k parts,

$V_1, \ldots, V_k$. We define the *partition index* to be

$$\text{ind}_G(\mathcal{P}) = \frac{1}{k^2} \sum_{1 \le i < j \le k} d_G^2(V_i, V_j).$$

Note that for every graph G and every partition $\mathcal{P}$ of its vertices, we have $\text{ind}_G(\mathcal{P}) < \frac{1}{2}$.

The usual Regularity Lemma is proved as follows. Suppose $G = (V, E)$ is a graph with an equipartition $\mathcal{P}$ of V into parts $V_1, \ldots, V_k$. We show that if $\mathcal{P}$ is not $(\epsilon)_G$-regular, then we can refine $\mathcal{P}$ by dividing each of $V_1, \ldots, V_k$ into a fixed number of smaller parts to obtain a new equipartition $\mathcal{Q}$. If we have divided $V_1, \ldots, V_k$ suitably, then we find that

$$\text{ind}_G(\mathcal{Q}) \ge \text{ind}_G(\mathcal{P}) + c,$$

where c is a constant that depends on ϵ only. As long as our equipartition is not ϵ-regular, we can keep refining it in this way. However, after a bounded number ($\le \lceil 1/2c \rceil$) of refinements, we obtain an equipartition with an index of at least $1/2$, which is a contradiction. Therefore, at some point before this happens, we must obtain an equipartition that is ϵ-regular, and this gives us a bound (depending on ϵ only) on the number of parts in such an equipartition.

Theorem 5.1.4 is proved in the same way, but with one difference. We start with an <u>ordered</u> equipartition $\mathcal{P}$, and as with graphs, if $\mathcal{P}$ is not ϵ-regular, then we try to find a refinement $\mathcal{Q}$ that is <u>ordered</u> and has a higher index.

Our first lemma shows that, given an equipartition, if we refine it arbitrarily, then its index may decrease, but not by much.

Lemma 5.2.1 *Let $G = (V, E)$ be a graph with $|V| = n$, and let $\mathcal{P}$ be an equipartition of G into k parts $V_1, \ldots, V_k$. Fix $q \in \mathbb{N}$. For each i, let $V_{i1}, \ldots, V_{iq}$ be an equipartition of V_i, and let $\mathcal{Q}$ be the resulting equipartition of V given by the V_{ix}. Then*

$$\text{ind}_G(\mathcal{Q}) \ge \text{ind}_G(\mathcal{P}) - \frac{2}{m+1},$$

where $m = \lfloor \frac{n}{kq} \rfloor$.

Proof We drop the subscript G . Since $\mathcal{P}$ and $\mathcal{Q}$ are equipartitions, it is clear that $m \leq |V_{ix}| \leq m+1$, and hence

$$|V_{ix}| \leq \frac{1}{q}\Big(\frac{m+1}{m}\Big)|V_i| \tag{5.1}$$

for all i, x. Fix i and j. Then since

$$\sum_{1 \leq x,y \leq q} e(V_{ix}, V_{jy}) = e(V_i, V_j),$$

using (5.1), we obtain

$$\begin{aligned}\frac{1}{q^2}\sum_{1\leq x,y\leq q} d(V_{ix}, V_{jy}) = \sum_{1\leq x,y\leq q}\frac{e(V_{ix}, V_{jy})}{q|V_{ix}|q|V_{jy}|} &\geq \Big(\frac{m}{m+1}\Big)^2 \sum_{1\leq x,y\leq q}\frac{e(V_{ix}, V_{jy})}{|V_i||V_j|} \\ &= \Big(\frac{m}{m+1}\Big)^2 d(V_i, V_j). \end{aligned} \tag{5.2}$$

Using (5.2) together with the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \operatorname{ind}(\mathcal{Q}) &\geq \frac{1}{k^2}\sum_{1\leq i<j\leq k}\frac{1}{q^2}\sum_{1\leq x,y\leq q} d^2(V_{ix}, V_{jy}) \\ &\geq \frac{1}{k^2}\sum_{1\leq i<j\leq k}\Big(\frac{1}{q^2}\sum_{1\leq x,y\leq q} d(V_{ix}, V_{jy})\Big)^2 && \text{(using Cauchy-Schwarz)} \\ &\geq \Big(\frac{m}{m+1}\Big)^4 \operatorname{ind}(\mathcal{P}) && \text{(using (5.2)).} \end{aligned}$$

Note that $(\frac{m}{m+1})^4 = (1 - \frac{1}{m+1})^4 \geq 1 - \frac{4}{m+1}$, from which we obtain

$$\operatorname{ind}(\mathcal{Q}) \geq \Big(1 - \frac{4}{m+1}\Big)\operatorname{ind}(\mathcal{P}) \geq \operatorname{ind}(\mathcal{P}) - \frac{2}{m+1}.$$

□

In order to find a refinement of an irregular partition that increases the index, we must somehow take advantage of the non-uniformity in density between some of the parts. In order to do this, we shall require a defect form of the Cauchy-Schwarz inequality. Some form of the Cauchy-Schwarz inequality is at the heart of any standard proof of the Regularity Lemma.

Lemma 5.2.2 **([61])** *Let $x_1, \ldots, x_n$ be non-negative real numbers, and let*

$m < n$. *Then*

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 \geq \Phi^2 + \frac{m}{n}(\Phi - \phi)^2,$$

where

$$\Phi = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{and} \quad \phi = \frac{1}{m}\sum_{i=1}^{m} x_i.$$

Proof The proof is a simple application of the usual Cauchy-Schwarz inequality.

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} x_i^2 &= \frac{1}{n}\left(\sum_{i=1}^{m} x_i^2 + \sum_{i=m+1}^{n} x_i^2\right) \\
&= \frac{m}{n}\left(\frac{1}{m}\sum_{i=1}^{m} x_i^2\right) + \frac{n-m}{n}\left(\frac{1}{n-m}\sum_{i=m+1}^{n} x_i^2\right) \\
&\geq \frac{m}{n}\left(\frac{1}{m}\sum_{i=1}^{m} x_i\right)^2 + \frac{n-m}{n}\left(\frac{1}{n-m}\sum_{i=m+1}^{n} x_i\right)^2 \text{(Cauchy-Schwarz)} \\
&= \frac{m}{n}\phi^2 + \frac{1}{n(n-m)}(n\Phi - m\phi)^2 \\
&= \Phi^2 + \frac{m}{n-m}(\Phi - \phi)^2 \\
&\geq \Phi^2 + \frac{m}{n}(\Phi - \phi)^2.
\end{aligned}$$

□

The next lemma is a technical one used for estimating the density of large induced subgraphs of bipartite graphs.

Lemma 5.2.3 ([61]) *Let* $G = (V, E)$ *be a graph and let* $V_1, V_2 \subseteq V$ *be disjoint. If* $U_1 \subseteq V_1$ *and* $U_2 \subseteq V_2$ *such that* $|U_1| \geq (1-\delta)|V_1|$ *and* $|U_2| \geq (1-\delta)|V_2|$, *then*

$$|d_G(V_1, V_2) - d_G(U_1, U_2)| \leq 2\delta.$$

Proof Again, we drop G as a subscript. We have that

$$\begin{aligned}
d(V_1, V_2) = \frac{e(V_1, V_2)}{|V_1||V_2|} \geq \frac{e(U_1, U_2)}{|U_1||U_2|}(1-\delta)^2 &\geq d(U_1, U_2)(1 - 2\delta) \\
&\geq d(U_1, U_2) - 2\delta.
\end{aligned}$$

By considering the complementary graph of G, we get

$$1 - d(V_1, V_2) \geq 1 - d(U_1, U_2) - 2\delta.$$

Together the inequalities give

$$|d(V_1, V_2) - d(U_1, U_2)| \leq 2\delta,$$

as required. $\square$

We now turn our attention to posets and ordered partitions. Let $P = (X, \prec)$ be a poset with $Y \subseteq X$. Suppose we have r ordered partitions of Y, which we call $\mathcal{P}_1, \ldots, \mathcal{P}_r$. Denote the parts of $\mathcal{P}_i$ by $Y_i(1), \ldots, Y_i(s_i)$. We define $\prod_{i=1}^{r} \mathcal{P}_i$, which we call the *product partition*, to be the partition with parts $Y(k_1, \ldots, k_r)$ where

$$Y(\mathbf{k}) = Y(k_1, \ldots, k_r) = \bigcap_{i=1}^{r} Y_i(k_i).$$

Here $\mathbf{k} = (k_1, \ldots, k_r) \in \prod_{i=1}^{r} [s_i]$.

As we have mentioned, Theorem 5.1.4 is proved by repeatedly refining partitions that are not ϵ-regular in a suitable manner until we obtain a partition that is ϵ-regular. However, we must maintain an ordered partition at each refinement. As we shall see, these refinements are based on a product partition. We ensure that our refinements remain ordered by proving, in the next lemma, that the product of ordered partitions is itself an ordered partition.

Lemma 5.2.4 *Let $P = (X, \prec)$ be a poset with $Y \subseteq X$ and let $\mathcal{P}_1, \ldots, \mathcal{P}_r$ be P-ordered partitions of Y. Then $\mathcal{P}^* = \prod_{i=1}^{r} \mathcal{P}_i$ is also a P-ordered partition of Y.*

Proof As before, let $Y_i(1), \ldots, Y_i(s_i)$ be the ordering of the parts of $\mathcal{P}_i$, and let $Y(\mathbf{k})$ give the parts of $\mathcal{P}^*$ for $\mathbf{k} \in \prod_{i=1}^{r} [s_i]$. We wish to find a linear ordering on the parts of $\mathcal{P}^*$ that demonstrates it is a P-ordered partition.

We partially order the parts of $\mathcal{P}^*$ by $\prec^*$ as follows. We let $Y(\mathbf{k}') \prec^* Y(\mathbf{k})$ if and only if $k_i' \leq k_i$ for all $i \in [r]$ and $k_j' < k_j$ for some $j \in [r]$. We show that $\prec^*$ respects $\prec$, that is, we show that if $a, b \in Y$ where $a \prec b$ with $a \in Y(\mathbf{u})$ and $b \in Y(\mathbf{v})$, then $Y(\mathbf{u}) \preceq^* Y(\mathbf{v})$. Once we have shown this, then any linear extension of $\prec^*$ gives a linear ordering on the parts of $\mathcal{P}^*$ that demonstrates $\mathcal{P}^*$ is a P-ordered partition.

It remains for us to show that $\prec^*$ respects $\prec$. Suppose $a, b \in Y$ where $a \prec b$ with $a \in Y(\mathbf{u})$ and $b \in Y(\mathbf{v})$. Looking at the ith components of $\mathbf{u}$ and $\mathbf{v}$, we have that $a \in Y_i(u_i)$ and $b \in Y_i(v_i)$. Since $Y_i(1), \ldots, Y_i(s_i)$ is a P-ordered partition, we have that $u_i \leq v_i$, and this is true for all i. Therefore $Y(\mathbf{u}) \preceq^* Y(\mathbf{v})$, as required. □

Our final preliminary result is a simple technical lemma. Here first is some notation.

Suppose $P = (X, \prec)$ is a poset and $Y_1, \ldots, Y_k$ is a P-ordered partition of $Y \subseteq X$. A linear extension, $L = (Y, \prec_L)$, of $P_Y = (Y, \prec)$ is said to *extend* the ordered partition $Y_1, \ldots, Y_k$ if, whenever $a \in Y_i$, $b \in Y_j$, and $i < j$, then $a \prec_L b$. Such linear extensions exist since we can simply concatenate linear extensions of each Y_i.

Lemma 5.2.5 *Let $P = (X, \prec)$ be a poset, let $Y_1, \ldots, Y_k$ be a P-ordered partition of $Y \subseteq X$, and fix $l \in \mathbb{N}$. Then there exists a P-ordered equipartition $Y_1', \ldots, Y_l'$ of Y such that for each i, there exists $i_1, \ldots, i_r$ where*

$$\bigcup_{j=1}^{r} Y'_{i_j} \subseteq Y_i \quad \textit{and} \quad \left|\bigcup_{j=1}^{r} Y'_{i_j}\right| \geq |Y_i| - 2\left\lceil \frac{|Y|}{l} \right\rceil.$$

Proof Let $L = (Y, \prec_L)$ extend the P-ordered partition $Y_1, \ldots, Y_k$. Use the order of L to partition the elements of Y into sets $Y_1', \ldots, Y_l'$, where Y_1' contains the first $\lfloor |Y|/l \rfloor$ elements, Y_2' contains the next $\lfloor (|Y|+1)/l \rfloor$ elements, Y_3' contains the next $\lfloor (|Y|+2)/l \rfloor$ elements, and so on. This is an L-ordered equipartition of Y, and hence, it is also a P-ordered equipartition of Y since $\prec_L$ is an extension of $\prec$.

Fix i and observe that each Y_j' (except at most two) is contained in or disjoint from Y_i. Let $Y'_{i_1}, \ldots, Y'_{i_r}$ be those Y_j' contained in Y_i and let Y'_{i_+} and Y'_{i_-} be the two exceptional (possibly empty) sets mentioned above. Then we have

$$|Y_i| - \left|\bigcup_{j=1}^{r} Y'_{i_j}\right| \leq |Y'_{i_+}| + |Y'_{i_-}| \leq 2\left\lceil \frac{|Y|}{l} \right\rceil,$$

as required. □

We are now ready to prove the key step in Theorem 5.1.4, that is, the existence of refinements of irregular partitions with larger index.

Lemma 5.2.6 *Fix $\epsilon, \epsilon' \in (0,1]$ and $l \in \mathbb{N}$. Let $P = (X, \prec)$ be a poset and let $G = (X, E)$ be a graph such that G is $(P, \epsilon, \epsilon', l)$-good. Suppose $X_1, \ldots, X_k$ is a P-ordered equipartition (which we denote by $\mathcal{P}$) that is not $(\epsilon)_G$-regular, where $k \geq 2$. If*

$$n := |X| \geq \frac{400kq}{\epsilon'^5} \quad \text{and} \quad q \geq \frac{800l^{k-1}}{\epsilon'^2},$$

then there exists an ordered equipartition $\mathcal{Q}$ with kq parts satisfying

$$\mathrm{ind}_G(\mathcal{Q}) \geq \mathrm{ind}_G(\mathcal{P}) + \frac{\epsilon'^5}{100}.$$

Proof Set $m = \lfloor n/(kq) \rfloor$ and note that $m \geq 1$. We begin by describing $\mathcal{Q}$.

Recall that since G is $(P, \epsilon, \epsilon', l)$-good, then, whenever (X_i, X_j) is not $(\epsilon, \epsilon)_G$-regular, there exist ordered partitions of X_i and X_j (which we denote by $\mathcal{P}_{ij}$ and $\mathcal{P}_{ji}$ respectively) into at most l parts with the following property. There exists $X_{ij} \subseteq X_i$, which is the union of some parts of $\mathcal{P}_{ij}$, and $X_{ji} \subseteq X_j$, which is the union of some parts of $\mathcal{P}_{ji}$, where (X_{ij}, X_{ji}) is witness to the $(\epsilon', \epsilon')_G$-irregularity of (X_i, X_j). If (X_i, X_j) is $(\epsilon, \epsilon)_G$-regular, then let $\mathcal{P}_{ij}$ and $\mathcal{P}_{ji}$ be the trivial partitions of X_i and X_j respectively, into one part.

Thus, keeping i fixed and varying j gives $k-1$ P-ordered partitions $\mathcal{P}_{ij}$ of X_i (some of which may be trivial). We know by Lemma 5.2.4 that

$$\mathcal{P}_i := \prod_{\substack{j=1 \\ j \neq i}}^{k} \mathcal{P}_{ij}$$

is a P-ordered partition of X_i. Note that $\mathcal{P}_i$ has at most l^{k-1} parts and that each X_{ij} is the union of some of these parts.

We divide each X_i into q parts as follows. By Lemma 5.2.5, there exists a P-ordered equipartition $Z_{i1}, \ldots, Z_{iq}$ of X_i such that, given any part A of $\mathcal{P}_i$, if A^* is the largest subset of A that can be written as the union of some of the Z_{ix}, then

$$|A^*| \geq |A| - 2\left\lceil \frac{n}{kq} \right\rceil = |A| - 2(m+1). \tag{5.3}$$

Note that $m \leq |Z_{ix}| \leq m+1$.

Now for each X_i, we have defined a P-ordered equipartition $Z_{i1}, \dots, Z_{iq}$. Concatenating these ordered equipartitions over i gives a P-ordered equipartition of X, which is a refinement of $\mathcal{P}$. This is the P-ordered equipartition $\mathcal{Q}$ described in the statement of the lemma.

It remains for us to estimate $\text{ind}_G(\mathcal{Q})$ in terms of $\text{ind}_G(\mathcal{P})$. This part of the proof follows closely the proof of the usual Regularity Lemma. It is somewhat involved, but standard. We drop G as a subscript.

Fix i and j. We start by estimating

$$\frac{1}{q^2} \sum_{1 \leq x,y \leq q} d(Z_{ix}, Z_{jy})$$

in terms of $d(X_i, X_j)$. We proceed as with the proof of Lemma 5.2.1. We observe that

$$\sum_{1 \leq x,y \leq q} e(Z_{ix}, Z_{jy}) = e(X_i, X_j) \quad \text{and} \quad q\frac{m}{m+1}|Z_{ix}| \leq |X_i|.$$

Therefore

$$\begin{aligned} \frac{1}{q^2} \sum_{1 \leq x,y \leq q} d(Z_{ix}, Z_{jy}) &= \sum_{1 \leq x,y \leq q} \frac{e(Z_{ix}, Z_{jy})}{q|Z_{ix}|q|Z_{jy}|} \\ &\geq \left(\frac{m}{m+1}\right)^2 \sum_{1 \leq x,y \leq q} \frac{e(Z_{ix}, Z_{jy})}{|X_i||X_j|} \\ &= \left(\frac{m}{m+1}\right)^2 d(X_i, X_j). \end{aligned}$$

Now

$$\left(\frac{m}{m+1}\right)^2 \geq 1 - \frac{2}{m+1} \geq 1 - \frac{2kq}{n} \geq 1 - \frac{a}{2}, \tag{5.4}$$

where a, which is to be determined later, must satisfy $4kq/n \leq a \leq 1$. Thus we have

$$\frac{1}{q^2} \sum_{1 \leq x,y \leq q} d(Z_{ix}, Z_{jy}) \geq \left(1 - \frac{a}{2}\right) d(X_i, X_j) \geq d(X_i, X_j) - \frac{a}{2}. \tag{5.5}$$

Applying the Cauchy-Schwarz inequality to (5.5), we obtain

$$\frac{1}{q^2}\sum_{1\le x,y\le q} d^2(Z_{ix},Z_{jy}) \ge \left(\frac{1}{q^2}\sum_{1\le x,y\le q} d(Z_{ix},Z_{jy})\right)^2 \ge d^2(X_i,X_j)-a. \tag{5.6}$$

Thus, before taking advantage of irregularity between parts, we find that we obtain a small loss in the mean square density. Note that a can be chosen to be arbitrarily small by insisting that n is very large.

Now fix i and j so that (X_i,X_j) is not $(\epsilon,\epsilon)_G$-regular. Recall that (X_{ij},X_{ji}) is witness to the $(\epsilon',\epsilon')_G$-irregularity of (X_i,X_j). Let X^*_{ij} (resp. X^*_{ji}) be the largest subset of X_{ij} (resp. X_{ji}) that can be written as the union of some of the Z_{ix} (resp. Z_{jy}). Since X_{ij} is the union of at most l^{k-1} parts of $\mathcal{P}_i$ and each part of $\mathcal{P}_i$ is approximately the union of some of the Z_{ix} via (5.3), so we have that

$$|X^*_{ij}| \ge |X_{ij}| - l^{k-1}(2(m+1)), \tag{5.7}$$

and similarly for X^*_{ji}. After reordering Z_{ix} and Z_{jy} suitably, let us assume that

$$X^*_{ij} = \bigcup_{x=1}^{r} Z_{ix} \quad \text{and} \quad X^*_{ji} = \bigcup_{y=1}^{s} Z_{jy}.$$

Our next task is to use irregularity to give a lower bound for

$$\left|\frac{1}{q^2}\sum_{1\le x,y\le q} d(Z_{ix},Z_{jy}) - \frac{1}{rs}\sum_{\substack{1\le x\le r\\1\le y\le s}} d(Z_{ix},Z_{jy})\right|.$$

We shall then apply Lemma 5.2.2 (the defect form of the Cauchy-Schwarz inequality) by setting

$$\Phi = \frac{1}{q^2}\sum_{1\le x,y\le q} d(Z_{ix},Z_{jy}) \quad \text{and} \quad \phi = \frac{1}{rs}\sum_{\substack{1\le x\le r\\1\le y\le s}} d(Z_{ix},Z_{jy}),$$

to obtain an increase in the mean square density. We prove a lower bound on $|\Phi-\phi|$ by breaking it down into four terms using the triangle inequality as follows:

$$\begin{aligned}
\left|\frac{1}{q^2}\sum_{1\le x,y\le q} d(Z_{ix},Z_{jy}) - \frac{1}{rs}\sum_{\substack{1\le x\le r\\1\le y\le s}} d(Z_{ix},Z_{jy})\right| \\
&\ge |d(X_i,X_j) - d(X_{ij},X_{ji})| && \text{(Term 1)}\\
&- \left|d(X_i,X_j) - \frac{1}{q^2}\sum_{1\le x,y\le q} d(Z_{ix},Z_{jy})\right| && \text{(Term 2)}\\
&- |d(X_{ij},X_{ji}) - d(X^*_{ij},X^*_{ji})| && \text{(Term 3)}\\
&- \left|d(X^*_{ij},X^*_{ji}) - \frac{1}{rs}\sum_{\substack{1\le x\le r\\1\le y\le s}} d(Z_{ix},Z_{jy})\right|. && \text{(Term 4)}
\end{aligned}$$

We now estimate each of these terms.

Term 1: Since (X_{ij}, X_{ji}) is witness to the $(\epsilon'\epsilon')_G$-irregularity of (X_i, X_j), we have

$$|d(X_i,X_j) - d(X_{ij},X_{ji}| > \epsilon'.$$

Term 2: The inequality (5.5) together with the inequality obtained from (5.5) by considering the complementary graph gives us

$$\left|d(X_i,X_j) - \frac{1}{q^2}\sum_{1\le x,y\le q} d(Z_{ix},Z_{jy})\right| \le \frac{a}{2}.$$

Term 3: We shall show that X^*_{ij} forms a large part of X_{ij} and apply Lemma 5.2.3. We have by (5.7) that

$$\begin{aligned}
|X^*_{ij}| \ge |X_{ij}| - l^{k-1}(2(m+1)) &= |X_{ij}|\Big(1 - \frac{2l^{k-1}(m+1)}{|X_{ij}|}\Big)\\
&\ge |X_{ij}|\Big(1 - \frac{4l^{k-1}m}{|X_{ij}|}\Big)\\
&> |X_{ij}|\Big(1 - \frac{4l^{k-1}m}{\epsilon'|X_i|}\Big) \quad \text{since } |X_{ij}| > \epsilon'|X_i|\\
&\ge |X_{ij}|\Big(1 - \frac{4l^{k-1}}{\epsilon' q}\Big) \quad \text{since } |X_i| \ge \left\lfloor\frac{n}{k}\right\rfloor \ge mq\\
&\ge |X_{ij}|\Big(1 - \frac{b}{2}\Big), \qquad (5.8)
\end{aligned}$$

where b, which we shall determine later, must satisfy $\frac{8l^{k-1}}{\epsilon' q} \le b \le 1$. Note that b can be chosen to be arbitrarily small by choosing q sufficiently large.

Now applying Lemma 5.2.3, we have

$$|d(X_{ij}, X_{ji}) - d(X_{ij}^*, X_{ji}^*)| \le 2(b/2) = b.$$

Term 4: This is obtained in a similar way to Term 2, but we give its derivation here for completeness. Since $X_{ij}^* = \cup_{x=1}^r Z_{ix}$ and $X_{ji}^* = \cup_{y=1}^s Z_{jy}$, we have that

$$\sum_{\substack{1\le x\le r\\1\le y\le s}} e(Z_{ix}, Z_{jy}) = e(X_{ij}^*, X_{ji}^*).$$

and furthermore, $|X_{ij}^*| \ge \frac{m}{m+1} r|Z_{ix}|$ for all x and $|X_{ji}^*| \ge \frac{m}{m+1} s|Z_{iy}|$ for all y. Thus

$$\begin{aligned}
\frac{1}{rs}\sum_{\substack{1\le x\le r\\1\le y\le s}} d(Z_{ix}, Z_{jy}) &= \sum_{\substack{1\le x\le r\\1\le y\le s}} \frac{e(Z_{ix}, Z_{jy})}{r|Z_{ix}|s|Z_{jy}|} \\
&\ge \left(\frac{m}{m+1}\right)^2 \sum_{\substack{1\le x\le r\\1\le y\le s}} \frac{e(Z_{ix}, Z_{jy})}{|X_{ij}^*||X_{ji}^*|} \\
&= \left(\frac{m}{m+1}\right)^2 d(X_{ij}^*, X_{ji}^*) \ge d(X_{ij}^*, X_{ji}^*) - \frac{a}{2},
\end{aligned}$$

where the last inequality follows using (5.4). This inequality, together with the inequality obtained from it by considering the complementary graph, gives us

$$\left| d(X_{ij}^*, X_{ji}^*) - \frac{1}{rs}\sum_{\substack{1\le x\le r\\1\le y\le s}} d(Z_{ix}, Z_{jy}) \right| \le \frac{a}{2}.$$

Now putting all four terms together, we obtain

$$\left| \frac{1}{q^2}\sum_{1\le x,y\le q} d(Z_{ix}, Z_{jy}) - \frac{1}{rs}\sum_{\substack{1\le x\le r\\1\le y\le s}} d(Z_{ix}, Z_{jy}) \right| > \epsilon' - \frac{a}{2} - b - \frac{a}{2} = \epsilon' - a - b, \tag{5.9}$$

provided that $\frac{4kq}{n} \le a \le 1$ and $\frac{8l^{k-1}}{\epsilon' q} \le b \le 1$. As we have noted, a and b can be chosen to be arbitrarily small by choosing n and q sufficiently large.

Applying the defect form of the Cauchy-Schwarz inequality (Lemma 5.2.2) to the numbers $d(Z_{ix}, Z_{jy})$ by setting

$$\Phi = \frac{1}{q^2}\sum_{1\le x,y\le q} d(Z_{ix}, Z_{jy}) \quad \text{and} \quad \phi = \frac{1}{rs}\sum_{\substack{1\le x\le r\\1\le y\le s}} d(Z_{ix}, Z_{jy}),$$

we obtain

$$\begin{aligned}\frac{1}{q^2}\sum_{1\le x,y\le q} d^2(Z_{ix},Z_{jy}) &\ge \Phi^2+\frac{rs}{q^2}(\Phi-\phi)^2\\ &> \left(d(X_i,X_j)-\frac{a}{2}\right)^2+\frac{rs}{q^2}(\epsilon'-a-b)^2 \quad \text{(using (5.5), (5.9))}\\ &\ge d^2(X_i,X_j)-a+\frac{rs}{q^2}(\epsilon'-a-b)^2. \qquad (5.10)\end{aligned}$$

Estimating for rs/q^2, we have

$$\begin{aligned}\frac{r}{q}=\left(1-\frac{1}{m+1}\right)\frac{(m+1)r}{mq} &\ge \left(1-\frac{1}{m+1}\right)\frac{|X_{ij}^*|}{|X_i|}\\ &\ge \left(1-\frac{1}{m+1}\right)\left(1-\frac{b}{2}\right)\frac{|X_{ij}|}{|X_i|} \quad \text{(using (5.8))}\\ &> \frac{1}{2}\left(1-\frac{b}{2}\right)\epsilon', \qquad (5.11)\end{aligned}$$

where the last inequality follows from $m\ge 1$ and the fact that (X_{ij},X_{ji}) is a witness to the $(\epsilon',\epsilon')_G$-irregularity of (X_i,X_j). We have the same inequality for s/q, so combining these with (5.10), we obtain

$$\frac{1}{q^2}\sum_{1\le x,y\le q} d^2(Z_{ix},Z_{jy}) > d^2(X_i,X_j)-a+\frac{1}{4}\left(1-\frac{b}{2}\right)^2\epsilon'^2(\epsilon'-a-b)^2. \quad (5.12)$$

Thus (5.12) holds for pairs (X_i,X_j) that are not $(\epsilon,\epsilon)_G$-regular, and (5.6) holds for pairs (X_i,X_j) that are $(\epsilon,\epsilon)_G$-regular. Noting that there are at least $\epsilon\binom{k}{2}\ge\epsilon'\binom{K}{2}$ pairs that are not $(\epsilon,\epsilon)_G$-regular, we have

$$\begin{aligned}\operatorname{ind}(\mathcal{Q}) &\ge \frac{1}{k^2}\sum_{1\le i<j\le k}\frac{1}{q^2}\sum_{1\le x,y\le q} d^2(Z_{ix}.Z_{jy})\\ &> \frac{1}{k^2}\left(\sum_{1\le i<j\le k}(d^2(X_i,X_j)-a)+\epsilon'\binom{k}{2}\frac{1}{4}\left(1-\frac{b}{2}\right)^2\epsilon'^2(\epsilon'-a-b)^2\right)\\ &\ge \operatorname{ind}(\mathcal{P})-\frac{1}{2}a+\frac{1}{64}\epsilon'^3(\epsilon'-a-b)^2,\end{aligned}$$

using that $b\le 1$ and $k\ge 2$ for the last inequality. By setting $a=\alpha\epsilon'^5$ and

$b = \beta\epsilon'$, we obtain

$$\begin{aligned}\mathrm{ind}(\mathcal{Q}) - \mathrm{ind}(\mathcal{P}) &> \left(\frac{1}{64}(1-\alpha-\beta)^2 - \frac{1}{2}\alpha\right)\epsilon'^5 \\ &\geq \left(\frac{1}{64}(1-2\alpha-2\beta) - \frac{1}{2}\alpha\right)\epsilon'^5 = \frac{\epsilon'^5}{100},\end{aligned}$$

where $\alpha = \beta = \frac{1}{100}$.

Thus, we have shown that $\mathrm{ind}(\mathcal{Q}) > \mathrm{ind}(\mathcal{P}) + \frac{\epsilon'^5}{100}$, provided that

$$\frac{4kq}{n} \leq a = \frac{\epsilon'^5}{100} \leq 1 \quad \text{and} \quad \frac{8l^{k-1}}{\epsilon' q} \leq b = \frac{\epsilon'}{100} \leq 1.$$

The conditions of the lemma ensure that such a and b exist. □

It is now an easy step to prove Theorem 5.1.4.

Proof (of Theorem 5.1.4) Recall that $\epsilon, \epsilon', l, m$ are given. Define $f(k) = \lceil 800 l^{k-1} \epsilon'^{-2} \rceil$. Define $f^* : \mathbb{N} \to \mathbb{N}$ inductively by setting $f^*(0) = m$ and $f^*(i+1) = f(f^*(i))$ for $i \in \mathbb{N}$. Set $r = f^*(\lceil 50\epsilon'^{-5} \rceil)$, and set

$$M = M(\epsilon, \epsilon', l, m) = 400 r f(r) \epsilon'^{-5}.$$

Given a graph $G = (X, E)$ and a poset $P = (X, \prec)$, if $|X| < M$, then partition X into vertices to obtain a P-ordered $(\epsilon)_G$-regular equipartition of X. If not, then $n = |X|$ is large enough that we may repeatedly apply Lemma 5.2.6 to G up to $\lceil 50\epsilon'^{-5} \rceil$ times. Start with $\mathcal{P}_0$, any P-ordered equipartition of X into $m = f^*(0)$ parts. If $\mathcal{P}_0$ is not $(\epsilon)_G$-regular, then apply Lemma 5.2.6 to obtain a P-ordered equipartition $\mathcal{P}_1$ with $f^*(1)$ parts and where

$$\mathrm{ind}_G(\mathcal{P}_1) \geq \mathrm{ind}_G(\mathcal{P}_0) + \frac{\epsilon'^5}{100}.$$

Repeat this process to obtain partitions $\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2, \ldots$, where $\mathcal{P}_i$ is a P-ordered equipartition with $f^*(i)$ parts and where

$$\mathrm{ind}_G(\mathcal{P}_i) \geq \mathrm{ind}_G(\mathcal{P}_0) + i\frac{\epsilon'^5}{100}.$$

Since $\mathrm{ind}(\mathcal{P}) < 1/2$ for all equipartitions $\mathcal{P}$, we must eventually reach a

P-ordered equipartition $\mathcal{P}_t$ that is $(\epsilon)_G$-regular, where

$$t \leq \frac{1/2}{\epsilon'^5/100} = 50\epsilon'^{-5}.$$

This partition has $f^*(t)$ parts, and $m \leq f^*(t) \leq M$, as required. □

5.3 Good Classes

In this section, we show that the classes $\mathcal{C}_1$ and $\mathcal{C}_2$, described in Section 5.1.1, are good classes. More precisely, we prove the following.

Lemma 5.3.1 *For every poset $P = (X, \prec)$ and every $\epsilon \in (0,1]$, the graph* Com(P) *is $(P, \epsilon, \epsilon, 2)$-good.*

Lemma 5.3.2 *For every poset $P = (X, \prec)$ and every $\epsilon \in (0,1]$, the graph* Cov(P) *is $(P, \epsilon, \frac{\epsilon^2}{2}, \lceil 8\epsilon^{-3} + 1 \rceil)$-good.*

These lemmas, together with Theorem 5.1.4, prove Theorems 5.1.1 and 5.1.2. Let us give the explicit proofs of these theorems.

Proof (of Theorem 5.1.1) We are given $\epsilon \in (0,1]$ and $m \in \mathbb{N}$. We set $M_1(\epsilon, m) = M(\epsilon, \epsilon, 2, m)$, (recall that M_1 is the function in the statement of Theorem 5.1.1 and that M is the function in the statement of Theorem 5.1.4). For $P = (X, \prec)$ a poset, we know (by Lemma 5.3.1) that Com(P) is $(P, \epsilon, \epsilon, 2)$-good, so by Theorem 5.1.4, there exists a P-ordered equipartition of X into k parts that is $(\epsilon)_{\mathrm{Com}(P)}$-regular, where $m \leq k \leq M(\epsilon, \epsilon, 2, m) = M_1(\epsilon, m)$.

□

Proof (of Theorem 5.1.2) We are given $\epsilon \in (0,1]$ and $m \in \mathbb{N}$. We set $M_2(\epsilon, m) = M(\epsilon, \frac{\epsilon^2}{2}, \lceil 8\epsilon^{-3} + 1 \rceil, m)$, (recall that M_2 is the function in the statement of Theorem 5.1.2 and that M is the function in the statement of Theorem 5.1.4).) For $P = (X, \prec)$ a poset, we know (by Lemma 5.3.1) that Cov(P) is $(P, \epsilon, \frac{\epsilon^2}{2}, \lceil 8\epsilon^{-3} + 1 \rceil)$-good, so by Theorem 5.1.4, there exists a P-ordered equipartition of X into k parts that is $(\epsilon)_{\mathrm{Cov}(P)}$-regular, where

$$m \leq k \leq M(\epsilon, \frac{\epsilon^2}{2}, \lceil 8\epsilon^{-3} + 1 \rceil, m) = M_2(\epsilon, m).$$

□

It remains for us to prove Lemma 5.3.1 and Lemma 5.3.2. Lemma 5.3.1 is natural and straightforward to prove. Lemma 5.3.2, although more unexpected, is not difficult to prove.

Proof (of Lemma 5.3.1) Given the poset $P = (X, \prec)$, let Y_1, Y_2 be a P-ordered equipartition of $Y \subseteq X$, where (Y_1, Y_2) is not $(\epsilon, \epsilon)_{\text{Com}(P)}$-regular. We show that there exist P-ordered partitions A_1, A_2 of Y_1 and B_1, B_2 of Y_2 such that either (A_1, B_2) or (A_2, B_1) is witness to the $(\epsilon, \epsilon)_{\text{Com}(P)}$-irregularity of $(Y_!, Y_2)$. This then proves the lemma.

We drop Com(P) as a subscript. Since (Y_1, Y_2) is not (ϵ, ϵ)-regular, there exists a witness (A, B) to the irregularity of (Y_1, Y_2), that is, there exists $A \subseteq Y_1$ and $B \subseteq Y_2$ with $|A| > \epsilon|Y_1|$ and $|B| > \epsilon|Y_2|$, where

$$|d(Y_1, Y_2) - d(A, B)| > \epsilon.$$

We consider two cases.

Case (i) Suppose that $d(A, B) > d(Y_1, Y_2) + \epsilon$. If $x \in A$ and $x' \in Y_1 \backslash A$ with $x' \prec x$, then replacing x with x' in A does not reduce $e(A, B)$ (by the transitivity of the poset) and hence does not reduce $d(A, B)$. Similarly if $y \in B$ and $y' \in Y_2 \backslash B$, with $y' \succ y$, then replacing y with y' in B does not reduce $d(A, B)$. Thus, after such replacements, (A, B) remains a witness to the (ϵ, ϵ)-irregularity of (Y_1, Y_2). Now repeatedly make such replacements until no more can be made and call the resulting sets A_1 and B_2. Thus A_1 is a down-set of Y_1; B_2 is an up-set of Y_2; and (A_1, B_2) is a witness to the (ϵ, ϵ)-irregularity of (Y_1, Y_2).

Let $A_2 = Y_1 \backslash A_1$ and $B_1 = Y_2 \backslash B_2$. Then A_1, A_2 is a P-ordered partition of Y_1; B_1, B_2 is a P-ordered partition of Y_2; and (A_1, B_2) is a witness to the (ϵ, ϵ)-irregularity of (Y_1, Y_2) as required.

Case (ii) Suppose that $d(A, B) < d(Y_1, Y_2) - \epsilon$. This case follows in a very similar way to Case (i). We give the details for completeness. If $x \in A$ and $x' \in Y_1 \backslash A$ with $x' \succ x$, then replacing x with x' in A does not increase $e(A, B)$ (by the transitivity of the poset) and hence does not increase $d(A, B)$. Similarly if $y \in B$ and $y' \in Y_1 \backslash B$, with $y' \prec y$, then replacing y with y' in B does not increase $d(A, B)$. Thus, after such replacements, (A, B) remains a witness to the (ϵ, ϵ)-irregularity of (Y_1, Y_2). Now repeatedly make such replacements until no more can be made and call the resulting sets A_2 and B_1. Thus A_2 is a up-set of Y_1; B_1 is an down-set of Y_2; and (A_2, B_1) is a witness to the (ϵ, ϵ)-irregularity of (Y_1, Y_2).

Let $A_1 = Y_1 \backslash A_2$ and $B_2 = Y_2 \backslash B_1$. Then A_1, A_2 is a P-ordered partition of Y_1; B_1, B_2 is an P-ordered partition of Y_2; and (A_2, B_1) is a witness to the (ϵ, ϵ)-irregularity of (Y_1, Y_2) as required. □

We need some preliminary lemmas before we can prove Lemma 5.3.2, but first we give some notation.

Let $P = (X, \prec)$ be a poset with $Y \subseteq X$. We define $ht_P(Y)$ to be the *height* of Y in P, that is, the size of the largest chain in Y.

We have the following easy lemma.

Lemma 5.3.3 *Let $P = (X, \prec)$ be a poset with $Y' \subseteq Y \subseteq X$ and let $h = ht_P(Y')$. Then there exists a P-ordered partition $\mathcal{P}$ of Y into $2h+1$ parts such that Y' is $\mathcal{P}$-unifiable (and hence $Y \backslash Y'$ is also $\mathcal{P}$-unifiable).*

Proof Recall from Proposition 1.2.1 that if we have a nested sequence of down-sets of $Q = (Y, \prec)$, $\phi = D_0 \subseteq D_1 \subseteq D_2 \subseteq \cdots \subseteq D_r = Y$, then by setting $X_i = D_i \backslash D_{i-1}$ for $i = 1, \ldots, r$, we have that $X_1, \ldots, X_r$ is a Q-ordered partition of Y. Trivially, this is also a P-ordered partition of Y.

Given $A \subseteq Y$, we define

$$D(A) = \{y \in Y : y \prec a \text{ for some } a \in A\}$$
$$\text{and } \bar{D}(A) = \{y \in Y : y \preceq a \text{ for some } a \in A\}.$$

These are both down-sets of Y by the transitivity of $\prec$.

Given that Y' has height h, we have the standard decomposition of Y' into h antichains $A_1, \ldots, A_h$ given by Theorem 1.2.3. Recall that these antichains are constructed inductively by setting $A_i = \max_P(Y'_i)$, where

$$Y'_h = Y' \quad \text{and} \quad Y'_i = Y' \backslash \bigcup_{j=i+1}^{h} A_j \quad \text{for} \quad i = 1, \ldots, h-1.$$

Now we have the following nested sequence of down-sets of Y:

$$\phi \subseteq D(A_1) \subseteq \bar{D}(A_1) \subseteq D(A_2) \subseteq \bar{D}(A_2) \subseteq \cdots \subseteq D(A_h) \subseteq \bar{D}(A_h) \subseteq Y.$$

It is clear that $D(A_i) \subseteq \bar{D}(A_i)$ for all i. To see that $\bar{D}(A_{i-1}) \subseteq D(A_i)$, note that $A_{i-1} \subseteq Y'_{i-1} \subseteq D(A_i)$ and therefore $\bar{D}(A_{i-1}) \subseteq D(A_i)$ (by transitivity).

Let $X_1, \ldots, X_{2h+1}$ be the P-ordered partition of Y induced by this nested sequence of down-sets. Then we have that

$$\bigcup_{i=1}^{h} X_{2i} = \bigcup_{i=1}^{h} \bar{D}(A_i)\backslash D(A_i) = \bigcup_{i=1}^{h} A_i = Y'$$

as required. □

In light of Lemma 5.3.3, one way to show that a graph G is $(P, \epsilon, \epsilon', l)$-good is the following. Whenever (A, B) is P-ordered and $(\epsilon, \epsilon)_G$-irregular, we try to find a witness (A', B') to the $(\epsilon', \epsilon')_G$-irregularity of (A, B) for which A' and B' (or their complements) have height at most $(l-1)/2$ in P.

When $G = \text{Cov}(P)$, we observe that two elements on the same chain in A have disjoint neighbourhoods in B; therefore elements with large neighbourhoods in B cannot form large chains in A. Thus for every irregular pair, we seek witnesses to irregularity where the elements in the witness sets (or their complements) have large neighbourhoods. This is the motivation behind the next lemma.

Lemma 5.3.4 *Let $G = (V, E)$ be a graph with $V_1, V_2 \subseteq V$ disjoint, where (V_1, V_2) is not $(\epsilon, \epsilon)_G$-regular and (V_1', V_2') is witness to this. Dropping G as a subscript, we have the following.*

(i) If $d(V_1', V_2') > d(V_1, V_2) + \epsilon$, then there exists $(\bar{V_1}, \bar{V_2})$ that is witness to the $(\epsilon, \frac{\epsilon^2}{2})$-irregularity of (V_1, V_2), where

$$e(x, V_2) \geq \frac{\epsilon^3}{4}|V_2|, \forall x \in \bar{V_1} \quad \textit{and} \quad e(y, V_1) \geq \frac{\epsilon^3}{4}|V_1|, \forall y \in \bar{V_2}.$$

(ii) If $d(V_1', V_2') < d(V_1, V_2) - \epsilon$, then there exists $(\bar{V_1}, \bar{V_2})$ that is witness to the $(\frac{\epsilon}{2}, \epsilon)$-irregularity of (V_1, V_2), where

$$e(x, V_2) \geq \frac{\epsilon^3}{4}|V_2|, \forall x \in V_1\backslash\bar{V_1} \quad \textit{and} \quad e(y, V_1) \geq \frac{\epsilon^3}{4}|V_1|, \forall y \in V_2\backslash\bar{V_2}.$$

Proof (i) Consider the vertices $x \in V_1'$ for which $d(x, V_2) < \epsilon^2/2$. We call these *low degree* vertices of V_1'. For such vertices, we have

$$d(x, V_2') = \frac{e(x, V_2')}{|V_2'|} \leq \frac{e(x, V_2)}{\epsilon|V_2|} = \frac{d(x, V_2)}{\epsilon} < \frac{\epsilon}{2}.$$

There are strictly fewer than $(1-\frac{\epsilon}{2})|V_1'|$ low degree vertices of V_1'. Indeed, if not, then

$$d(V_1', V_2') = \frac{\sum_{x \in V_1'} d(x, V_2')}{|V_1'|} \leq \frac{1}{|V_1'|}\left(\left(1-\frac{\epsilon}{2}\right)|V_1'| \cdot \frac{\epsilon}{2} + \frac{\epsilon}{2}|V_1'| \cdot 1\right) < \epsilon,$$

contradicting that $d(V_1', V_2') > d(V_1, V_2) + \epsilon \geq \epsilon$.

Remove the low degree vertices of V_1' from V_1' to form $\bar{V_1}$ and observe that $d(\bar{V_1}, V_2') \geq d(V_1', V_2')$ (since we have removed vertices of lower than average degree) and that $|\bar{V_1}| > \frac{\epsilon^2}{2}|V_1|$.

Similarly, any vertex $y \in V_2'$ for which $d(y, V_1) < \epsilon^3/4$ is called a *low degree* vertex of V_2'. For such vertices, we have

$$d(y, \bar{V_1}) = \frac{e(y, \bar{V_1})}{|\bar{V_1}|} \leq \frac{e(y, V_1)}{\frac{1}{2}\epsilon^2|V_1|} = \frac{d(y, V_1)}{\frac{1}{2}\epsilon^2} < \frac{\epsilon}{2}.$$

There are strictly fewer than $(1-\frac{\epsilon}{2})|V_2'|$ low degree vertices of V_2'. Indeed, if not, then

$$d(V_2', \bar{V_1}) = \frac{\sum_{y \in V_2'} d(y, \bar{V_1})}{|V_2'|} \leq \frac{1}{|V_2'|}\left(\left(1-\frac{\epsilon}{2}\right)|V_2'| \cdot \frac{\epsilon}{2} + \frac{\epsilon}{2}|V_2'| \cdot 1\right) < \epsilon,$$

contradicting that $d(\bar{V_1}, V_2') \geq d(V_1', V_2') > \epsilon$.

Remove the low degree vertices of V_2' from V_2' to form $\bar{V_2}$. As before, observe that

$$d(\bar{V_1}, \bar{V_2}) \geq d(\bar{V_1}, V_2') \geq d(V_1', V_2') > d(V_1, V_2) + \epsilon,$$

where the first inequality holds because we have removed vertices of smaller than average degree. Furthermore, note that $|\bar{V_2}| > \frac{\epsilon^2}{2}|V_2|$.

Thus, $(\bar{V_1}, \bar{V_2})$ is witness to the $(\epsilon, \frac{\epsilon^2}{2})$-irregularity of (V_1, V_2), and

$$e(x, V_2) \geq \frac{\epsilon^2}{2}|V_2| \geq \frac{\epsilon^3}{4}|V_2|, \forall x \in \bar{V_1} \quad \text{and} \quad e(y, V_1) \geq \frac{\epsilon^3}{4}|V_1|, \forall y \in \bar{V_2}.$$

(ii) Add to V_1' all vertices $x \in V_1$ for which $d(x, V_2) < \epsilon^2/4$ to form $\bar{V_1}$. For such vertices, using a similar calculation as before, we have $d(x, V_2') < \epsilon/4$, and hence $d(\bar{V_1}, V_2') < d(V_1', V_2') + \frac{\epsilon}{4}$.

Add to V_2' all vertices $y \in V_2$ for which $d(y, V_1) < \epsilon^2/4$ to form $\bar{V_2}$. For such vertices, using a similar calculation as before, we have $d(y, \bar{V_1}) < \epsilon/4$, and hence

$$d(\bar{V_1}, \bar{V_2}) < d(\bar{V_1}, V_2') + \frac{\epsilon}{4} < d(V_1', V_2') + \frac{\epsilon}{2} < d(V_1, V_2) - \frac{\epsilon}{2}.$$

Thus, $(\bar{V_1}, \bar{V_2})$ is witness to the $(\frac{\epsilon}{2}, \epsilon)$-irregularity of (V_1, V_2), and

$$e(x, V_2) \geq \frac{\epsilon^2}{4}|V_2|, \forall x \in V_1 \backslash \bar{V_1} \quad \text{and} \quad e(y, V_1) \geq \frac{\epsilon^2}{4}|V_1|, \forall y \in V_2 \backslash \bar{V_2}.$$

□

Finally, we apply Lemma 5.3.3 and Lemma 5.3.4 to the covering graph of a poset to obtain Lemma 5.3.2.

Proof (of Lemma 5.3.2) We are given a poset $P = (X, \prec)$. Let Y_1, Y_2 be an ordered partition of $Y \subseteq X$, where (Y_1, Y_2) is not $(\epsilon, \epsilon)_{\mathrm{Cov}(P)}$-regular. Let (Y_1', Y_2') be witness to this. It is sufficient for us to prove that there exists a witness $(\bar{Y_1}, \bar{Y_2})$ to the $(\frac{\epsilon^2}{2}, \frac{\epsilon^2}{2})_{\mathrm{Cov}(P)}$-irregularity of (Y_1, Y_2) and ordered partitions $\mathcal{P}_1$ of Y_1 and $\mathcal{P}_2$ of Y_2 into at most $8\epsilon^{-3} + 1$ parts such that $\bar{Y_1}$ is $\mathcal{P}_1$-unifiable and $\bar{Y_2}$ is $\mathcal{P}_2$-unifiable. We consider two cases. Throughout, we work with $\mathrm{Cov}(P)$, so we drop it as a subscript.

(i) Suppose $d(Y_1', Y_2') > d(Y_1, Y_2) + \epsilon$. By Lemma 5.3.4(i), we can find a witness $(\bar{Y_1}, \bar{Y_2})$ to the $(\epsilon, \frac{\epsilon^2}{2})$-irregularity (and hence the $(\frac{\epsilon^2}{2}, \frac{\epsilon^2}{2})$-irregularity) of (Y_1, Y_2), where

$$e(x, Y_2) \geq \frac{\epsilon^3}{4}|Y_2|, \forall x \in \bar{Y_1} \quad \text{and} \quad e(y, Y_1) \geq \frac{\epsilon^3}{4}|Y_1|, \forall y \in \bar{Y_2}.$$

For $x, x' \in Y_1$ with $x' \prec x$, we have that $E(x', Y_2) \cap E(x, Y_2) = \phi$. Therefore, if C is any chain in $\bar{Y_1}$, then

$$|C|\frac{\epsilon^3}{4}|Y_2| \leq \sum_{x \in C} e(x, Y_2) \leq |Y_2|,$$

and hence $|C| \leq 4\epsilon^{-3}$. Therefore $ht_P(\bar{Y_1}) \leq 4\epsilon^{-3}$. Similarly $ht_P(\bar{Y_2}) \leq 4\epsilon^{-3}$. Applying Lemma 5.3.3, we deduce that there exist partitions $\mathcal{P}_i$ of Y_i into at most $8\epsilon^{-3} + 1$ parts such that $\bar{Y_i}$ is $\mathcal{P}_i$-unifiable for $i = 1, 2$, as required.

(ii) This is similar to (i). Suppose $d(Y_1', Y_2') < d(Y_1, Y_2) - \epsilon$. Using Lemma 5.3.4(ii), we can find a witness $(\bar{Y_1}, \bar{Y_2})$ to the $(\frac{\epsilon}{2}, \epsilon)$-irregularity (and hence the $(\frac{\epsilon^2}{2}, \frac{\epsilon^2}{2})$-irregularity) of (Y_1, Y_2), where

$$e(x, Y_2) \geq \frac{\epsilon^3}{4}|Y_2|, \forall x \in Y_1 \backslash \bar{Y_1} \quad \text{and} \quad e(y, Y_1) \geq \frac{\epsilon^3}{4}|Y_1|, \forall y \in Y_2 \backslash \bar{Y_2}.$$

As before, if C is any chain in $Y_1 \backslash \bar{Y_1}$, then

$$|C|\frac{\epsilon^3}{4}|Y_2| \leq \sum_{x \in C} e(x, Y_2) \leq |Y_2|,$$

and hence $|C| \leq 4\epsilon^{-3}$. Therefore $ht_P(Y_1 \backslash \bar{Y_1}) \leq 4\epsilon^{-3}$. Similarly, we have $ht_P(Y_2 \backslash \bar{Y_2}) \leq 4\epsilon^{-3}$. Applying Lemma 5.3.3, we deduce that there exist partitions $\mathcal{P}_i$ of Y_i into at most $8\epsilon^{-3} + 1$ parts such that $Y_i \backslash \bar{Y_i}$ (and hence $\bar{Y_i}$) is $\mathcal{P}_i$-unifiable for $i = 1, 2$, as required. □

5.4 Simultaneous Regular Ordered Partitions and Further Properties

We begin this section with the proof of Theorem 5.1.6. The proof of this is an easy application of Lemma 5.2.6 and Lemma 5.2.1.

Proof (of Theorem 5.1.6) We are given ϵ and m. Let $P = (X, \prec)$ be a poset. We know that $\mathrm{Com}(P)$ and $\mathrm{Cov}(P)$ are both $(P, \epsilon, \epsilon', l)$-good, where $\epsilon' = \frac{\epsilon^2}{2}$ and $l = \lceil 8\epsilon^{-3} + 1 \rceil$.

Suppose $\mathcal{P}$ is a P-ordered equipartition of X into k parts. Let $q = f(k) := \lceil 800 l^{k-1} \epsilon'^{-2} \rceil$ and suppose that

$$n = |X| \geq 400kq\epsilon'^{-5}.$$

Define $\bar{m} := \lfloor n/kq \rfloor \geq \lfloor 400\epsilon'^{-5} \rfloor$. Now $\mathcal{P}$ can be refined in one of two ways using Lemmas 5.2.1 and 5.2.6.

(i) If $\mathcal{P}$ is not $(\epsilon)_{\mathrm{Cov}(P)}$-regular, then we apply Lemma 5.2.6 to $\mathrm{Cov}(P)$ and Lemma 5.2.1 to $\mathrm{Com}(P)$ to obtain a refinement $\mathcal{Q}_1$ of $\mathcal{P}$, where $\mathcal{Q}_1$ is

P-ordered, has kq parts, and where

$$\begin{aligned}
&\text{ind}_{\text{Com}(P)}(\mathcal{Q}_1) + \text{ind}_{\text{Cov}(P)}(\mathcal{Q}_1) \\
&\qquad > \text{ind}_{\text{Com}(P)}(\mathcal{P}) - \frac{2}{\bar{m}+1} + \text{ind}_{\text{Cov}(P)}(\mathcal{P}) + \frac{\epsilon'^5}{100} \\
&\qquad \geq \text{ind}_{\text{Com}(P)}(\mathcal{P}) + \text{ind}_{\text{Cov}(P)}(\mathcal{P}) + \frac{\epsilon'^5}{200},
\end{aligned}$$

the last inequality following from the condition on $\bar{m}$.

(ii) Similarly, if $\mathcal{P}$ is not $(\epsilon)_{\text{Com}(P)}$-regular, then we apply Lemma 5.2.6 to $\text{Com}(P)$ and Lemma 5.2.1 to $\text{Cov}(P)$ to obtain a refinement $\mathcal{Q}_2$ of $\mathcal{P}$, where $\mathcal{Q}_2$ is P-ordered, has kq parts, and where

$$\text{ind}_{\text{Com}(P)}(\mathcal{Q}_2) + \text{ind}_{\text{Cov}(P)}(\mathcal{Q}_2) \geq \text{ind}_{\text{Com}(P)}(\mathcal{P}) + \text{ind}_{\text{Cov}(P)}(\mathcal{P}) + \frac{\epsilon'^5}{200}.$$

Now, provided n is large enough, we construct partitions $\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2, \ldots$ as follows. Let $\mathcal{P}_0$ be any P-ordered equipartition into m parts. Given $\mathcal{P}_{j-1}$, if $\mathcal{P}_{j-1}$ is not $(\epsilon)_{\text{Cov}(P)}$-regular, then refine it according to (i) to give $\mathcal{P}_j$; if $\mathcal{P}_{j-1}$ is not $(\epsilon)_{\text{Com}(P)}$-regular, then refine it according to (ii) to give $\mathcal{P}_j$; if $\mathcal{P}_{j-1}$ is (ϵ)-regular for both graphs, then stop.

We claim that the above process terminates at the kth iteration, where $k \leq 200\epsilon^{-5}$. Indeed, we know that $\text{ind}_{\text{Com}(P)}(\mathcal{P}) + \text{ind}_{\text{Cov}(P)}(\mathcal{P}) < 1$ for all equipartitions $\mathcal{P}$ and that at each iteration, $\text{ind}_{\text{Com}(P)}(\mathcal{P}) + \text{ind}_{\text{Cov}(P)}(\mathcal{P})$ increases by at least $\epsilon'^5/200$. Hence after at most $200\epsilon'^{-5}$ steps, the process terminates, and we obtain an ordered partition that is ϵ-regular for both $\text{Com}(P)$ and $\text{Cov}(P)$.

The number of parts in the final partition, described above, is at most $r := f^*(\lceil 200\epsilon^{-5} \rceil)$, where f^* is the function described in the proof of Theorem 5.1.4. In order to meet the conditions on n, it is sufficient that

$$n > M_3 = 400rf(r)\epsilon'^{-5}.$$

Now, if $n < M_3 = M_3(\epsilon, m)$, then we partition X into elements; otherwise, we partition X as previously described into at most $r \leq M_3$ parts. □

Having proved that there exists an ordered equipartition that is regular simultaneously for both the covering graph and the comparability graph of a poset, we now prove some simple properties of such partitions. Here is the proof of Theorem 5.1.7

Proof (of Theorem 5.1.7) Recall that we are given a poset $P = (X, \prec)$ and a P-ordered equipartition $X_1, \ldots, X_k$ of X that is ϵ-regular for both $\mathrm{Com}(P)$ and $\mathrm{Cov}(P)$.

(i) We work exclusively with the graph $\mathrm{Com}(P)$ in this part of the theorem, so we drop it as a subscript. We are given parts X_r, X_s, and X_t with $r < s < t$, where (X_r, X_s) and (X_s, X_t) are both (ϵ, ϵ)-regular, and where $p := d(X_r, X_s) > 2\epsilon$ and $q := d(X_s, X_t) > 2\epsilon$. We show that $d(X_r, X_t) \geq 1 - 2\epsilon$.

Let $\bar{A} \subseteq X_r$ be the set consisting of elements $x \in X_r$ for which $d(x, X_s) \leq \epsilon$. Thus $d(\bar{A}, X_s) \leq \epsilon < p - \epsilon$, and so $|\bar{A}| \leq \epsilon|X_r|$, otherwise the (ϵ, ϵ)-regularity of (X_r, X_s) would be violated. Let $A = X_r \backslash \bar{A}$.

For each $x \in A$, define $B_x = \{y \in X_s : y \succ x\} \subseteq X_s$. We know that $|B_x| > \epsilon|X_s|$ since $x \in A$.

For each $x \in A$, let $\bar{C}_x \subseteq X_t$ be the set consisting of elements $z \in X_t$ for which z is incomparable to every element in B_x. Thus $d(\bar{C}_x, B_x) = 0 < q - \epsilon$, and since $|B_x| > \epsilon|X_s|$, we must have that $|\bar{C}_x| \leq \epsilon|X_t|$, otherwise the (ϵ, ϵ)-regularity of (X_s, X_t) is violated. Let $C_x = X_t \backslash \bar{C}_x$.

For every $x \in A$ and every $z \in C_x$, there exists $y \in B_x$ such that $x \prec y \prec z$. Therefore, since $|C_x| \geq (1 - \epsilon)|X_t|$, we have $d(x, X_t) \geq 1 - \epsilon$ for all $x \in A$. Since $|A| \geq (1-\epsilon)|X_r|$, we have that $d(X_r, X_t) \geq (1-\epsilon)^2 \geq 1-2\epsilon$. This completes the proof of (i).

(ii) Again, we deal exclusively with the graph $\mathrm{Com}(P)$ in this part of the theorem, so we drop it as a subscript. We are given parts X_r and X_s, where (X_r, X_s) is (ϵ, ϵ)-regular and $p := d(X_r, X_s)$ satisfies $2\epsilon < p < 1 - 2\epsilon$. We show that $e(X_r) \leq 2\epsilon|X_r|^2$. (By considering the dual poset, we can deduce that $e(X_s) \leq 2\epsilon|X_s|^2$.)

Assume, for a contradiction, that $e(X_r) > 2\epsilon|X_r|^2$. As before, let $\bar{A} \subseteq X_r$ be the set consisting of elements $x \in X_r$ for which $d(x, X_s) \leq \epsilon$, and let $A = X_r \backslash \bar{A}$. We have seen that $|\bar{A}| \leq \epsilon|X_r|$ (otherwise $(\bar{A}, X_s)$ is a witness to the (ϵ, ϵ)-irregularity of (X_r, X_s)).

We know that $e(A) + e(\bar{A}, X_r) = e(X_r) > 2\epsilon|X_r|^2$. Since $\bar{A} \leq \epsilon|X_r|$ then $e(\bar{A}, X_r) \leq \epsilon|X_r|^2$, hence $e(A) > \epsilon|X_r|^2$.

For each $x \in A$, define $D_A(x) = \{y \in A : y \prec x\}$. Observe that

$$\sum_{x \in A} |D_A(x)| = e(A) > \epsilon |X_r|^2 \geq \epsilon |A||X_r|,$$

so there exists some $x^* \in A$ such that $|D_A(x^*)| > \epsilon |X_r|$. Let $U_{X_s}(x^*) = \{y \in X_s : y \succ x^*\}$. We have that $d(D_A(x^*), U_{X_s}(x^*)) = 1 > p + \epsilon$, where $|D_A(x^*)| > \epsilon |X_r|$ and $|U_{X_s}(x^*)| > \epsilon |X_s|$, violating the (ϵ, ϵ) regularity of (X_r, X_s). Hence we have a contradiction and this completes the proof of (ii).

(iii) We are given parts X_r and X_s, where (X_r, X_s) is (ϵ, ϵ)-regular for both $\mathrm{Com}(P)$ and $\mathrm{Cov}(P)$, and where $p := d_{\mathrm{Com}(P)}(X_r, X_s)$ and $q := d_{\mathrm{Cov}(P)}(X_r, X_s) > 3\epsilon$. We show that there exists an antichain $A_r \subseteq X_r$ for which $|A_r| \geq (1 - 2\epsilon)|X_r|$. (By considering the dual poset, we can deduce that there exists an antichain $A_s \subseteq X_s$ for which $|A_s| \geq (1 - 2\epsilon)|X_s|$.)

Assume, for a contradiction, that the largest antichain in X_r has size strictly smaller than $(1 - 2\epsilon)|X_r|$. Let I be a maximal set of independent edges of $\mathrm{Com}(P)$ in X_r. The elements of X_r not incident with any edges in I form an antichain (since I maximal), hence there are strictly fewer than $(1 - 2\epsilon)|X_r|$ such elements. Therefore, we have $|I| > \epsilon |X_r|$. Let $x_i^- x_i^+$ be the edges in I for $i = 1, \ldots, l$, $l > \epsilon |X_r|$, where $x_i^- \prec x_i^+$.

Observe that

$$\begin{aligned} E_{\mathrm{Cov}(P)}(x_i^-, X_s) \cap E_{\mathrm{Com}(P)}(x_i^+, X_s) = \phi \\ \text{and } E_{\mathrm{Com}(P)}(x_i^-, X_s) \supseteq E_{\mathrm{Cov}(P)}(x_i^-, X_s) \cup E_{\mathrm{Com}(P)}(x_i^+, X_s). \end{aligned}$$

Therefore, we have

$$d_{\mathrm{Com}(P)}(x_i^-, X_s) \geq d_{\mathrm{Cov}(P)}(x_i^-, X_s) + d_{\mathrm{Com}(P)}(x_i^+, X_s).$$

Let $A^- = \{x_i^- : i = 1, \ldots, l\}$ and $A^+ = \{x_i^+ : i = 1, \ldots, l\}$. We know that $|A^+|, |A^-| > \epsilon |X_r|$, so using the regularity of $\mathrm{Com}(P)$ and $\mathrm{Cov}(P)$, we obtain

$$\begin{aligned}
p+\epsilon &\geq d_{\mathrm{Com}(P)}(A^-, X_s) \\
&\geq d_{\mathrm{Cov}(P)}(A^-, X_s) + d_{\mathrm{Com}(P)}(A^+, X_s) \\
&\geq (q-\epsilon) + (p-\epsilon) \\
&> p+\epsilon, \quad \text{(since } q > 3\epsilon\text{)}
\end{aligned}$$

which is a contradiction. This completes the proof of (iii) and the proof of the theorem.

□

Chapter 6

An Exact and a Stability Result for Hypergraphs

6.1 Introduction

In Chapter 1, we introduced and discussed Turán-type problems. Although the problem we explore in this chapter is not quite a Turán-type problem, it has many similarities to such problems. It is closely related to the problem of determining the Turán density of K_4^-, as we shall discuss.

We begin by defining a family of 3-graphs derived from a geometric construction. Before we can describe this family, we define what we mean by the *convex hull* of a set of vectors in $\mathbb{R}^m$. Let $U = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$, where $\mathbf{u}_i \in \mathbb{R}^m$ for $i = 1, \ldots, k$. We define *conv*(U), the *convex hull* of U, to be

$$\left\{ \sum_{i=1}^{k} \lambda_i \mathbf{u}_i : \sum_{i=1}^{k} \lambda_i = 1,\ \lambda_i \geq 0 \text{ for all } i \right\}.$$

For a k-graph $H = (V_H, E_H)$ and a set $S \subseteq V$, we say that S *induces* or *spans* r edges of H if $|E_H \cap S^{(k)}| = r$, that is, there are exactly r edges of H contained entirely inside S.

Let $\mathcal{C}$ be the set of 3-graphs constructed as follows. We have $H \in \mathcal{C}$ if the vertices of H can be placed (at distinct points) on the circumference of a circle in $\mathbb{R}^2$ with centre O so that the edges of H are precisely those triples whose convex hull contains O. (We insist that the convex hull of every pair of points does not contain O.) Every $H \in \mathcal{C}$ has the following interesting property: every set of four vertices of H spans exactly zero or two edges of

H. Indeed, suppose $S = \{v_1, v_2, v_3, v_4\}$ is a set of four vertices of $H \in \mathcal{C}$ ordered clockwise around the circle. Let us identify the vertices with their position vectors. Observe that the convex hulls of $\{v_1, v_2, v_3\}$ and $\{v_1, v_3, v_4\}$ partition the convex hull of S (except for the intersection along $conv\{v_1, v_3\}$). Similarly, the convex hulls of $\{v_1, v_2, v_4\}$ and $\{v_2, v_3, v_4\}$ partition the convex hull of S (except for the intersection along $conv\{v_2, v_4\}$). Thus, if the convex hull of S contains O, then S induces exactly two edges of H; if not, then clearly S induces zero edges of H.

This was observed by Frankl and Füredi in [22], where they classified all 3-graphs for which every set of four vertices induces exactly zero or two edges. They found that the only other 3-graphs having this property are the *blowups* of F_6. Let us define these terms.

Let $F = (V_F, E_F)$ be a 3-graph, where $V_F = [n]$. A blowup of F is obtained by expanding each vertex $v_i \in V_F$ into a class V_i of vertices and placing an edge between vertices in different classes if and only if F has a corresponding edge. Formally, $B = (V_B, E_B)$ is a blowup of $F = ([n], E_F)$ if we can partition V_B into nonempty sets $V_1, \dots, V_n$ such that

$$E_B = \bigcup_{abc \in E_F} V_a V_b V_c,$$

where

$$V_a V_b V_c = \{xyz : x \in V_a, y \in V_b, z \in V_c\}.$$

Thus, any set of n vertices from V_B that contains exactly one vertex from each V_i induces a copy of F. An *equipartitioned blowup* of F is a blowup of F in which the class sizes are as equal as possible.

We define F_6 to be the 3-graph with vertex set $\{1, \dots, 6\}$ and edge set $E_6 = \{123, 234, 345, 451, 512, 613, 624, 635, 641, 652\}$. One can easily check that every set of four vertices of F_6 spans exactly two edges. Furthermore, any blowup G of F_6 also has this property. Indeed, let S be a set of four vertices of G. If the vertices of S are taken from four different vertex classes of G, then S will span two edges of G; if the vertices of S are taken from three different vertex classes V_a, V_b, V_c of G and $abc \in E_6$, then S will span two edges. In all other cases, S spans zero edges of G. Thus, all blowups of F_6 and all 3-graphs in $\mathcal{C}$ satisfy the property that every set of four vertices spans exactly zero or two edges. Frankl and Füredi [22] proved that these are the only 3-graphs satisfying that property. In the sections that follow, we give a new proof of this result and a related stability result. Before stating the results, we give some background and motivation.

Recall that a hypergraph is *k-colourable* if its vertices can be partitioned into k colour classes so that no edge of H is contained entirely within a colour class. We write $\chi(H)$ for the *chromatic number* of a hypergraph H, that is, the least k for which H is k-colourable. We note that every hypergraph in $\mathcal{C}$ is 2-colourable. This is because any diameter of the circle on which the vertices of $H \in \mathcal{C}$ are placed separates the vertices of H into two colour classes. Also, we find that $\chi(F_6) = 3$, and consequently, any blowup of F_6 also has chromatic number 3.

For a 3-graph $H = (V_H, E_H)$, let $q_i(H)$ be the number of elements of $V_H^{(4)}$ inducing exactly i edges of H. Here is the result of Frankl and Füredi stated in the form in which we shall prove it.

Theorem 6.1.1 *Suppose $H = (V_H, E_H)$ is a 3-graph such that $q_i(H) = 0$ for $i = 1, 3, 4$.*

(a) If H is 2-colourable, then $H \in \mathcal{C}$.

(b) If H is not 2-colourable, then H is a blowup of F_6.

As noted implicitly in [22], the densest 3-graph in $\mathcal{C}$ on n (odd) vertices is given by placing the vertices evenly around the circumference of the circle. The density of these graphs is asymptotically $\frac{1}{4}$. The densest blowup of F_6 on n vertices is the equipartitioned blowup, whose density is asymptotically $\frac{5}{18}$.

The result of Frankl and Füredi is closely related to the problem of determining the Turán density of K_4^-. Recall that K_4^- is the 3-graph on four vertices with three edges. Determining $\pi(K_4^-)$ is one of the most basic problems in extremal hypergraph theory since K_4^- is the smallest 3-graph that has a non-zero Turán density. As mentioned in Chapter 1, the best known bounds for this problem are

$$\frac{2}{7} \leq \pi(K_4^-) < \frac{1}{3} - \frac{1}{280}. \tag{6.1}$$

The upper bound, which was recently proved by Talbot [59], is an improvement on a bound given by Mubayi [42] using a supersaturation argument, which in turn improved the bound $\pi(K_4^-) \leq \frac{1}{3}$ given by de Caen [8]. The lower bound is due to a construction given by Frankl and Füredi [22] and is derived from the graph F_6 as follows. We take an equipartitioned blowup of F_6 on n vertices and insert an equipartitioned blowup of F_6 within each vertex class. We continue iteratively inserting equipartitioned blowups of F_6 into each new vertex class that we create until the vertex classes have fewer than six vertices. At this point we stop. The asymptotic density of this construction is $\frac{2}{7}$; see [22].

Let us discuss the methods used by Talbot [59] to prove his upper bound. Let $ex_k(n, F)$ denote the maximum number of edges in a k-colourable n-vertex r-graph not containing a copy of F, and let

$$\pi_k(F) = \lim_{n\to\infty} \frac{ex_k(n, F)}{\binom{n}{r}}.$$

These parameters were first introduced in [59] and are further investigated in [45]. It is shown in [59] that $\pi_2(K_4^-) < \frac{3}{10}$ and this is used to show that

$$\pi_3(K_4^-) < \frac{3 + \sqrt{11/3}}{15}.$$

This in turn is used to prove the upper bound in (6.1). Moreover, any improvement in the upper bound on $\pi_2(K_4^-)$ translates into an improvement in the upper bound on $\pi(K_4^-)$.

Below, we prove the upper bound on $\pi_2(K_4^-)$, following Talbot [59]. The proof is a variation of a simple counting argument used by de Caen [8]. Let F be a 2-colourable 3-graph containing no copy of K_4^-, where F has n vertices and m edges. Thus, we have $q_3(F) = q_4(F) = 0$. Let A and B be the colour classes of $F = (V_F, E_F)$, where $|A| = cn$ and $|B| = (1-c)n$ and $c \leq \frac{1}{2}$. For vertices x and y of F, let

$$d_{xy} = |\{z \in V_F : xyz \in E_F\}|,$$

and observe that $\sum_{xy\in V_F^{(2)}} d_{xy} = 3m$. By double counting pairs (x, abc), where $x \in V_F$ and $abc \in E_F$, we have

$$\begin{aligned}(n-3)m &= q_1(F) + 2q_2(F)\\ &= q_1(F) + 2\Big(\sum_{xy \in A^{(2)} \cup B^{(2)}} \binom{d_{xy}}{2} + \sum_{xy \in A \times B} \binom{d_{xy}}{2}\Big)\\ &= q_1(F) + \sum_{xy \in A^{(2)} \cup B^{(2)}} d_{xy}^2 + \sum_{xy \in A \times B} d_{xy}^2 - \sum_{xy \in V_F^{(2)}} d_{xy}.\end{aligned}$$

Hence we obtain

$$nm = q_1(F) + \sum_{xy \in A^{(2)} \cup B^{(2)}} d_{xy}^2 + \sum_{xy \in A \times B} d_{xy}^2.$$

Since A and B are independent sets, we have

$$\sum_{xy \in A^{(2)} \cup B^{(2)}} d_{xy} = m \quad \text{and} \quad \sum_{xy \in A \times B} d_{xy} = 2m,$$

and so, using the Cauchy-Schwarz inequality, we have

$$\begin{aligned}nm &\geq q_1(F) + \frac{m^2}{\binom{cn}{2} + \binom{(1-c)n}{2}} + \frac{4m^2}{c(1-c)n^2}\\ &\geq q_1(F) + \Big(\frac{2}{c^2 + (1-c)^2} + \frac{4}{c(1-c)}\Big)\frac{m^2}{n^2}\\ &\geq q_1(F) + 20\frac{m^2}{n^2}.\end{aligned}$$

The last line follows since $\frac{2}{c^2+(1-c)^2} + \frac{4}{c(1-c)}$ is minimised when $c = \frac{1}{2}$. Using the fact that $q_1(F) \geq 0$, we obtain Talbot's bound that $m \leq n^3/20$, that is, $\pi_2(K_4^-) \leq \frac{3}{10}$. However, we know that $q_1(F)$ cannot be zero; otherwise we would have $F \in \mathcal{C}$, and the density of F could be at most $\frac{1}{4}$. Indeed, using supersaturation (a simple counting argument to quantify a lower bound on $q_1(F)$, cf. [42]), Talbot claims (but does not explicitly prove) that the upper bound can be improved to $\pi_2(K_4^-) \leq \frac{3}{10} - 10^{-4}$. Our next theorem, a stability result, was an attempt to further improve the lower bound on $q_1(F)$, which would then improve the upper bound on $\pi_2(F)$ and filter through to an improvement in the upper bound on $\pi(K_4^-)$. However, the bounds in our result are not strong enough to accomplish this.

The result is proved in Section 6.3 by using the ideas in our new proof of Theorem 6.1.1.

Theorem 6.1.2 *Fix $\epsilon \geq 0$. Suppose $H = (V_H, E_H)$ is a 2-colourable 3-graph, where $|V_H| = n$ and where $q_i(H) < \epsilon n^4$ for $i = 1, 3, 4$. Then we can construct a hypergraph $H' = (V_H, E_{H'})$ such that $H' \in \mathcal{C}$ and*

$$E_H \triangle E_{H'} \leq 1620\epsilon^{\frac{1}{32}} n^3.$$

A general stability theorem (of the type given above) for hereditary properties of hypergraphs has recently been proved by Rödl and Schacht [51]. A special case of their result is the following: the conditions of Theorem 6.1.2 imply that $E_H \triangle E_{H'} \leq f(\epsilon)n^3$, where $f(\epsilon) \to 0$ as $\epsilon \to 0$. However, the result of Rödl and Schacht makes use of a hypergraph regularity lemma, and consequently, $f(\epsilon) \to 0$ extremely slowly as $\epsilon \to 0$.

As well as giving upper bounds on the values of $\pi_2(K_4^-)$ and $\pi_3(K_4^-)$, Talbot [59] also gives the lower bounds

$$\pi_2(K_4^-) > 0.25682 \quad \text{and} \quad \pi_3(K_4^-) > \frac{5}{18}.$$

The lower bound on $\pi_3(K_4^-)$ is achieved by the equipartitioned blowup of F_6, and Talbot [59] conjectures that this is the correct value for $\pi_3(K_4^-)$. In the final section, we give a simple counterexample to this conjecture, which shows that $\pi_3(K_4^-) \geq 0.28153$.

6.2 The Exact Result

We begin with some notation. Throughout the rest of the section, $H = (V_H, E_H)$ will be a 3-graph and $B \subseteq V_H$ will be an independent set of H, that is, a set that does not contain any edge of H. We write A for $V_H \backslash B$.

For each $x \in A$, let $E_x = \{yz \in B^{(2)} : xyz \in E_H\}$, and let G_x be the graph (B, E_x).

For each $xy \in A^{(2)}$, let $E_{xy} = \{z \in B : xyz \in E_H\}$.

Given a graph $G = (V, E)$, we write $\Gamma_G(v)$ as a shorthand for $E_G(v, V)$, the set of neighbours of v in G.

We are now ready to prove Theorem 6.1.1. The proof is written so as to facilitate easy comparison with corresponding steps in the stability result. It is therefore possible to simplify some parts of the proof.

Proof (of Theorem 6.1.1) If (a) H is 2-colourable, then let $A, B \subseteq V_H$ be the colour classes of a 2-colouring of H, and if (b) H is not 2-colourable, then let B be a maximal independent set with $A = V_H \backslash B$. For now, we assume only that B is an independent set; later we shall distinguish between the cases. We break the proof down into various claims. The proofs of the claims are straightforward, but are sometimes cumbersome to describe.

Claim 1 For each $x \in A$, G_x is a complete bipartite graph.

Proof This is easy to prove, but tedious to explain. Nevertheless, we give the details below.

Since $q_1(H) = q_3(H) = 0$, then each $rst \in B^{(3)}$ induces either zero or two edges of G_x. Fixing $rs \in E_x$, we have that $\Gamma_{G_x}(r), \Gamma_{G_x}(s)$ forms a partition of B: if not, then either there exists a vertex $t \in \Gamma_{G_x}(r) \cap \Gamma_{G_x}(s)$, in which case rst induces three edges of G_x; or there exists a vertex $t \in B \backslash (\Gamma_{G_x}(r) \cup \Gamma_{G_x}(s))$, in which case rst spans one edge of G_x. Furthermore, we have that G_x is the complete bipartite graph between $\Gamma_{G_x}(r)$ and $\Gamma_{G_x}(s)$: if not, then any edge of G_x within parts together with one of r or s spans three edges of G_x; any edge between $\Gamma_{G_x}(r)$ and $\Gamma_{G_x}(s)$ missing from G_x together with one of r or s forms a triple spanning exactly one edge of G_x. This completes the proof of the claim.

We now fix some notation. For each $x \in A$, let $B_x, \bar{B_x} \subseteq B$ be the two parts of the complete bipartite graph G_x, with $|B_x| \leq |\bar{B_x}|$. We shall find that knowing B_x for each $x \in A$ determines H completely.

For $x, y \in A$, we let $B_{xy} = B_x \cap B_y$, $B_{\bar{x}y} = \bar{B_x} \cap B_y$, $B_{x\bar{y}} = B_x \cap \bar{B_y}$, and $B_{\bar{x}\bar{y}} = \bar{B_x} \cap \bar{B_y}$. Note that these four sets form a partition of B. We write B_{ij} for a general one of the above sets, that is, we think of i and j as variables taking values from $\{x, \bar{x}\}$ and $\{y, \bar{y}\}$ respectively. When i takes the value x (resp. $\bar{x}$), we define $\bar{i}$ to take the value $\bar{x}$ (resp. x), and similarly for j. We call $B_{\bar{i}j}$ and $B_{i\bar{j}}$ the *neighbouring parts* of B_{ij}, and we call $B_{\bar{i}\bar{j}}$ the *opposing part* of B_{ij}.

Our next two claims relate E_{xy} to B_x and B_y.

Claim 2 Fix $x, y \in A$. We have the following four properties.

(i) For each i, j, we have $B_{ij} \cap E_{xy} = B_{ij}$ or ϕ.

(ii) For each i, j, if $N \neq \phi$ is a neighbour part of $B_{ij} \neq \phi$, then exactly one of $B_{ij} \cap E_{xy}$ and $N \cap E_{xy}$ is empty.

(iii) For each i, j, if $B_{ij} \cap E_{xy} \neq \phi$, then $B_{\bar{i}\bar{j}} = \phi$.

(iv) We have $|B_x| \leq |\bar{B}_x|$ and $|B_y| \leq |\bar{B}_y|$.

Proof (i) If not, then taking $r, s \in B_{ij}$ with $r \in E_{xy}$ and $s \notin E_{xy}$, we have that $xyrs$ spans one edge of H (namely xyr), a contradiction.

(ii) If both $B_{ij} \cap E_{xy}$ and $N \cap E_{xy}$ are empty, then taking $r \in B_{ij}$ and $s \in N$, we find that $xyrs$ spans exactly one edge of H (namely either xrs or yrs depending on which neighbour part N is), a contradiction. If both $B_{ij} \cap E_{xy}$ and $N \cap E_{xy}$ are nonempty, then taking $r \in B_{ij} \cap E_{xy}$ and $s \in N \cap E_{xy}$, we find that $xyrs$ spans three edges of H (namely xyr, xys, and one of xrs or yrs depending on which neighbour part N is), a contradiction.

(iii) If not, then $B_{ij} \cap E_{xy} \neq \phi$ and $B_{\bar{i}\bar{j}} \neq \phi$. Taking $r \in B_{ij} \cap E_{xy}$ and $s \in B_{\bar{i}\bar{j}}$, we find that $xyrs$ spans at least three edges of H (namely xrs, yrs, and xyr), a contradiction.

(iv) This is true by our choice of B_x, B_y, but we emphasise this here because the corresponding step in the stability result needs a proof.

This completes the proof of Claim 2.

From Claim 2, we deduce the following.

Claim 3 For each $x, y \in A$, if B_x and B_y are both nonempty, then one of the following three conditions holds. Either

(a) $B_x \subseteq B_y$ and $E_{xy} = B_y \backslash B_x = B_{\bar{x}y}$; or

(b) $B_y \subseteq B_x$ and $E_{xy} = B_x \backslash B_y = B_{x\bar{y}}$; or

(c) B_x, B_y are disjoint and $E_{xy} = B_{\bar{x}\bar{y}}$.

If $B_x = \phi$ or $B_y = \phi$, then at least one of conditions (a), (b), or (c) holds.

Proof We assume for the moment that B_x and B_y are both nonempty; we consider only at the end the case when one or both sets are empty. First we show that either $B_x \subseteq B_y$, or $B_y \subseteq B_x$, or B_x, B_y are disjoint. Suppose not. Then B_{xy}, $B_{\bar{x}y}$, and $B_{x\bar{y}}$ are all nonempty. We cannot have $B_{\bar{x}\bar{y}} = \phi$, otherwise condition (iv) of Claim 2 is violated; hence B_{xy}, $B_{\bar{x}y}$, $B_{x\bar{y}}$, and

$B_{\bar{x}\bar{y}}$ are all nonempty. Now properties (i) and (ii) of Claim 2 imply that there exists B_{ij} such that $B_{ij} \cap E_{xy} = B_{ij}$. Then property (iii) of Claim 2 implies that $B_{\bar{i}\bar{j}} = \phi$, contradicting that all four parts are nonempty. Thus, either $B_x \subseteq B_y$, or $B_y \subseteq B_x$, or B_x, B_y are disjoint, as required.

Next, we show that E_{xy} is as stated for each case. If $B_x \subsetneq B_y$, then $B_{x\bar{y}} = \phi$ and the other three parts are nonempty ($B_{\bar{x}\bar{y}} \neq \phi$ by property (iv) of Claim 2). Now properties (i), (ii), and (iii) force that $E_{xy} = B_{\bar{x}y} = B_y \backslash B_x$, as required. Similarly for the case $B_y \subsetneq B_x$. If $B_x = B_y \neq \phi$, then $B_{\bar{x}y} = \phi$ and $B_{x\bar{y}} = \phi$. Also $B_{xy} \neq \phi$ and $B_{\bar{x}\bar{y}} \neq \phi$ (by property (iv) of Claim 2). Properties (i) and (iii) of Claim 2 force that $E_{xy} = \phi = B_y \backslash B_x = B_x \backslash B_y$, as required.

If B_x, B_y are disjoint and $B_x \cup B_y \neq B$, then $B_{xy} = \phi$ and the other three parts are nonempty. Properties (i), (ii), and (iii) of Claim 2 force that $E_{xy} = B_{\bar{x}\bar{y}} = B \backslash (B_x \cup B_y)$, as required. If B_x, B_y are disjoint and $B_x \cup B_y = B$, then $B_{xy} = B_{\bar{x}\bar{y}} = \phi$ and the other two parts are nonempty. Properties (i) and (iii) of Claim 2 force that $E_{xy} = \phi = B \backslash (B_x \cup B_y)$, as required.

Finally, suppose that one of B_x or B_y is empty; without loss of generality $B_x = \phi$ and $B_y \neq \phi$. Then $B_{xy} = B_{x\bar{y}} = \phi$, and by properties (i) and (ii) of Claim 2, we have that either $E_{xy} = B_{\bar{x}y}$, or $E_{xy} = B_{\bar{x}\bar{y}}$, as required. If $B_x = B_y = \phi$, then $B_{\bar{x}\bar{y}} = B$ and the other three parts are empty. Thus, condition (i) of Claim 2 implies that either $E_{xy} = B = B_{\bar{x}\bar{y}}$ or $E_{xy} = \phi = B_{\bar{x}y} = B_{x\bar{y}}$, as required. This completes the proof of Claim 3.

We see that when $B_x = \phi$ or $B_y = \phi$ (without loss of generality $B_x = \phi$), we have the degenerate situation where $B_x \subseteq B_y$ and B_x, B_y are disjoint. In order to keep this case consistent with the conditions of Claim 3, we declare that $B_x \subseteq B_y$ if $E_{xy} = B_{\bar{x}y}$; $B_y \subseteq B_x$ if $E_{xy} = B_{x\bar{y}}$; and B_x, B_y are disjoint if $E_{xy} = B_{\bar{x}\bar{y}}$.

Next we shall see how the three sets B_x, B_y, B_z are related for three vertices $x, y, z \in A$. At this point, we distinguish between cases (a) and (b) of the Theorem. For case (a), we have that A is an independent set. We define a colouring on the elements of $A^{(2)}$, which we call the *containment colouring.* Given $xy \in A^{(2)}$, we colour xy red if $B_x \subseteq B_y$ or $B_y \subseteq B_x$, and we colour xy blue if B_x, B_y are disjoint. Thus every element of $A^{(2)}$ is coloured and we have the following.

Claim 4 For every triple $xyz \in A^{(3)}$, either xyz is a red triangle, or two of the edges are blue and the other is red.

Proof We observe that for each $r \in B$, r cannot be in exactly one or exactly three of the sets E_{xy}, E_{yz}, and E_{xz}, otherwise $xyzr$ spans exactly one or exactly three edges of H.

If xyz is a blue triangle, then B_x, B_y, and B_z are pairwise disjoint. If one of these sets is nonempty (B_x without loss of generality), then taking $r \in B_x$, we have that $r \in E_{yz}$ and $r \notin E_{xy}, E_{xz}$ by Claim 3. Hence $xyzr$ spans exactly one edge of H (namely ryz), a contradiction. If all three of B_x, B_y, and B_z are empty, then $E_{xy} = E_{xz} = E_{yz} = B$. Hence taking any $r \in B$, we have that $xyzr$ spans three edges of H (namely rxy, rxz, and ryz), a contradiction.

If xyz consists of two red edges and one blue edge, then without loss of generality, let xy and xz be the red edges. Thus B_y, B_z are disjoint. Using Claim 3, for any $r \in B \backslash (B_x \cup B_y \cup B_z)$, we have that $r \in E_{yz}$ and $r \notin E_{xy}, E_{xz}$; thus $xyzr$ spans exactly one edge of H, a contradiction. If $B \backslash (B_x \cup B_y \cup B_z) = \phi$, then it must be the case that $B_y \sqcup B_z = B$ and $B_x = \phi$ (recall that $|B_x| \leq \frac{1}{2}|B|$). In this case, we have that $E_{yz} = \phi$ and $E_{xy} = B_y$, $E_{xz} = B_z$. Thus each $r \in B$ is contained in exactly one of E_{xy}, E_{xz}, and E_{yz}, so that $xyzr$ spans exactly one edge of H, a contradiction.

Therefore the only possibilities remaining are those stated in the claim. This completes the proof of Claim 4.

Consider the graph G_b induced on A by the blue edges of the containment colouring. By Claim 4, any triple $xyz \in A^{(3)}$ induces either zero or two edges of G_b. Thus, by the proof of Claim 1, G_b is a complete bipartite graph. Let $A_1, A_2 \subseteq A$ be the parts of this graph. Let

$$\mathcal{A}_1 = \{B_x : x \in A_1\} \text{ and } \mathcal{A}_2 = \{B_x : x \in A_2\}$$

Each of $\mathcal{A}_1$ and $\mathcal{A}_2$ are nested families, and each set in $\mathcal{A}_1$ is disjoint from each set in $\mathcal{A}_2$. For each $x \in A$, define

$$C_x = \begin{cases} B_x & \text{if } x \in A_1; \\ \bar{B}_x & \text{if } x \in A_2. \end{cases}$$

It is clear that the C_x form a nested family. Moreover, for each $x \in A$, E_x

is the set of edges between C_x and $\bar{C_x} := B \backslash C_x$, and for each $xy \in A^{(2)}$, if $C_x \subseteq C_y$, then $E_{xy} = C_y \backslash C_x = C_x \bigtriangleup C_y$ (as a consequence of Claim 3). Thus, H is completely determined by the sets $(C_x)_{x \in A}$.

These properties of H imply that $H \in \mathcal{C}$. Indeed, let $x_1, \ldots, x_k$ be the vertices of A ordered such that $C_{x_1} \subseteq \ldots \subseteq C_{x_k}$, and let $y_1, \ldots, y_l$ be the vertices of B ordered such that each C_{x_i} is an initial segment of $y_1, \ldots, y_l$. For each $1 \leq i \leq k$, define $r(i)$ such that $C_{x_i} = \{y_1, \ldots, y_{r(i)}\}$. Arrange $y_1, \ldots, y_l$ in order on one half of the circumference of a circle, and place each x_i diametrically opposite a point on the circumference between $y_{r(i)}$ and $y_{r(i)+1}$. Now we see that if $x_i \in A$ and $p, q \in B$, then the convex hull of $\{x_i, p, q\}$ contains the origin of the circle if and only if one of p or q is in $C_{x_i} = \{y_1, \ldots, y_{r(i)}\}$ and the other is in $\bar{C}_{x_i} = \{y_{r(i)+1}, \ldots, y_l\}$. Also for $x_i, x_j \in A$ with $i < j$ and $p \in B$, the convex hull of $\{x_i, x_j, p\}$ contains the origin of the circle if and only if $p \in \{y_{r(i)+1}, \ldots, y_{r(j)}\} = C_{x_j} \backslash C_{x_i}$. This corresponds exactly to the edges of H.

For (b), we assume that B is a maximal independent set; hence for each $x \in A$, we have that $B_x \neq \phi$ (if not, then $E_x = \phi$, and we can add x to B). Recall that A is not an independent set. We have the following claim.

Claim 5 We have that xyz is an edge of H in A if and only if, either

(i) $B = B_x \sqcup B_y \sqcup B_z$; or

(ii) without loss of generality $B_x = B_y \sqcup B_z$.

(These two cases are essentially the same: it is simply because we choose to have $|B_x| \leq \frac{1}{2}|B|$ that creates case (ii) when $|B_y \sqcup B_z| \leq \frac{1}{2}|B|$.)

Proof Suppose xyz is an edge of H in A. For every $r \in B$, we have that $xyzr$ spans exactly two edges of H. Therefore r is an element of exactly one of E_{xy}, E_{xz}, or E_{yz}, and so $B = E_{xy} \sqcup E_{xz} \sqcup E_{yz}$. From Claim 3, any pair of sets from B_x, B_y, B_z are nested or disjoint. We cannot have all three sets nested; if $B_x \subseteq B_y \subseteq B_z$, then by claim 3, we find that every element of $B_x \neq \phi$ is not an element of $E_{xy} \sqcup E_{xz} \sqcup E_{yz}$, a contradiction. Thus we may assume, without loss of generality, that B_y, B_z are disjoint. Then, by Claim 3, either

(a) B_x is also disjoint from B_y and B_z; or

(b) $B_y, B_z \subseteq B_x$.

For (a), using Claim 3, we have that $B \backslash (B_x \cup B_y \cap B_z) = E_{xy} \cap E_{xz} \cap E_{yz} (= \phi)$; hence $B = B_x \sqcup B_y \sqcup B_z$, giving us case (i) of the claim conclusion. For (b), using Claim 3, we have that $B_x \backslash (B_y \cup B_z) = E_{xz} \cap E_{xy} (= \phi)$; hence $B_x = B_y \sqcup B_z$, giving us case (ii) of the claim conclusion.

For the converse, assume that $xyz \in A^{(3)}$ and that either case (i) or (ii) from the claim statement holds. Using Claim 3, we find that for both cases, we have $B = E_{xy} \sqcup E_{xz} \sqcup E_{yz}$. Thus, taking any $r \in B$, we find that $xyzr$ spans exactly one edge of H unless xyz is an edge of H. This completes the proof of the claim.

Note: Claim 5 tells us that if xyz is an edge of H in A and we know B_y and B_z, then B_x is uniquely determined (except for the ambiguity that arises when $|B_x| = \frac{1}{2}|B|$).

Let XYZ be a fixed edge of H in A. For each $a \in A$, we have that $aXYZ$ must span exactly two edges of H. Hence, one of aXY, aXZ, or aYZ is an edge of H, and so by Claim 5, it must be (respectively) the case that either $B_a = B_Z$, $B_a = B_Y$, or $B_a = B_X$ (one might have to change the choice of B_a if $|B_a| = \frac{1}{2}|B|$). Let

$$A_X = \{a \in A : B_a = B_X\},$$

and define A_Y and A_Z similarly. We have $A = A_X \sqcup A_Y \sqcup A_Z$, and by Claim 5, the edges of H contained in A are precisely those triples that have exactly one vertex in each of A_X, A_Y, and A_Z. Knowing $A_X, A_Y, A_Z, B_X, B_Y, B_Z$ determines H completely: indeed, we know that $A_X A_Y A_Z$ gives all the edges in A and that B is an independent set, and the edges between A and B are determined by the sets $(B_x)_{x \in A}$ (all of which are known by our choice of A_X, A_Y, A_Z) via Claim 1 and Claim 3.

Finally, let us check that the structural description of H we have deduced implies that H is a blowup of F_6. From Claim 5, although there appear to be two possibilities to consider, they turn out to be essentially the same:

(i) $B = B_X \sqcup B_Y \sqcup B_Z$; and

(ii) $B_X = B_Y \sqcup B_Z$.

For case (i), we have that E_H is the union of the following sets of edges:

$$A_X A_Y A_Z,$$
$$A_X B_X \bar{B_X},\ A_Y B_Y \bar{B_Y},\ A_Z B_Z \bar{B_Z},$$
$$A_X A_Y B_Z,\ A_Y A_Z B_X,\ A_Z A_X B_Y;$$

or equivalently, E_H is the union of the following sets of edges:

$$A_X A_Y A_Z,$$
$$A_X B_X B_Y,\ A_X B_X B_Z,\ A_Y B_Y B_X,\ A_Y B_Y B_Z,\ A_Z B_Z B_X,\ A_Z B_Z B_Y,$$
$$A_X A_Y B_Z,\ A_Y A_Z B_X,\ A_Z A_X B_Y.$$

Setting $V_1, \ldots, V_6$ equal to $A_X, A_Y, A_Z, B_X, B_Z, B_Y$ respectively shows that H is a blowup of F_6. For case (ii), if we replace B_X with $\bar{B_X}$ (so that B_X, B_Y, B_Z partition B), then the edges of H are precisely as described above. □

6.3 The Stability Result

Throughout this section, H is a hypergraph satisfying the hypothesis of Theorem 6.1.2, that is, $H = (V_H, E_H)$ is a 2-colourable 3-graph, where $|V_H| = n$ and $q_i(H) < \epsilon n^4$ for $i = 1, 3, 4$. We make no attempt to optimise the bounds in this section since we have found that even the most optimistic bounds using our method would not be sufficient for the purpose of improving the upper bound on $\pi(K_4^-)$. Therefore, our goal in this section is simply to demonstrate our method.

Let A and B be the colour classes of H, where $|A| = cn$, $|B| = (1-c)n$, and $c \leq \frac{1}{2}$. Note that there are at most $c(1-c)n^3$ edges in H. Therefore, we may assume that

$$c(1-c) > 1620\epsilon^{\frac{1}{32}}; \tag{6.2}$$

otherwise we can delete all edges of H to leave the empty hypergraph H', which satisfies the conclusion of Theorem 6.1.2.

Let $Q_i(H)$ be the set of $rstu \in V_H^{(4)}$ spanning exactly i edges; thus

$q_i(H) = |Q_i(H)|$. For $x, y, z \in A$, let

$$\begin{aligned} Q_i(x) &= \{rst \in B^{(3)} : xrst \in Q_i(H)\}, \\ Q_i(xy) &= \{rs \in B^{(2)} : xyrs \in Q_i(H)\}, \text{ and} \\ Q_i(xyz) &= \{r \in B : xyzr \in Q_i(H)\}. \end{aligned}$$

Let $q_i(x)$, $q_i(xy)$, and $q_i(xyz)$ be the respective sizes of these sets. Thus we have that

$$\sum_{x \in A} q_i(x) + \sum_{xy \in A^{(2)}} q_i(xy) + \sum_{xyz \in A^{(3)}} q_i(xyz) = q_i(H) < \epsilon n^4,$$

for $i = 1, 3, 4$. It is clear, since H is 2-colourable, that $q_4(x) = 0$ for all $x \in A$. Setting $\delta_1 = \sqrt{\epsilon}$, we have that more than $(1 - \delta_1)cn$ vertices $x \in A$ satisfy

$$q_1(x) < \frac{\epsilon n^4}{\delta_1(cn)} = \frac{\epsilon}{\delta_1 c} n^3 = \epsilon_1((1-c)n)^3,$$

where

$$\epsilon_1 = \frac{\epsilon}{\delta_1 c(1-c)^3} = \frac{\sqrt{\epsilon}}{c(1-c)^3}. \tag{6.3}$$

Similarly,

$$q_3(x) < \epsilon_1((1-c)n)^3$$

for more than $(1 - \delta_1)cn$ vertices of A. Therefore, more than $(1 - 2\delta_1)cn$ vertices of A satisfy

$$q_1(x) < \epsilon_1((1-c)n)^3 \text{ and } q_3(x) < \epsilon_1((1-c)n)^3.$$

Such vertices of A are referred to as *good* vertices. Any vertex of A that is not good is referred to as a *bad* vertex, and so we have

$$|\{x \in A : x \text{ bad }\}| < 2\delta_1 cn =: \eta_1 cn. \tag{6.4}$$

Our first lemma is an analogue of Claim 1 (from the proof of Theorem 6.1.1) for good vertices. We show that if $x \in A$ is a good vertex, then $G_x = (B, E_x)$ is close to a complete bipartite graph. We prove the lemma in a general form. First we need some notation.

Let $G = (V, E)$ be a graph. For $v \in V$, recall that $\Gamma(v)$ is the set of neighbours of v in G. For $i = 1, 2, 3$, define

$$T_i(G) = \{xyz \in V^{(3)} : |\{xy, yz, xz\} \cap E| = i\},$$

so for example, $T_3(G)$ is the set of triangles in G. Let $t_i(G) = |T_i(G)|$. Note that for the graph G_x, we have $t_i(G_x) = q_i(x)$.

Lemma 6.3.1 *Suppose that $G = (V, E)$ is a graph with $|V| = n$ and $|E| = m$. Suppose that $t_i(G) \leq tn^3$ for $i = 1, 3$. Then there is a complete bipartite graph $G' = (V, E')$ with $|E \triangle E'| \leq 3\sqrt{t}n^2$.*

Proof Given an edge $e = uv \in E$, we define

$$V_1^e = \Gamma(u) \backslash \Gamma(v), \qquad V_2^e = \Gamma(v) \backslash \Gamma(u),$$
$$R_1^e = \Gamma(u) \cap \Gamma(v), \text{ and } \qquad R_2^e = V \backslash \Gamma(u) \cup \Gamma(v).$$

Note that these sets form a partition of V. We show that on average (over all edges $e \in E$), R_1^e and R_2^e are small. Further, we show that on average, the number of edges of G within V_1^e and V_2^e is small, and that the number of edges of G missing between V_1^e and V_2^e is small. Thus for a suitably chosen edge e, the complete bipartite graph between $V_1^e \cup R_1^e \cup R_2^e$ and V_2^e is close to G.

It is easy to see that

$$\sum_{e \in E} |R_1^e| = 3t_3(G) \leq 3tn^3; \tag{6.5}$$

$$\sum_{e \in E} |R_2^e| = t_1(G) \leq tn^3. \tag{6.6}$$

For $uv = e \in E$, define

$$F(e) = \{ab \in V^{(2)} \backslash E : a \in V_1^e, b \in V_2^e \text{ or } b \in V_1^e, a \in V_2^e\},$$

and let $f(e) = |F(e)|$. We estimate the average of $f(e)$ by double counting induced paths of length three.

Let P_3 be the set of $(uv, ab) \in E \times V^{(2)}$ for which $auvb$ induces a path of length three and uv is the middle edge. Then $(uv, ab) \in P_3$ if and only if

$ab \in F(uv)$; thus

$$\sum_{e \in E} f(e) = |P_3|.$$

Suppose that $aub \in T_1(G)$ with $au \in E$. Then $(uv, ab) \in P_3$ if and only if $v \in (\Gamma(u) \cap \Gamma(b)) \backslash \Gamma(a)$. Thus

$$\sum_{e \in E} f(e) = |P_3| = \sum_{\substack{aub \in T_1(G) \\ au \in E}} |(\Gamma(u) \cap \Gamma(b)) \backslash \Gamma(a)| \leq \sum_{\substack{aub \in T_1(G) \\ au \in E}} |V| \leq tn^4. \quad (6.7)$$

For $uv = e \in E$, define

$$G(e) = \{ab \in E : a, b \in V_1^e \text{ or } a, b \in V_2^e\},$$

and let $g(e) = |G(e)|$. We estimate the average of $g(e)$ by double counting induced copies of F_4. We define F_4 to be the graph on four vertices $v_1, \ldots, v_4$ with edge set $\{v_1v_2, v_1v_3, v_2v_3, v_1v_4\}$. Thus $v_1v_2v_3$ is a triangle, and we refer to v_1v_4 as the *hanging edge* and to v_4 as the *hanging vertex.*

Let Q be the set of $(uv, ab) \in E \times V^{(2)}$ such that $uvab$ induces a copy of F_4 in G and uv is the hanging edge. We have that $(uv, ab) \in Q$ if and only if $ab \in G(uv)$. Also, for $abc \in T_3(G)$, we have that $xabc$ induces a copy of F_4 if and only if x is an element of exactly one of $\Gamma(a)$, $\Gamma(b)$, or $\Gamma(c)$. Thus, we have

$$\sum_{e \in E} g(e) = |Q| \leq \sum_{abc \in T_3(G)} |\Gamma(a) \cup \Gamma(b) \cup \Gamma(c)| \leq \sum_{abc \in T_3(G)} |V| \leq tn^4. \quad (6.8)$$

Fix an edge $e \in E$. The number of edges we have to change in order that G becomes complete bipartite with parts $V_1^e \cup R_1^e \cup R_2^e$ and V_2^e is at most

$$f(e) + g(e) + n(|R_1^e| + |R_2^e|).$$

Using (6.5), (6.6), (6.7), and (6.8), we have that

$$\sum_{e \in E} f(e) + g(e) + n(|R_1^e| + |R_2^e|) \leq 6tn^4,$$

and so there exists an edge e^* for which

$$f(e^*) + g(e^*) + n(|R_1^{e^*}| + |R_2^{e^*}|) \leq \frac{6tn^4}{m}.$$

If $m \leq 3\sqrt{t}n^2$, then delete all edges to obtain the empty graph, which is complete bipartite. Otherwise, apply the argument above, changing at most

$$\frac{6tn^4}{3\sqrt{t}n^2} < 3\sqrt{t}n^2$$

suitable edges to obtain a complete bipartite graph. $\square$

If we apply the above lemma to the graph G_x, where x is a good vertex of A, we obtain the following corollary.

Corollary 6.3.2 *If x is a good vertex of A, then by changing at most $3\sqrt{\epsilon_1}|B|^2$ suitable edges of G_x (as in Lemma 6.3.1), we obtain a complete bipartite graph.*

Proof Recall that if x is good then $t_i(G_x) = q_i(x) < \epsilon_1((1-c)n)^3$ for $i = 1, 3$. Now apply Lemma 6.3.1. $\square$

If $x \in A$ is good, we define B_x, $\bar{B}_x$ to be the parts of the complete bipartite graph given by the above corollary, where $|B_x| \leq |\bar{B}_x|$.

Let us assume that $|A| = cn > 1$, so that $\binom{cn}{2} \geq (cn)^2/4$. Let $\delta_2 = \sqrt{\epsilon}$. We say $xy \in A^{(2)}$ is *good* if x and y are good, and for $i = 1, 3, 4$, we have

$$\begin{aligned} q_i(xy) &< \frac{\epsilon n^4}{\delta_2\binom{cn}{2}} \\ &\leq \frac{4\epsilon}{\delta_2 c^2(1-c)^2}((1-c)n)^2 = \epsilon_2((1-c)n)^2, \end{aligned} \tag{6.9}$$

where

$$\epsilon_2 = \frac{4\epsilon}{\delta_2 c^2(1-c)^2} = \frac{4\sqrt{\epsilon}}{c^2(1-c)^2}. \tag{6.10}$$

For each fixed $i = 1, 3, 4$, the above inequality (6.9) holds for more than $(1-\delta_2)\binom{cn}{2}$ pairs in $A^{(2)}$, where we have used the fact that

$$\sum_{xy \in A^{(2)}} q_i(xy) < \epsilon n^4 \quad \text{for } i = 1, 3, 4.$$

Thus, more than $(1 - 3\delta_2)\binom{cn}{2}$ pairs in $A^{(2)}$ satisfy (6.9) for $i = 1, 3, 4$ simultaneously. Since, by (6.4), fewer than $\eta_1(cn) = 2\delta_1(cn)$ vertices of A are bad, we have that

$$|\{xy \in A^{(2)} : xy \text{ bad}\}| < 3\delta_2\binom{cn}{2} + 2\delta_1(cn)^2 \leq \left(2\delta_1 + \frac{3}{2}\delta_2\right)(cn)^2,$$

and since $\delta_1 = \delta_2 = \sqrt{\epsilon}$, we have

$$|\{xy \in A^{(2)} : xy \text{ bad}\}| < \frac{7}{2}\epsilon^{\frac{1}{2}}(cn)^2 =: \eta_2(cn)^2. \qquad (6.11)$$

For a good pair $xy \in A^{(2)}$, we let $B_{xy} = B_x \cap B_y$, $B_{\bar{x}y} = \bar{B}_x \cap B_y$, $B_{x\bar{y}} = B_x \cap \bar{B}_y$, and $B_{\bar{x}\bar{y}} = \bar{B}_x \cap \bar{B}_y$, just as in the proof of Theorem 6.1.1. As before, we write B_{ij} for a general one of the above sets, that is, we think of i and j as variables taking values from $\{x, \bar{x}\}$ and $\{y, \bar{y}\}$ respectively. When i takes the value x (resp. $\bar{x}$), we define $\bar{i}$ to take the value $\bar{x}$ (resp. x), and similarly for j. Recall that $B_{\bar{i}j}$ and $B_{i\bar{j}}$ are the *neighbouring parts* of B_{ij}.

We have the following lemma, which is the approximation corresponding to Claim 2 of Theorem 6.1.1.

Lemma 6.3.3 *Fix a good pair $xy \in A^{(2)}$ and set $r = (6\sqrt{\epsilon_1} + \epsilon_2)^{1/2}$. Then there exists a set $R(xy) \subseteq B$ satisfying*

$$|R(xy)| \leq 27r|B|$$

such that the following holds. Define $B^* = B \backslash R(xy)$, and let B^*_x, $\bar{B}^*_x$, B^*_y, $\bar{B}^*_y$ be respectively the intersection of B_x, $\bar{B}_x$, B_y, $\bar{B}_y$ with B^*. Let $B^*_{ij} = B_{ij} \cap B^*$ for $i \in \{x, \bar{x}\}$, $j \in \{y, \bar{y}\}$. We have the following properties for the starred sets, which are identical to those of Claim 2 of Theorem 6.1.1.

(i) For each i, j, we have $B^*_{ij} \cap E_{xy} = B^*_{ij}$ or ϕ.

(ii) For each i, j, if $N \neq \phi$ is a neighbouring part of $B^*_{ij} \neq \phi$, (that is, $N = B^*_{\bar{i}j}$ or $N = B^*_{i\bar{j}}$) then exactly one of $B^*_{ij} \cap E_{xy}$ and $N \cap E_{xy}$ is empty.

(iii) For each i, j, if $B^*_{ij} \cap E_{xy} \neq \phi$, then $B^*_{\bar{i}\bar{j}} = \phi$.

(iv) We have $|B^*_x| \leq |\bar{B}^*_x|$ and $|B^*_y| \leq |\bar{B}^*_y|$.

We bound r in terms of ϵ for later. Using (6.3) and (6.10), we have

$$\begin{aligned} r = (6\sqrt{\epsilon_1} + \epsilon_2)^{\frac{1}{2}} \leq \sqrt{6}\epsilon_1^{\frac{1}{4}} + \epsilon_2^{\frac{1}{2}} &\leq 3\Big(\frac{\sqrt{\epsilon}}{c(1-c)^3}\Big)^{\frac{1}{4}} + \Big(\frac{4\sqrt{\epsilon}}{c^2(1-c)^2}\Big)^{\frac{1}{2}} \\ &\leq \frac{3\epsilon^{\frac{1}{8}}}{c(1-c)} + \frac{2\epsilon^{\frac{1}{8}}}{c(1-c)}; \end{aligned}$$

thus

$$r \leq \frac{5\epsilon^{\frac{1}{8}}}{c(1-c)}. \tag{6.12}$$

Proof We start by finding elements whose removal ensures condition (i) of the lemma.

Write $\bar{E}_{xy}$ for $B \backslash E_{xy}$. For $r \in B_{ij} \cap E_{xy}$ and $s \in B_{ij} \cap \bar{E}_{xy}$, we have $xyrs \in Q_1(xy)$ unless $rs \in E_x$ or E_y (or both). By Corollary 6.3.2, we know that at most $3\sqrt{\epsilon_1}|B|^2$ edges of E_x and $3\sqrt{\epsilon_1}|B|^2$ edges of E_y lie within B_{ij} for each i and j. Hence

$$|B_{ij} \cap E_{xy}||B_{ij} \cap \bar{E}_{xy}| - 2(3\sqrt{\epsilon_1})|B|^2 \leq q_1(xy) < \epsilon_2|B|^2.$$

Thus

$$|B_{ij} \cap E_{xy}||B_{ij} \cap \bar{E}_{xy}| < (6\sqrt{\epsilon_1} + \epsilon_2)|B|^2,$$

and so at least one of $B_{ij} \cap E_{xy}$ and $B_{ij} \cap \bar{E}_{xy}$ has strictly fewer than $(6\sqrt{\epsilon_1} + \epsilon_2)^{1/2}|B| = r|B|$ elements. The elements of the smaller set are discarded into a residual set $R^1(xy)$. This process needs to be carried out sequentially for all i, j, and each time, the discarded elements are added to $R^1(xy)$. By the end, we can only have discarded at most $4r|B|$ elements; hence $|R^1(xy| \leq 4r|B|$. We define $B^1 = B \backslash R^1(xy)$ and we define $B^1_x = B_x \cap B^1$, etc. We see that the B^1_{ij} satisfy condition (i) of the lemma.

Next, we remove elements from the B^1_{ij} so that condition (ii) of the lemma is also satisfied.

Suppose $N \neq \phi$ is a neighbour part of $B^1_{ij} \neq \phi$. We shall assume that $N = B^1_{\bar{i}j}$; the case $N = B^1_{i\bar{j}}$ is similar. If $B^1_{ij} \cap E_{xy}$ and $B^1_{\bar{i}j} \cap E_{xy}$ are both nonempty (resp. empty), then by part (i), $B^1_{ij} \cap E_{xy} = B^1_{ij}$ (resp. ϕ) and $B^1_{\bar{i}j} \cap E_{xy} = B^1_{\bar{i}j}$ (resp. ϕ). By Corollary 6.3.2, we know that all but at most $3\sqrt{\epsilon_1}|B|^2$ edges between B_x and $\bar{B}_x$ are elements of E_x. One of B^1_{ij} and $B^1_{\bar{i}j}$ is a subset of B_x and the other is a subset of $\bar{B}_x$; thus all but at most $3\sqrt{\epsilon_1}|B|^2$ edges between B^1_{ij} and $B^1_{\bar{i}j}$ are in E_x. Furthermore $B^1_{ij} \cup B^1_{\bar{i}j} \subseteq B_y$ or $\bar{B}_y$, and so again by Corollary 6.3.2, at most $3\sqrt{\epsilon_1}|B|^2$ edges of E_y go between B^1_{ij} and $B^1_{\bar{i}j}$. Hence all but at most $6\sqrt{\epsilon_1}|B|^2$ edges between B^1_{ij} and $B^1_{\bar{i}j}$ are elements of $E_x \backslash E_y$.

Taking $rs \in E_x \backslash E_y$ with $r \in B^1_{ij}$ and $s \in B^1_{\bar{i}j}$, we have that $xyrs \in$

$Q_3(xy)$ (resp. $Q_1(xy)$). Thus

$$|B^1_{ij}||B^1_{\bar{i}j}| - 6\sqrt{\epsilon_1}|B|^2 \leq q_{3\ (\text{resp. } 1)}\ (xy) < \epsilon_2|B|^2,$$

and so

$$|B^1_{ij}||B^1_{\bar{i}j}| < (6\sqrt{\epsilon_1} + \epsilon_2)|B|^2.$$

Thus, the smaller of B^1_{ij} and $B^1_{\bar{i}j}$ has strictly fewer than $r|B|$ elements. The elements of the smaller set are discarded into a residual set $R^2(xy)$. This process needs to be carried out sequentially for all pairs of nonempty neighbour parts, and each time, the discarded elements are added to $R^2(xy)$. By the end, we can only have discarded at most three of the B^1_{ij}; hence $|R^2(xy)| \leq 3r|B|$. We define $B^2 = B^1\backslash R^2(xy)$ and we define $B^2_x = B^1_x \cap B^2$, etc. We see that the B^2_{ij} satisfy conditions (i) and (ii) of the lemma.

Next, we show that the removal of a small number of elements from B^2 ensures that condition (iii) is satisfied.

Suppose that $B^2_{ij} \cap E_{xy} \neq \phi$ and $B^2_{\bar{i}\bar{j}} \neq \phi$ for some i, j. Thus, $B^2_{ij} \cap E_{xy} = B^2_{ij}$. We know that one of B^2_{ij} and $B^2_{\bar{i}\bar{j}}$ is a subset of B_x, and the other is a subset of $\bar{B}_x$. Similarly, one is a subset of B_y and the other is a subset of $\bar{B}_y$. As before, using Corollary 6.3.2, we know that all but at most $3\sqrt{\epsilon_1}|B|^2$ edges between B^2_{ij} and $B^2_{\bar{i}\bar{j}}$ are elements of E_x, and likewise, all but at most $3\sqrt{\epsilon_1}|B|^2$ edges between B^2_{ij} and $B^2_{\bar{i}\bar{j}}$ are elements of E_y. Therefore, all but at most $6\sqrt{\epsilon_1}|B|^2$ edges between B^2_{ij} and $B^2_{\bar{i}\bar{j}}$ are elements of $E_x \cap E_y$.

Taking $r \in B^2_{ij}$ and $s \in B^2_{\bar{i}\bar{j}}$ with $rs \in E_x \cap E_y$, we find that if $B^2_{\bar{i}\bar{j}} \cap E_{xy} = \phi$, then $xyrs \in Q_3(xy)$, and if $B^2_{\bar{i}\bar{j}} \cap E_{xy} = B^2_{\bar{i}\bar{j}}$, then $xyrs \in Q_4(xy)$. Since $q_3(xy), q_4(xy) < \epsilon_2|B|^2$, we have that

$$|B^2_{ij}||B^2_{\bar{i}\bar{j}}| - 6\sqrt{\epsilon_1}|B|^2 < \epsilon_2|B|^2,$$

and therefore

$$|B^2_{ij}||B^2_{\bar{i}\bar{j}}| < 6\sqrt{\epsilon_1}|B|^2 + \epsilon_2|B|^2.$$

Thus the smaller of B^2_{ij} and $B^2_{\bar{i}\bar{j}}$ has size strictly smaller than $r|B|$. The smaller of the sets is discarded by adding its elements to a residual set $R^3(xy)$, and this is done for each i, j. By the end, we can only have discarded at most two of the B^2_{ij}, and therefore, $|R^3(xy)| \leq 2r|B|$. We set $B^3 = B^2\backslash R^3(xy)$, and define $B^3_x = B_x \cap B^3$, etc. Thus the B^3_{ij} satisfy conditions (i), (ii), and (iii) of the lemma.

Finally, we remove a small number of elements from B_x^3 and B_y^3 to ensure condition (iv) holds. So far, we have removed a total of at most $9r|B|$ elements from $|B|$. Since we know that $|B_x| \leq |\bar{B}_x|$ and $|B_y| \leq |\bar{B}_y|$, then we have that

$$|B_x^3| \leq |\bar{B}_x^3| + 9r|B| \text{ and } |B_y^3| \leq |\bar{B}_y^3| + 9r|B|.$$

By removing $9r|B|$ elements from each of B_x^3 and B_y^3 and adding them to a residual set $R^4(xy)$, we have that $|R^4(xy)| \leq 18r|B|$. Let $B^4 = B^3 \backslash R^4(xy)$, and let $B_x^4 = B_x \cap B^4$, etc. Then $|B_x^4| \leq |\bar{B}_x^4|$ and $|B_y^4| \leq |\bar{B}_y^4|$.

Finally, set

$$R(xy) = \bigcup_{i=1}^{4} R^i(xy).$$

Note that $|R(xy)| \leq 27r|B|$. Recalling that $B^* = B \backslash R(xy)$, and that $B_{ij}^* = B_{ij} \cap B^*$, we see that the B_{ij}^* satisfy all the conditions of the lemma. □

We have the following corollary.

Corollary 6.3.4 *Fix a good pair $xy \in A^{(2)}$, and let B^*, B_x^*, $\bar{B}_x^*$, B_y^*, $\bar{B}_y^*$, and B_{ij}^* be as described in the statement of Lemma 6.3.3. If B_x^* and B_y^* are both nonempty, then one of the following three conditions holds. Either*

(a) $B_x^* \subseteq B_y^*$ and $E_{xy} \cap B^* = B_y^* \backslash B_x^* = B_{\bar{x}y}^*$; or

(b) $B_y^* \subseteq B_x^*$ and $E_{xy} \cap B^* = B_x^* \backslash B_y^* = B_{x\bar{y}}^*$; or

(c) B_x^*,B_y^* are disjoint and $E_{xy} \cap B^* = B_{\bar{x}\bar{y}}^*$.

If $B_x^* = \phi$ or $B_y^* = \phi$, then at least one of the conditions (a), (b), or (c) holds.

Proof The proof is exactly that of Claim 3 in the proof of Theorem 6.1.1, with each set replaced by the corresponding starred set. □

As in the proof of Theorem 6.1.1, we maintain consistency as follows. Given a fixed good pair $xy \in A^{(2)}$, we declare that

(a) $B_x^* \subseteq B_y^*$ if $E_{xy} \cap B^* = B_{\bar{x}y}^*$;

(b) $B_y^* \subseteq B_x^*$ if $E_{xy} \cap B^* = B_{x\bar{y}}^*$; and

(c) B_x^*, B_y^* are disjoint if $E_{xy} \cap B^* = B_{\bar{x}\bar{y}}^*$.

As a consequence of Corollary 6.3.4 we note that for a good pair $xy \in A^{(2)}$, we have

$$\min\{|B_x \backslash B_y|, |B_y \backslash B_x|, |B_x \cap B_y|\} \leq 27r|B|.$$

For the next step in our stability result, we define an analogue of the containment colouring given in the proof of Theorem 6.1.1. We call this analogous colouring the *approximate containment colouring.* Again, it is a colouring of the elements of $A^{(2)}$. We colour $xy \in A^{(2)}$

green if xy is not a good pair;

red if xy is a good pair and $B_x^* \subseteq B_y^*$ or $B_y^* \subseteq B_x^*$;

blue if xy is a good pair and B_x^*, B_y^* are disjoint.

Our next task is to show that the graph induced by the blue edges of the approximate containment colouring is close to complete bipartite.

A triple $xyz \in A^{(3)}$ is *good* if each of the pairs xy, yz, and xz is good, and if, for $i = 1, 3$, we have

$$q_i(xyz) < \epsilon_3|B| = \epsilon_3(1-c)n,$$

where

$$\epsilon_3 = \frac{1}{2}(1 - 162r). \tag{6.13}$$

By substituting (6.12) into the above and rearranging, we have that

$$\epsilon_3 > \frac{1}{4} \text{ if } 1620\epsilon^{\frac{1}{8}} < c(1-c), \tag{6.14}$$

which holds by (6.2).

Any triple that is not good is referred to as *bad.* Recall that since

$$\sum_{xyz \in A^{(3)}} q_i(xyz) < \epsilon n^4,$$

for $i = 1, 3$, we have that $q_1(xyz) < \epsilon_3(1-c)n$ for more than $(1 - \frac{\epsilon}{\epsilon_3(1-c)})n^3$ triples of A. Similarly, $q_3(xyz) < \epsilon_3(1-c)n$ for more than $(1 - \frac{\epsilon}{\epsilon_3(1-c)})n^3$ triples of A. Also, there are fewer than $\eta_2(cn)^2$ bad pairs, and so there are fewer than $\eta_2(cn)^3$ triples that involve a bad pair. Putting all this together, we have

$$\begin{aligned}|\{xyz \in A^{(3)} : xyz \text{ bad}\}| &< \Big(\eta_2 + \frac{2\epsilon}{\epsilon_3(1-c)c^3}\Big)(cn)^3 \\ &\le \Big(\frac{7}{2}\epsilon^{\frac{1}{2}} + \frac{8\epsilon}{(1-c)c^3}\Big)(cn)^3,\end{aligned}$$

where we have used (6.11) and (6.14). Thus, we have that

$$|\{xyz \in A^{(3)} : xyz \text{ bad}\}| < \Big(\frac{7}{2} + \frac{8}{(1-c)c^3}\Big)\epsilon^{\frac{1}{2}}(cn)^3 =: \eta_3(cn)^3. \qquad (6.15)$$

We have the following lemma, which is the analogue of Claim 4 in the proof of Theorem 6.1.1.

Lemma 6.3.5 *Fix a good triple $xyz \in A^{(3)}$. Then in the approximate containment colouring of $A^{(2)}$, either xyz is a red triangle, or two of its edges are blue and the other is red.*

Proof Suppose the conclusion of the lemma fails. Then either (a) xyz is a blue triangle, or (b) xyz has two red edges and one blue edge. We derive contradictions for each case.

Define $R(xyz) = R(xy) \cup R(yz) \cup R(xz)$, where $R(xy)$, $R(yz)$, and $R(xz)$ are the sets obtained from Lemma 6.3.3. Note that $|R(xyz)| \le 81r|B| \le 162r|B|$. Let $D = B \backslash R(xyz)$ and let D_x, $\bar{D}_x$, D_y, $\bar{D}_y$, D_z, $\bar{D}_z$ be respectively the intersections of B_x, $\bar{B}_x$, B_y, $\bar{B}_y$, B_z, $\bar{B}_z$ with D. Thus, if any one of the edges, say yz, is red, then either $D_y \subseteq D_z$ or $D_z \subseteq D_y$; if yz is blue, then D_y, D_z are disjoint (and likewise for xy and xz). We have a similar notion of containment and disjointness for empty sets as we had before. Specifically, taking yz as an example, we have $D_y \subseteq D_z$ if and only if $E_{yz} \cap D = D_z \backslash D_y$; and D_y, D_z are disjoint if and only if $E_{yz} \cap D = D \backslash (D_y \cup D_z)$.

Case (a) If xyz is a blue triangle, then D_x, D_y, and D_z are pairwise disjoint. Hence, letting $\bar{D} = D \backslash (D_x \cup D_y \cup D_z)$, we have that

$$E_{xy} \cap D = D_z \sqcup \bar{D}, \quad E_{xz} \cap D = D_y \sqcup \bar{D}, \quad E_{yz} \cap D = D_x \sqcup \bar{D}.$$

Thus, for $r \in D_x \sqcup D_y \sqcup D_z$, $xyzr \in Q_1(xyz)$, and for $r \in D \backslash (D_x \cup D_y \cup D_z)$, $xyzr \in Q_3(xyzr)$. Taking the larger of the two sets, we have

$$q_i(xyz) \ge \frac{1}{2}|D| = \frac{1}{2}(|B| - |R(xyz)|) = \frac{1}{2}(1 - 81r)|B| = \epsilon_3|B|,$$

for $i = 1$ or $i = 3$, which contradicts xyz being a good triple.

Case (b) We assume without loss of generality that xz and yz are red edges and that xy is a blue edge. There are four possible cases.

1. We have $D_x, D_y \subseteq D_z$, where D_x, D_y disjoint.

2. We have $D_z \subseteq D_x, D_y$, where D_x, D_y disjoint.

3. We have $D_x \subseteq D_z \subseteq D_y$, where D_x, D_y disjoint.

4. We have $D_y \subseteq D_z \subseteq D_x$, where D_x, D_y disjoint.

Cases 3 and 4 are essentially the same. They need to be checked for our definition of containment and disjointness, that is, we need to check that transitivity holds for our definition of containment.

Case (b1) Taking $r \in D \backslash D_z$, we find that $r \in E_{xy}$, but $r \notin E_{xz}, E_{yz}$. Hence for each $r \in D \backslash D_z$, $rxyz \in Q_1(xyz)$. Therefore

$$q_1(xyz) \geq |D \backslash D_z| = |B \backslash B_z| - |R(xyz)| \geq \frac{1}{2}|B| - 81r|B| \geq \epsilon_3 |B|,$$

which contradicts xyz being a good triple.

Case (b2) We must have that $D_z = \phi$, and thus we have that $E_{yz} \cap D = D_y$; $E_{xz} \cap D = D_x$; and $E_{xy} \cap D = D \backslash (D_x \cup D_y)$. Hence, for each $r \in D$, r belongs to exactly one of E_{xy}, E_{xz}, or E_{yz}, so that $xyzr \in Q_1(xyz)$ for all $r \in D$. Therefore

$$q_1(xyz) \geq |D| > \epsilon_3 |B|,$$

contradicting that xyz is a good triple.

Case (b3) The only way we can have $D_x \subseteq D_z \subseteq D_y$ with D_x, D_y disjoint is if $D_x = \phi$. Then $E_{xz} \cap D = D_z$, $E_{yz} \cap D = D_y \backslash D_z$, and $E_{xy} \cap D = D \backslash D_y$; hence $xyzr \in Q_1(xyz)$ for every $r \in D$. Thus

$$q_1(xyz) \geq |D| > \epsilon_3 |B|,$$

contradicting that xyz is a good triple. This completes the proof of the lemma. □

As before, we define G_b to be the graph on A induced by the blue edges of the approximate containment colouring. Combining Lemma 6.3.1 and Lemma 6.3.5, we have the following corollary.

Corollary 6.3.6 *By changing at most* $3\sqrt{\eta_3}(cn)^2$ *(suitable) edges of* G_b*, we obtain a complete bipartite graph.*

Proof From the previous lemma, we see that the only contribution to $t_1(G_b)$ and $t_3(G_b)$ is from bad triples, of which there are at most $\eta_3(cn)^3$. Now apply Lemma 6.3.1.

□

Let $A_1, A_2 \subseteq A$ be the parts of the complete bipartite graph given by Corollary 6.3.6. As in the proof of Theorem 6.1.1, for each $x \in A$, we define

$$C_x = \begin{cases} B_x & \text{if } x \in A_1; \\ \bar{B}_x & \text{if } x \in A_2. \end{cases}$$

We write $\mathcal{C} = \{C_x : x \in A\}$. We expect that $\mathcal{C}$ is approximately a nested family. A measure of how close $\mathcal{C}$ is to being nested is the quantity

$$m(\mathcal{C}) = \sum_{xy \in A^{(2)}} \min(|C_x \backslash C_y|, |C_y \backslash C_x|). \tag{6.16}$$

A good pair $xy \in A^{(2)}$ is called good* if either

1. $xy \in A_1^{(2)} \cup A_2^{(2)}$ and is coloured red in the approximate containment colouring; or
2. $xy \in A \backslash A_1^{(2)} \cup A_2^{(2)}$ and is coloured blue in the approximate containment colouring.

We shall show that most pairs in $A^{(2)}$ are good*, but first we give a corollary which gives us a way of bounding $m(\mathcal{C})$. It also gives us control over E_{xy} in terms of C_x and C_y when xy is a good* pair.

Corollary 6.3.7 *Fix a good* pair* $xy \in A^{(2)}$*. We have that*

(i) $\min(|C_x \backslash C_y|, |C_y \backslash C_x|) \leq |R(xy)| \leq 27r(1-c)n$*; and*

(ii) $E_{xy} \backslash R(xy) = (C_x \triangle C_y) \backslash R(xy)$,

where $R(xy)$ *is as defined in the statement of Lemma 6.3.3.*

Proof Define

$$C_x^{(xy)} = C_x \backslash R(xy) \quad \text{and} \quad C_y^{(xy)} = C_y \backslash R(xy).$$

Note that $C_x^{(xy)} = B_x^*$ if $x \in A_1$ and $C_x^{(xy)} = \bar{B}_x^*$ if $x \in A_2$, where B_x^*, $\bar{B}_x^*$ are as defined in the statement of Lemma 6.3.3.

We prove part (i). As a consequence of Corollary 6.3.4 and the definition of the containment colouring, we have the following: if $xy \in A_1^{(2)} \cup A_2^{(2)}$ is coloured red, then B_x^*, B_y^* are nested, and hence $C_x^{(xy)}$, $C_y^{(xy)}$ are nested; if $xy \in A \backslash (A_1^{(2)} \cup A_2^{(2)})$ is coloured blue, then B_x^*, B_y^* are disjoint, but since exactly one of x or y is in A_2, we still have $C_x^{(xy)}$, $C_y^{(xy)}$ nested. In both cases $C_x^{(xy)}$, and $C_y^{(xy)}$ are nested and so

$$\min(|C_x \backslash C_y|, |C_y \backslash C_x|) \leq |R(xy)|.$$

We prove part (ii). As a consequence of Corollary 6.3.4 and the definition of the containment colouring, we have the following: if $xy \in A_1^{(2)} \cup A_2^{(2)}$ is coloured red, then B_x^*, B_y^* are nested, and hence

$$E_{xy} \backslash R(xy) = B_x^* \triangle B_y^* = C_x^{(xy)} \triangle C_y^{(xy)};$$

if $xy \in A \backslash (A_1^{(2)} \cup A_2^{(2)})$ is coloured blue, then B_x^*, B_y^* are disjoint, and hence

$$E_{xy} \backslash R(xy) = B^* \backslash (B_x^* \cup B_y^*) = C_x^{(xy)} \triangle C_y^{(xy)}$$

since exactly one of x or y is in A_2. Thus the conclusion of part (ii) of the corollary is satisfied for both cases. □

The good* pairs account for almost all elements of $A^{(2)}$; indeed, the only pairs not accounted for, which we call *bad** pairs, are green elements of $A^{(2)}$, blue elements of $A_1^{(2)} \cup A_2^{(2)}$, and red elements of $A \backslash A_1^{(2)} \cup A_2^{(2)}$. There are fewer than $\eta_2(cn)^2$ green elements of $A^{(2)}$ by (6.11), and the number of blue elements of $A_1^{(2)} \cup A_2^{(2)}$ together with red elements of $A \backslash A_1^{(2)} \cup A_2^{(2)}$ is fewer than $3\sqrt{\eta_3}(cn)^2$ by Corollary 6.3.6. Thus, the number of bad* pairs is fewer than $(\eta_2 + 3\sqrt{\eta_3})(cn)^2$, and using (6.11) and (6.15), we have

$$\begin{aligned}(\eta_2 + 3\sqrt{\eta_3})(cn)^2 &\leq \left(\frac{7}{2}\epsilon^{\frac{1}{2}} + 3\sqrt{\left(\frac{7}{2} + \frac{8}{(1-c)c^3}\right)\epsilon^{\frac{1}{2}}}\right)(cn)^2 \\ &\leq \left(\frac{7}{2} + 3\sqrt{\frac{7}{2} + \frac{8}{(1-c)c^3}}\right)\epsilon^{\frac{1}{4}}(cn)^2 \\ &\leq \left(\frac{7}{2} + 3\sqrt{\frac{7}{2}} + 3\sqrt{\frac{8}{(1-c)c^3}}\right)\epsilon^{\frac{1}{4}}(cn)^2 \\ &\leq 10\Big(1 + \frac{1}{(1-c)^{\frac{1}{2}}c^{\frac{3}{2}}}\Big)\epsilon^{\frac{1}{4}}(cn)^2.\end{aligned}$$

Therefore

$$\{xy \in A^{(2)} : xy \text{ bad*}\} \leq 10\Big(1 + \frac{1}{(1-c)^{\frac{1}{2}}c^{\frac{3}{2}}}\Big)\epsilon^{\frac{1}{4}}(cn)^2 =: \eta_4(cn)^2. \qquad (6.17)$$

Recalling (6.16), and using Corollary 6.3.7 and (6.17), we have that

$$\begin{aligned}m(\mathcal{C}) &\leq |\{\text{good* pairs}\}||R(xy)| + |\{\text{bad* pairs}\}||B| \\ &\leq (cn)^2 27r|B| + \eta_4(cn)^2|B| \\ &= (27r + \eta_4)|A|^2|B| =: s|A|^2|B|.\end{aligned}$$

Using (6.12) and (6.17), we have

$$\begin{aligned}s = (27r + \eta_4) &\leq 27\frac{5\epsilon^{\frac{1}{8}}}{c(1-c)} + 10\Big(1 + \frac{1}{(1-c)^{\frac{1}{2}}c^{\frac{3}{2}}}\Big)\epsilon^{\frac{1}{4}} \\ &\leq \Big(\frac{135}{c(1-c)} + 10 + \frac{10}{(1-c)^{\frac{1}{2}}c^{\frac{3}{2}}}\Big)\epsilon^{\frac{1}{8}} \\ &\leq \frac{155}{c^{\frac{3}{2}}(1-c)}\epsilon^{\frac{1}{8}}. \qquad (6.18)\end{aligned}$$

Our next lemma says that if a family of sets is close to being nested, in the above sense, then we can make a small number of changes to the sets and obtain a nested family.

Lemma 6.3.8 *Suppose we have a family $\mathcal{S}$ of sets, $S_1, \ldots, S_k \subseteq [n]$, where*

$$m(\mathcal{S}) := \sum_{1 \leq i < j \leq k} \min(|S_i \backslash S_j|, |S_j \backslash S_i|) \leq \gamma k^2 n.$$

By changing (adding and deleting) a total of at most $2\gamma^{1/2}kn$ elements in sets of $\mathcal{S}$, we obtain a nested family.

Proof Without loss of generality, assume that $|S_1| \geq \ldots \geq |S_k|$. Then, for $1 \leq i < j \leq k$, we have that

$$\min(|S_i \backslash S_j|, |S_j \backslash S_i|) = |S_j \backslash S_i|,$$

and thus

$$m(\mathcal{S}) = \sum_{1 \leq i < j \leq k} |S_j \backslash S_i|.$$

For $x \in [n]$, define $S_x^* = \{p \in [k] : x \in S_p\}$. Note that if each S_x^* is an initial segment of $[k]$, then $S_1, \ldots, S_k$ form a nested family. Furthermore, $x \in [n]$ contributes to $m(\mathcal{S})$ whenever $x \notin S_i$, but $x \in S_j$ for $i < j$. More precisely, defining

$$M_x = \{(i,j) \in [k]^2 : i < j,\ x \notin S_i,\ x \in S_j\},$$

and $m_x = |M_x|$, we have that

$$m(\mathcal{S}) = \sum_{x \in [n]} m_x.$$

For $x \in [n]$ and $q \in [k]$, define

$$N_{x,q}^+ = ([k] \backslash S_x^*) \cap [q] \quad \text{and} \quad N_{x,q}^- = S_x^* \cap ([k] \backslash [q]).$$

Observe that $N_{x,q}^+ \times N_{x,q}^- \subseteq M_x$. Observe further that if we increase the value of q by one, then either $|N_{x,q}^+|$ increases by one or $|N_{x,q}^-|$ decreases by one, but not both. Thus for x fixed, we can choose $q(x)$ such that $|N_{x,q(x)}^+| = |N_{x,q(x)}^-|$, so that both sets have size at most $\sqrt{m_x}$. For each $x \in [n]$, we modify $S_1, \ldots, S_k$ by adding x to every S_p for which $p \in N_{x,q(x)}^+$ and deleting x from S_p for every $p \in N_{x,q(x)}^-$. By doing this, we find that for the modified family of sets, we have $S_x^* = [q(x)]$ for each $x \in [n]$, so that each S_x^* is an initial segment of $[k]$. Thus, after modification, $S_1, \ldots, S_k$ is a nested family. The number of changes made is bounded by

$$\sum_{x \in [n]} |N_{x,q(x)}^+| + |N_{x,q(x)}^-| \leq \sum_{x \in [n]} 2\sqrt{m_x} \leq 2n\sqrt{\Big(\sum_{x \in [n]} m_x\Big)/n} = 2\sqrt{\gamma}kn,$$

where we have made use of the Cauchy-Schwarz inequality. $\square$

Let us apply Lemma 6.3.8 to the sets $(C_x)_{x\in A}$. By making a total of at most $2\sqrt{s}|A||B|$ changes to the sets $(C_x)_{x\in A}$, we obtain a family $(C_x^*)_{x\in A}$ that is nested. For each $x \in A$, let $\bar{C_x^*} = B\backslash C_x^*$. Setting $\delta_3 = \epsilon^{\frac{1}{32}}$, we have for at least $(1-\delta_3)cn$ elements of A that

$$|C_x^* \triangle C_x| \leq \frac{2\sqrt{s}}{\delta_3}(1-c)n \leq 2\sqrt{\frac{155}{c^{\frac{3}{2}}(1-c)}}\epsilon^{\frac{1}{32}}(1-c)n \\ \leq \frac{26}{c}\epsilon^{\frac{1}{32}}n, \quad (6.19)$$

where we have used (6.18). If $x \in A$ is good and the above inequality holds for x, then x is referred to as a *good** element of A. Any element of A that is not good* is referred to as *bad**. Using (6.4), the number of bad* elements is at most

$$(\eta_1 + \delta_3)cn = (2\epsilon^{\frac{1}{2}} + \epsilon^{\frac{1}{32}})cn \leq 3\epsilon^{\frac{1}{32}}cn := \eta_5 cn. \quad (6.20)$$

We are now ready to complete the proof of Theorem 6.1.2.

Proof (of Theorem 6.1.2) Let us start by describing the graph H' given in the statement of Theorem 6.1.2. As with H, A and B are independent sets for H'. For $x \in A$, define

$$E'_x = \{rs \in B^{(2)} : xrs \in E_{H'}\},$$

and for $xy \in A^{(2)}$, define

$$E'_{xy} = \{r \in B : xyr \in E_{H'}\}.$$

Note that by giving E'_x for all $x \in A$ and E'_{xy} for all $xy \in A^{(2)}$, we specify H' completely; this is what we now do. For each $x \in A$, let

$$E'_x = E(C_x^*, \bar{C_x^*}) := \{rs : r \in C_x^*, s \in \bar{C_x^*}\},$$

that is, E'_x consists of the edges of the complete bipartite graph between C_x^* and $\bar{C_x^*}$. For each $xy \in A^{(2)}$, let

$$E'_{xy} = C_x^* \triangle C_y^*.$$

Since $(C_x^*)_{x\in A}$ is a nested family, the properties above imply that $H' \in \mathcal{H}$ (see the end of the proof of Theorem 6.1.1(a)).

Finally, we estimate $E_H \triangle E_{H'}$. Note that

$$|E_H \triangle E_{H'}| = \sum_{x\in A} |E_x \triangle E'_x| + \sum_{xy\in A^{(2)}} |E_{xy} \triangle E'_{xy}|.$$

For the first sum, we note that if $x \in A$ is good*, then by (6.19), we have $|C_x^* \triangle C_x| \leq \frac{26}{c}\epsilon^{\frac{1}{32}} n$, and therefore,

$$|E(C_x^*, \bar{C_x^*}) \triangle E(C_x, \bar{C_x})| \leq \frac{26\epsilon^{\frac{1}{32}}}{c} n(1-c)n \leq \frac{26\epsilon^{\frac{1}{32}}}{c} n^2.$$

We have $E(C_x, \bar{C_x}) = E(B_x, \bar{B_x})$. By Corollary 6.3.2 and (6.3), we have

$$|E(B_x, \bar{B_x}) \triangle E_x| \leq 3\sqrt{\epsilon_1}(1-c)^2 n^2 \leq \frac{3\epsilon^{\frac{1}{4}}}{c^{\frac{1}{2}}(1-c)^{\frac{3}{2}}}(1-c)^2 n^2 \leq \frac{3\epsilon^{\frac{1}{32}}}{c} n^2.$$

Putting all this together, we have, for all good* $x \in A$, that

$$\begin{aligned}
|E'_x \triangle E_x| &= |E(C_x^*, \bar{C_x^*}) \triangle E_x| \\
&\leq |E(C_x^*, \bar{C_x^*}) \triangle E(C_x, \bar{C_x})| + |E(C_x, \bar{C_x}) \triangle E_x| \\
&\leq |E(C_x^*, \bar{C_x^*}) \triangle E(C_x, \bar{C_x})| + |E(B_x, \bar{B_x}) \triangle E_x| \\
&\leq \Big(\frac{26\epsilon^{\frac{1}{32}}}{c} + \frac{3\epsilon^{\frac{1}{32}}}{c}\Big) n^2 \leq \frac{29}{c}\epsilon^{\frac{1}{32}} n^2.
\end{aligned}$$

For each bad* $x \in A$, we have that $|E_x \triangle E'_x| \leq (1-c)^2 n^2 \leq n^2$. Using this and the above, we have

$$\begin{aligned}
\sum_{x\in A} |E_x \triangle E'_x| &\leq |\{\text{good* x}\}| \frac{29}{c}\epsilon^{\frac{1}{32}} n^2 + |\{\text{bad* x}\}| n^2 \\
&\leq (cn)\frac{29}{c}\epsilon^{\frac{1}{32}} n^2 + (3\epsilon^{\frac{1}{32}} cn) n^2 \qquad \text{(by (6.20))} \\
&\leq 32\epsilon^{\frac{1}{32}} n^3.
\end{aligned}$$

We say that a pair $xy \in A^{(2)}$ is *good*** if xy is a good* pair and both x and y are good* vertices. Any element of $A^{(2)}$ that is not good** is referred to as *bad***. Using (6.17) and (6.20), the number of bad** pairs is at most

$$10\Big(1+\frac{1}{(1-c)^{\frac{1}{2}}c^{\frac{3}{2}}}\Big)\epsilon^{\frac{1}{4}}(cn)^2+3\epsilon^{\frac{1}{32}}cn(cn)\le 20\epsilon^{\frac{1}{4}}n^2+3\epsilon^{\frac{1}{32}}c^2n^2$$
$$\le 23\epsilon^{\frac{1}{32}}n^2, \qquad (6.21)$$

where we have used that $c \le \frac{1}{2}$

If a pair $xy \in A^{(2)}$ is bad**, then we know that $|E_{xy} \triangle E'_{xy}| \le |B| \le n$. If a pair $xy \in A^{(2)}$ is good**, then we have

$$\begin{aligned}|E'_{xy} \triangle E_{xy}| &= |(C^*_x \triangle C^*_y) \triangle E_{xy}|\\ &\le |(C^*_x \triangle C^*_y) \triangle (C_x \triangle C_y)| + |(C_x \triangle C_y) \triangle E_{xy}|\\ &\le |C^*_x \triangle C_x| + |C^*_y \triangle C_y| + |(C_x \triangle C_y) \triangle E_{xy}|,\end{aligned}$$

where, for the last inequality, we have used the fact that the operation of symmetric difference is commutative and associative. By (6.19), we have that $|C^*_x \triangle C_x| \le \frac{26}{c}\epsilon^{\frac{1}{32}}n$, and similarly for $|C^*_y \triangle C_y|$. We also have that $(C_x \triangle C_y)\backslash R(xy) = E_{xy}\backslash R(xy)$ (by Corollary 6.3.7); hence $|(C_x \triangle C_y) \triangle E_{xy}| \le |R(xy)| \le 27r|B|$ by Lemma 6.3.3. Thus if $xy \in A^{(2)}$ is good**, then

$$|E'_{xy} \triangle E_{xy}| \le \frac{52}{c}\epsilon^{\frac{1}{32}}n + 27r|B| \le \frac{52}{c}\epsilon^{\frac{1}{32}}n + 27\frac{5\epsilon^{\frac{1}{8}}}{c(1-c)}(1-c)n \le \frac{187}{c}\epsilon^{\frac{1}{32}}n.$$

Using the above, we have

$$\begin{aligned}\sum_{xy\in A^{(2)}} |E_{xy} \triangle E'_{xy}| &\le |\{\text{good** xy}\}|\frac{187}{c}\epsilon^{\frac{1}{32}}n + |\{\text{bad** xy}\}|n\\ &\le (cn)^2\frac{187}{c}\epsilon^{\frac{1}{32}}n^2 + 23\epsilon^{\frac{1}{32}}(cn)^2 n \qquad \text{(by (6.21))}\\ &\le 210\epsilon^{\frac{1}{32}}n^3.\end{aligned}$$

Therefore we have that

$$\begin{aligned}|E_H \triangle E_{H'}| &\le \sum_{x\in A}|E_x \triangle E'_x| + \sum_{xy\in A^{(2)}} |E_{xy} \triangle E'_{xy}|\\ &\le 32\epsilon^{\frac{1}{32}}n^3 + 210\epsilon^{\frac{1}{32}}n^3 = 242\epsilon^{\frac{1}{32}}n^3 \le 1620\epsilon^{\frac{1}{32}}n^3.\end{aligned}$$

as required. □

6.4 A Counterexample

In this section, we show that $\pi_3(K_4^-) > \frac{5}{18}$, disproving a conjecture of Talbot [59]. Recall that the asymptotic density of equipartitioned blowups of F_6 is $\frac{5}{18}$ and that these blowups are K_4^--free and 3-colourable.

Our counterexample is constructed by adding edges to the equipartitioned blowups of F_6 and showing that the resulting hypergraph remains K_4^--free and 3-colourable. We are able to do this because there are many 3-colourings of the equipartitioned blowups of F_6; in particular, there are 3-colourings for which vertices in the same vertex class of a blowup can be coloured with two, and even three different colours.

Recall that F_6 is the the 3-graph on the vertex set $\{1, \ldots, 6\}$ with edge set $\{123, 234, 345, 451, 512, 613, 624, 635, 641, 652\}$. Let F be an equipartitioned blowup of F_6 with vertex set V and vertex classes $V_1, \ldots, V_6$. Let $|V| = n$. Consider the following partial colouring of F. All the vertices of V_1 and V_2 are coloured red; all the vertices of V_3 and V_4 are coloured green; and all the vertices of V_5 are coloured blue. Then the vertices of V_6 can coloured red, green, or blue (independently of one another) to give a 3-colouring of F. Since blowups of F_6 are 3-colourable, we can insert an equipartitioned blowup of F_6 inside V_6 to give a 3-graph F' that is 3-colourable. It is easy to check that F' is K_4^--free. The number of edges in F' is

$$f(n) + f(\lceil n/6 \rceil),$$

where $f(k)$ is the number of edges in an equipartitioned blowup of F_6 on k vertices. Since $f(k) = 10(k/6)^3 + o(k^3) = 5k^3/108 + o(k^3)$, then the number of edges in F' is

$$\frac{5n^3}{108} + \frac{5(n/6)^3}{108} + o(n^3) = \frac{1085}{23328}n^3 + o(n^3).$$

The asymptotic density of F' is therefore $\frac{1085}{3888} = \frac{5}{18} + \frac{5}{3888}$.

There are many ways in which to use the above idea to create examples of 3-graphs that are 3-colourable and K_4^--free with density greater than 5/18. This can be done by investigating colourings of F that permit more than one colour in a given colour class; by altering vertex class sizes; and by considering iterative constructions (note that in the example above, we can place a blowup of F_6 inside a vertex class in the blowup of F_6 placed in V_6).

Although none of these constructions is likely to give an extremal hypergraph for this problem, we give below the best construction we have been able to find. We note that any improvement in the lower bound on $\pi_2(K_4^-)$ would lead to an improvement in the bound below.

Theorem 6.4.1 *We have*

$$\pi_3(K_4^-) \geq 0.28153.$$

Proof Let F be a blowup of F_6 with vertex classes $V_1, \dots, V_6$ whose relative sizes are to be decided. We observe that we obtain a 3-colouring of F by colouring all vertices in V_1 red; all vertices in V_3 blue; all vertices in V_4 green; vertices in V_2 red or blue; vertices in V_5 green or blue; and vertices in V_6 red or green. Therefore, we can insert 2-colourable K_4^--free 3-graphs within the vertex classes V_2, V_5, and V_6 to obtain a 3-colourable K_4^--free 3-graph. Talbot [59] gives a family of 2-colourable K_4^--free 3-graphs which show that $\pi_2(K_4^-) = 0.25682$. Thus, for each k, there exists a 2-colourable K_4^--free 3-graph on k vertices that has

$$\frac{0.25682}{6}k^3 + o(k^3)$$

edges. Let $t = 0.25682/6$. We insert these 2-colourable 3-graphs of appropriate size into V_2, V_5, and V_6, and choose the relative sizes of $V_1, \dots, V_6$ to maximize the density of our construction. We find that the density of our construction is 0.28153... if we choose the proportion of vertices in $V_1, \dots, V_6$ to be respectively a, b, a, a, b, b, where $a = 0.16072$ and $b = 0.17261\dot{3}$. Indeed, a 3-graph on n vertices as described above has

$$(b^3 + 3ab^2 + 6a^2b)n^3 + 3t(bn)^3 + o(n^3) \geq 0.046922n^3 + o(n^3)$$

edges, giving an asymptotic density of at least 0.28153. □

Bibliography

[1] Martin Aigner and Günter M. Ziegler. *Proofs from The Book.* Springer-Verlag, Berlin, third edition, 2004. Including illustrations by Karl H. Hofmann.

[2] Noga Alon. Bipartite subgraphs. *Combinatorica*, 16(3):301–311, 1996.

[3] Noga Alon and Asaf Shapira. Every monotone graph property is testable. *SIAM J. Comput.*, 38(2):505–522, 2008.

[4] Francisco Barahona, Martin Grtschel, Michael Jnger, and Gerhard Reinelt. An application of combinatorial optimization to statistical physics and circuit layout design. *Oper. Res.*, 36(3):493–513, 1988.

[5] B. Bollobás and A. D. Scott. Better bounds for Max Cut. In *Contemporary combinatorics*, volume 10 of *Bolyai Soc. Math. Stud.*, pages 185–246. János Bolyai Math. Soc., Budapest, 2002.

[6] B. Bollobás and A. D. Scott. Judicious partitions of bounded-degree graphs. *J. Graph Theory*, 46(2):131–143, 2004.

[7] Fan Chung and Linyuan Lu. An upper bound for the Turán number $t_3(n,4)$. *J. Combin. Theory Ser. A*, 87(2):381–389, 1999.

[8] D. de Caen. Extension of a theorem of Moon and Moser on complete subgraphs. *Ars Combin.*, 16:5–10, 1983.

[9] Dominique De Caen and Zoltán Füredi. The maximum size of 3-uniform hypergraphs not containing a Fano plane. *J. Combin. Theory Ser. B*, 78(2):274–276, 2000.

[10] Jozep Diaz, Jordi Petit, and Maria Serna. A survey of graph layout problems. *ACM Computing Surveys*, 34(3):313–356, 2002.

[11] Reinhard Diestel. *Graph theory*, volume 173 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, third edition, 2005.

[12] R. P. Dilworth. A decomposition theorem for partially ordered sets. *Ann. of Math. (2)*, 51:161–166, 1950.

[13] C. S. Edwards. Some extremal properties of bipartite subgraphs. *Canad. J. Math.*, 25:475–485, 1973.

[14] C. S. Edwards. An improved lower bound for the number of edges in a largest bipartite subgraph. In *Recent advances in graph theory (Proc. Second Czechoslovak Sympos., Prague, 1974)*, pages 167–181. Academia, Prague, 1975.

[15] P. Erdős. Some recent results on extremal problems in graph theory. Results. In *Theory of Graphs (Internat. Sympos., Rome, 1966)*, pages 117–123 (English); pp. 124–130 (French). Gordon and Breach, New York, 1967.

[16] P. Erdős and M. Simonovits. A limit theorem in graph theory. *Studia Sci. Math. Hungar*, 1:51–57, 1966.

[17] P. Erdős and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.

[18] Paul Erdős and Paul Turán. On some sequences of integers. *Journal of the London Mathematical Society*, 11:261–264, 1936.

[19] Peter C. Fishburn, Paul J. Tanenbaum, and Ann N. Trenk. Linear discrepancy and bandwidth. *Order*, 18(3):237–245, 2001.

[20] L. R. Ford, Jr. and D. R. Fulkerson. Maximal flow through a network. *Canad. J. Math.*, 8:399–404, 1956.

[21] András Frank. On chain and antichain families of a partially ordered set. *J. Combin. Theory Ser. B*, 29(2):176–184, 1980.

[22] P. Frankl and Z. Füredi. An exact result for 3-graphs. *Discrete Math.*, 50(2-3):323–328, 1984.

[23] Zoltán Füredi. Turán type problems. In *Surveys in combinatorics, 1991 (Guildford, 1991)*, volume 166 of *London Math. Soc. Lecture Note Ser.*, pages 253–300. Cambridge Univ. Press, Cambridge, 1991.

[24] Zoltán Füredi, Oleg Pikhurko, and Miklós Simonovits. The Turán density of the hypergraph $\{abc, ade, bde, cde\}$. *Electron. J. Combin.*, 10:Research Paper 18, 7 pp. (electronic), 2003.

[25] Zoltán Füredi, Oleg Pikhurko, and Miklós Simonovits. On triple systems with independent neighbourhoods. *Combin. Probab. Comput.*, 14(5-6):795–813, 2005.

[26] Michael R. Garey and David S. Johnson. *Computers and intractability.* W. H. Freeman and Co., San Francisco, Calif., 1979. A guide to the theory of NP-completeness, A Series of Books in the Mathematical Sciences.

[27] W. T. Gowers. Lower bounds of tower type for Szemerédi's uniformity lemma. *Geom. Funct. Anal.*, 7(2):322–337, 1997.

[28] Curtis Greene and Daniel J. Kleitman. The structure of Sperner k-families. *J. Combinatorial Theory Ser. A*, 20(1):41–68, 1976.

[29] M. Grötschel, L. Lovász, and A. Schrijver. Corrigendum to our paper: "The ellipsoid method and its consequences in combinatorial optimization" [Combinatorica **1** (1981), no. 2, 169–197; MR0625550 (84a:90044)]. *Combinatorica*, 4(4):291–295, 1984.

[30] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2 of *Algorithms and Combinatorics.* Springer-Verlag, Berlin, second edition, 1993.

[31] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers.* Oxford Univ. Press, fifth edition, 1979.

[32] Thomas Hofmeister and Hanno Lefmann. On k-partite subgraphs. *Ars Combin.*, 50:303–308, 1998.

[33] Satoru Iwata, Lisa Fleischer, and Satoru Fujishige. A strongly polynomial-time algorithm for minimizing submodular functions. *Sūrikaisekikenkyūsho Kōkyūroku*, (1120):11–23, 1999. Algorithm engineering as a new paradigm (Japanese) (Kyoto, 1999).

[34] Gyula Katona, Tibor Nemetz, and Miklós Simonovits. On a problem of Turán in the theory of graphs. *Mat. Lapok*, 15:228–238, 1964.

[35] Peter Keevash and Benny Sudakov. On a hypergraph Turán problem of Frankl. *Combinatorica*, 25(6):673–706, 2005.

[36] Peter Keevash and Benny Sudakov. The Turán number of the Fano plane. *Combinatorica*, 25(5):561–574, 2005.

[37] Y. Kohayakawa. Szemerédi's regularity lemma for sparse graphs. In *Foundations of computational mathematics (Rio de Janeiro, 1997)*, pages 216–230. Springer, Berlin, 1997.

[38] 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.

[39] A. V. Kostochka. A class of constructions for Turán's (3, 4)-problem. *Combinatorica*, 2(2):187–192, 1982.

[40] Daniela Kühn and Deryk Osthus. Maximizing several cuts simultaneously. *Combin. Probab. Comput.*, 16(2):277–283, 2007.

[41] Linyuan Lu and Yi Zhao. An exact result and its applications on hypergraph turán numbers. *preprint (submitted)*, 2008.

[42] Dhruv Mubayi. On hypergraphs with every four points spanning at most two triples. *Electron. J. Combin.*, 10:Note 10, 4 pp. (electronic), 2003.

[43] Dhruv Mubayi. A hypergraph extension of Turán's theorem. *J. Combin. Theory Ser. B*, 96(1):122–134, 2006.

[44] Dhruv Mubayi and Vojtêch Rödl. On the Turán number of triple systems. *J. Combin. Theory Ser. A*, 100(1):136–152, 2002.

[45] Dhruv Mubayi and John Talbot. Extremal problems for t-partite and t-colorable hypergraphs. *Electron. J. Combin.*, 15(1):Research Paper 26, 9, 2008.

[46] Viresh Patel. Cutting two graphs simultaneously. *J. Graph Theory*, 57(1):19–32, 2008.

[47] Viresh Patel. Partitioning posets. *Order*, 25(2):131–152, 2008.

[48] Oleg Pikhurko. Exact computation of the hypergraph Turán function for expanded complete 2-graphs. *J. Combin. Theory Ser. B.* accepted; publication suspended for an indefinite period.

[49] D. Rautenbach and Z. Szigeti. Simultaneous large cuts. *manuscript*, 2004.

[50] Vojtěch Rödl and Mathias Schacht. Generalizations of the removal lemma. *Combinatorica (to appear).*

[51] Vojtech Rödl and Mathias Schacht. Property testing in hypergraphs and the removal lemma. In *STOC*, pages 488–495, 2007.

[52] I. Z. Ruzsa and E. Szemerédi. Triple systems with no six points carrying three triangles. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II*, volume 18 of *Colloq. Math. Soc. János Bolyai*, pages 939–945. North-Holland, Amsterdam, 1978.

[53] Alexander Schrijver. A combinatorial algorithm minimizing submodular functions in strongly polynomial time. *J. Combin. Theory Ser. B*, 80(2):346–355, 2000.

[54] Alex Scott. Judicious partitions and related problems. In *Surveys in combinatorics 2005*, volume 327 of *London Math. Soc. Lecture Note Ser.*, pages 95–117. Cambridge Univ. Press, Cambridge, 2005.

[55] M. Simonovits. A method for solving extremal problems in graph theory, stability problems. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 279–319. Academic Press, New York, 1968.

[56] Vera T. Sós. Remarks on the connection of graph theory, finite geometry and block designs. In *Colloquio Internazionale sulle Teorie Combinatorie (Roma, 1973), Tomo II*, pages 223–233. Atti dei Convegni Lincei, No. 17. Accad. Naz. Lincei, Rome, 1976.

[57] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975. Collection of articles in memory of Juriĭ Vladimirovič Linnik.

[58] Endre Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay,*