2020

Extremal Graph Theory and Dimension Theory for Partial Orders

David Lewis

EXTREMAL GRAPH THEORY

AND

DIMENSION THEORY FOR PARTIAL ORDERS

by

David C. Lewis

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Mathematical Sciences

The University of Memphis

May 2020

ACKNOWLEDGEMENTS

First I would like to thank my supervisor, Béla Bollobás, for his constant encouragement during my Ph.D. programme. I am grateful to Béla for inviting me to Cambridge, Lucca, and Rio de Janeiro, and for the countless opportunities he has given me.

I would like to thank Gabriella Bollobás for her kindness and hospitality. I am also infinitely grateful to Tricia Simmons for all her hard work, and for making sure I didn't miss any deadlines.

I am also grateful to Jeffrey Wheeler, without whom I never would have thought to apply to the University of Memphis in the first place.

I would like to thank my collaborators, António Girão, Kamil Popielarz, and Victor Souza, and the many other friends I made in Memphis, Cambridge, Lucca, and Rio de Janeiro. I would also like to thank Hunter Spink and Wang Zhiyu for their helpful comments and advice.

I would like to thank my parents, my brothers, and especially Zimmy.

Finally, I am thankful to Đặng Thị Siêm for her love and support. Em là thức ăn của anh!

The research in this dissertation was supported by NSF-DMS grants #1855745, "Applications of Probabilistic Combinatorial Methods" and #1600742, "Probabilistic and Extremal Combinatorics."

ABSTRACT

Lewis, David C. Ph.D. The University of Memphis. May 2020. Extremal Graph Theory and Dimension Theory for Partial Orders. Major Professor: Béla Bollobás, Ph.D.

This dissertation analyses several problems in extremal combinatorics.

In Part I, we study the following problem proposed by Barrus, Ferrara, Vandenbussche, and Wenger. Given a graph H and an integer t, what is the minimum number of coloured edges in a t-edge-coloured graph G on n vertices such that G does not contain a rainbow copy of H, but adding a new edge to G in any colour creates a rainbow copy of H? We determine the growth rates of these numbers for almost all graphs H and all $t \geq e(H)$.

In Part II, we study dimension theory for finite partial orders. In Chapter 1, we introduce and define the concepts we use in the succeeding chapters.

In Chapter 2, we determine the dimension of the divisibility order on $[n]$ up to a factor of $\Theta(\log \log n)$.

In Chapter 3, we answer a question of Kim, Martin, Masařík, Shull, Smith, Uzzell, and Wang on the local bipartite covering numbers of difference graphs.

In Chapter 4, we prove some bounds on the local dimension of any pair of layers of the Boolean lattice. In particular, we show that the local dimension of the first and middle layers is asymptotically $\frac{n}{\log_2 n}$.

In Chapter 5, we introduce a new poset parameter called local t-dimension. We also discuss the fractional variants of this and other dimension-like parameters.

All of Part I is joint work with António Girão of the University of Cambridge and Kamil Popielarz of the University of Memphis.

Chapter 2 of Part II is joint work with Victor Souza of IMPA (Instituto de Matemática Pura e Aplicada, Rio de Janeiro).

Chapter 3 of Part II is joint work with António Girão.

Contents

List of Figures

INTRODUCTION

The purpose of this dissertation is to solve problems in two areas of extremal combinatorics, namely extremal graph theory and dimension theory for finite partial orders.

In the first part of this dissertation, we study the rainbow saturation numbers of graphs. Given a graph H and an integer $t \geq e(H)$, the n^{th} t-rainbow saturation number of H is the minimum number of edges in a t-edge-coloured graph on n vertices that does not contain a rainbow copy of H, but to which the addition of a new coloured edge between any pair of vertices creates a rainbow copy of H. We completely characterise the growth rates of the t-rainbow saturation numbers for all graphs belonging to a large class of connected graphs. In particular, we prove that, for every $r \in \mathbb{N}$ and $t \geq \binom{r}{2}$, the t-rainbow saturation numbers of the complete graph on r vertices are asymptotically $\Theta_{r,t}(n \log n)$, confirming a conjecture of Barrus, Ferrara, Vandenbussche, and Wenger [3]. We also show that the only graphs (without isolated vertices) whose t-rainbow saturation numbers are quadratic are stars, answering another question of Barrus et al.

The second part of this dissertation concerns some dimension-like parameters of finite posets. These parameters are formally defined in the first chapter. The classic Dushnik-Miller dimension was introduced by Dushnik and Miller [43] in 1941. The Dushnik-Miller dimension (or simply the dimension) of a poset P is defined as the minimum cardinality of a set of linear orders whose intersection is P. If P is finite, we can also define the dimension of P as the smallest nonnegative integer n such that we can assign n $\mathbb{N}$–valued coordinates to each point in P in such a way that, for every x and y in P, $x \leq y$ if and only if each coordinate for x is less than or equal to the corresponding coordinate for y. Given an integer $t \geq 2$, the t-dimension of P is defined in the same way as the dimension, except that the coordinates are required to take values in $[t] = \{1, 2, \ldots t\}$.

In the second chapter, we look at the divisibility order on the set $[n] = \{1, 2, \dots n\}$. We show that the dimension of this poset is $(\log n)^2(\log\log n)^{-\Theta(1)}$ as $n \to \infty$, and that the same is true for the t-dimension for each integer $t \geq 2$. We also show that this result holds for other sets of integers, such as the arithmetic progression $a[n]+b$, and for the set of ideals in a number field with norm at most n. We prove a similar result for monic polynomials over a finite field with degree at most n, showing that this poset has dimension and t-dimension $n^2(\log n)^{-\Theta(1)}$. We also prove an analogue of a result of Füredi and Kahn [24] for the 2-dimension of posets of bounded degree, and use it to show that, for every $\alpha \in (0,1)$ and every integer $t \geq 2$, the t-dimension of the divisibility order on $(\alpha, n]$ is $\Theta_{\alpha,t}(\log n)$.

In the succeeding chapters, we study a variant of dimension called local dimension, introduced in 2016 by Ueckerdt [58]. Given a finite poset P, the local dimension of P is the smallest n such that we can assign each $x \in P$ a subset $S(x)$ of $\mathbb{N}$ with cardinality at most n and a sequence of $\mathbb{N}$–valued coordinates $\{x_i : i \in S(x)\}$ so that, for all x and y in P, $x \leq y$ if and only if $x_i \leq y_i$ for all $i \in S(x) \cap S(y)$.

The third chapter focuses on a problem related to local dimension proposed by Kim, Martin, Masařík, Shull, Smith, Uzzell, and Wang [32], namely, given a so-called *difference graph* H, what is the smallest d such that H has an edge covering with complete bipartite subgraphs that covers each vertex at most d times? In this chapter, we also prove that the poset induced by the first two layers of the n-dimensional hypercube has local dimension $(1+o(1))\log_2\log_2 n$, which is asymptotically equivalent to its Dushnik-Miller dimension.

In the fourth chapter, we prove some upper and lower bounds on the local dimension of the poset induced by any pair of layers in the n-dimensional hypercube, and show that under certain conditions these bounds are asymptotically correct. In particular, we show that the poset induced by the first and middle layer of the hypercube has local dimension $(1+o(1))\frac{n}{\log_2 n}$.

In the fifth chapter, we introduce a new parameter called the local t-dimension, defined in the same way as local dimension except that the coordinates are required to be $[t]$–valued rather than $\mathbb{N}$–valued. We prove some basic results about this parameter, and also discuss the fractional linear programming relaxation of this and other dimension-like parameters.

Graph theory notation follows [7]. In the sequel, we use log to denote the base–2 logarithm and ln to denote the natural logarithm.

Part I

Rainbow Saturation of Graphs

CHAPTER 1

INTRODUCTION TO RAINBOW SATURATION

All work in this part of the present dissertation is joint work with António Girão and Kamil Popielarz. Another version of this part has been accepted for publication in the *Journal of Graph Theory* [26].

In extremal graph theory, over many decades, much attention has been paid to the following two types of question. One is the classical Turán-type problem [57] which asks for the maximum number of edges a graph on n vertices can have provided it does not contain as a subgraph any member of a fixed class of graphs $\mathcal{H}$. The other question is concerned with another extreme, namely to determine the minimum number of edges in a graph G on n vertices which does not contain any member of $\mathcal{H}$ as a subgraph, but for which the addition of any edge between two non-adjacent vertices of G creates a copy of some graph $H \in \mathcal{H}$.

Given a class $\mathcal{H}$ of graphs, a graph G is called $\mathcal{H}$*-free* if none of its subgraphs are in $\mathcal{H}$. An $\mathcal{H}$-free graph G on n vertices that is maximal with respect to inclusion among n-vertex graphs is said to be $\mathcal{H}$*-saturated*. The latter question can then be reformulated: what is the smallest number of edges in a $\mathcal{H}$-saturated graph on n vertices? This number, called the n^{th} *saturation number* of $\mathcal{H}$, is usually denoted by $\text{sat}(n, \mathcal{H})$. When $\mathcal{H}$ is the isomorphism class of a single graph H, we write it as H. Saturation numbers were studied by Zykov [60] and independently by Erdős, Hajnal, and Moon [17] who proved that $\text{sat}(n, K_r) = (r-2)(n-1) - \binom{r-2}{2}$. By making the obvious changes, one can also define saturation numbers for ℓ-uniform hypergraphs for any $\ell \in \mathbb{N}$. Soon after saturation numbers were first introduced, Bollobás [5] showed that $\text{sat}(n, K_s^\ell) = \binom{n}{\ell} - \binom{n-(s-\ell)}{\ell}$, where K_s^ℓ is the complete ℓ-uniform hypergraph on s vertices. Note that the highest-order terms cancel and the resulting polynomial has degree $\ell - 1$ as a function of n. Bollobás also conjectured that $\mathrm{sat}(n, \mathcal{H}) = O(n)$ for every class of graphs $\mathcal{H}$. Kászonyi and Tuza [30], in 1986, confirmed this conjecture. For more information on saturation numbers we refer the reader to the survey of Faudree, Faudree, and Schmitt [20].

In Part I of this dissertation, we will be interested in a variation of the saturation numbers, following the approach of Hanson and Toft [28], who extended this notion to edge-coloured graphs.

First, we introduce some definitions. We define a t-edge-coloured graph to be an ordered pair (G, c), where G is a graph and c is a t-edge-colouring of G, i.e., a function from the edge set of G to the set $\{1, 2, 3, \ldots, t\}$, whose elements we call colours. An edge-coloured subgraph of G is a pair $(H, c|_{E(H)})$, where H is any subgraph of G. Throughout this part, we will usually identify the coloured graph (G, c) with the graph G, especially when it is clear from the context which colouring is being used. Note that we do not require edge colourings to be proper. Given an integer t and a family $\mathcal{F}$ of t-edge-coloured graphs, we say that a t-edge-coloured graph (G, c) is $(\mathcal{F}, t)$-saturated if (G, c) contains no member of $\mathcal{F}$ as an edge-coloured subgraph, but the addition of any non-edge in any colour from the set $\{1, 2, \ldots, t\}$ creates a copy of a coloured graph in $\mathcal{F}$. Similarly to the usual saturation problem, one denotes by $\mathrm{sat}_t(n, \mathcal{F})$ the minimum number of edges in a $(\mathcal{F}, t)$-saturated t-edge-coloured graph on n vertices. In [28], Hanson and Toft proved that, for any sequence of positive integers $2 \leq k_1 \leq k_2 \leq \ldots \leq k_m$,

$$\mathrm{sat}_t(n, \mathcal{M}(K_{k_1}, K_{k_2} \ldots, K_{k_m})) = \begin{cases} \binom{n}{2} & \text{if } n \leq k - 2m \\ \binom{k-2m}{2} + (k-2m)(n-k+2m) & \text{if } n > k - 2m, \end{cases} \tag{1.1}$$

where $k = \sum_{i=1}^{t} k_i$ and $\mathcal{M}(K_{k_1}, K_{k_2} \ldots, K_{k_m})$ is the collection of coloured graphs consisting of a monochromatic copy of K_{k_i} in colour i, for each $i \in \{1, 2, \ldots, m\}$.

In this part of the present dissertation, we investigate some problems proposed by Barrus, Ferrara, Vandenbussche, and Wenger [3]. Given a graph H and $t \geq e(H)$, we let $\mathfrak{R}(H)$ to be the collection of all rainbow copies of H, i.e. all t-edge-coloured graphs (H, c) where each edge is assigned a different colour from $\{1, 2 \ldots, t\}$. We shall call $\mathrm{sat}_t\,(n, \mathfrak{R}(H))$ the n^{th} *t-rainbow saturation number of H*, and, if the set of colours is infinite (say the set of natural numbers) we shall simply write $\mathrm{sat}\,(n, \mathfrak{R}(H))$ and call it the n^{th} *rainbow saturation number of H*. Our goal throughout is to determine the value of $\mathrm{sat}_t\,(n, \mathfrak{R}(H))$ for a fixed graph H.

The authors of [3] proved several beautiful and surprising results concerning these numbers. In particular, they showed a rather interesting phenomenon, namely that there are graphs whose t-rainbow saturation numbers grow considerably faster as a function of n then the usual saturation numbers. For example, they proved that for every integer r and $t \geq \binom{r}{2}$ there exist two positive constants c_1 and c_2 such that

$$c_1 \frac{n \log n}{\log \log n} \leq \mathrm{sat}_t\,(n, \mathfrak{R}(K_r)) \leq c_2 n \log n. \tag{1.2}$$

In the same paper, the authors determined the t-rainbow saturation numbers of stars, showing that $\mathrm{sat}_t\,(n, \mathfrak{R}(K_{1,k})) = \Theta(n^2)$ for any positive integers $t \geq k \geq 2$. This result confirms that the growth rates of rainbow saturation numbers behave very differently from the usual saturation numbers. They also state the following conjecture.

Conjecture 1 ([3])**.** *For any integers r and t with $t \geq \binom{r}{2}$,*

$$\mathrm{sat}_t\,(n, \mathfrak{R}(K_r)) = \Theta(n \log n).$$

One of our aims is to prove this lovely conjecture. Moreover, we show that any graph H without isolated vertices satisfying $\mathrm{sat}_t\,(n, \mathfrak{R}(H)) = \Theta(n^2)$, for some

$t \geq e(H)$, must be a star. This answers a question posed in [3] asking if stars were the only graphs with quadratic t-rainbow saturation numbers. Observe that the function $\operatorname{sat}_t(n, \mathfrak{R}(H))$ is monotonically decreasing in t for every graph H. Therefore, one just needs to show $\operatorname{sat}_t(n, \mathfrak{R}(H)) = o(n^2)$ when $t = e(H)$. Indeed, we show the following stronger result.

Theorem 1. *Let H be a graph without isolated vertices which is not a star. Then, for any $t \geq e(H)$,*

$$\operatorname{sat}_t(n, \mathfrak{R}(H)) = O(n \log n).$$

Observe trivially that adding isolated vertices to H does not change the rainbow saturation numbers when n is sufficiently large.

Given a graph H, we say that a vertex $x \in V(H)$ is *conical* if its degree is $|H| - 1$ and we say an edge is *pendant* if one of its endpoints has degree 1. For any $r \geq 4$, we define *K_r with a rotated edge* to be the graph obtained by taking with a copy of K_r, adding a new vertex, and "rotating" one edge by replacing one of its endpoints with the new vertex, as in Figure 1.1.

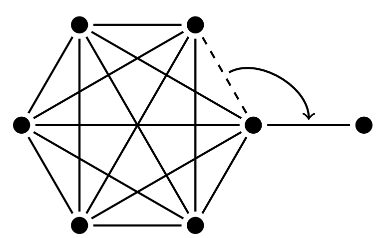

Figure 1.1: K_6 with a rotated edge. The dashed line represents the removed edge.

In the next result, we completely characterise the growth rates of t-rainbow saturation numbers of every connected graph H with no leaves, for every $t \geq e(H)$. Actually, we prove a slightly stronger result.

Theorem 2. *Let H be a connected graph of order at least 3. Then, for every $t \geq e(H)$, $\operatorname{sat}_t(n, \mathfrak{R}(H))$ equals:*

1. *$\Theta(n^2)$, if H is a star.*

2. $\Theta(n \log n)$, *if H has a conical vertex but is not a star.*

3. $\Theta(n \log n)$, *if every edge of H is in a triangle.*

4. $\Theta(n)$, *if H contains a non-pendant edge which does not belong to a triangle.*

5. $\Theta(n)$, *if H is a K_r with a rotated edge, for some even $r \geq 4$.*

We note that if H is connected with no pendant edges, then, for any $t \geq e(H)$, $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) = \Theta(n \log n)$ if every edge belongs to a triangle (by 3) and $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) = \Theta(n)$ otherwise (by 4).

We confirm Conjecture 1 as a direct consequence of Theorem 2:

Theorem 3. *For any integers r and t with $t \geq \binom{r}{2}$, $\mathrm{sat}_t\left(n, \mathfrak{R}(K_r)\right) = \Theta_{r,t}(n \log n)$.*

We would like to note that Conjecture 1 was independently proved by Korándi [37] and by Ferrara, Johnston, Loeb, Pfender, Schulte, Smith, Sullivan, Tait, and Tompkins [22].

It is easy to check that all graphs excluded from the classification of Theorem 2 can be constructed by starting with a connected graph in which every edge lies in a triangle and adding pendant edges to the graph. Note that not all graphs constructed in this way are excluded, as the class of such graphs includes all cliques with a rotated edge and some graphs with a conical vertex. For simplicity, we denote by $\mathcal{B}$ the class of all connected graphs excluded from the classification of Theorem 2.

Although we have not determined the correct order of magnitude of the t-rainbow saturation numbers of any graph H in $\mathcal{B}$ for all $t \geq e(H)$, in almost all cases, we were able to determine the order of magnitude of $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right)$ for all sufficiently large values of t. The authors of [3] also showed that if H is a graph on at least five vertices with a leaf whose neighbour is not a conical vertex and the rest of the vertices do not induce a clique then for any $t \geq \binom{|H|-1}{2}$ we have $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) = \Theta(n)$. Our next result covers almost all the remaining graphs containing a pendant edge. We show that for every H in $\mathcal{B}$ (with the exception of K_r with a rotated edge, r odd), the t-rainbow saturation number of H is linear in n, for all t sufficiently large.

Theorem 4. *Let H be a connected graph with no conical vertex and containing at least one pendant edge. Moreover, suppose H is not a copy of K_r with a rotated edge for odd $r \geq 5$. Then, for every $t \geq |H|^2$,*

$$\operatorname{sat}_t(n, \mathfrak{R}(H)) = \Theta(n).$$

In all results discussed above, we assumed that the number of available colours, t, is fixed and does not grow with n. In Theorem 20 we scratch the surface of the case when $t = t(n)$ grows with n and prove that for any $r \geq 3$ there exists a constant $c_r > 0$ such that, for any $t \geq \binom{r}{2}$, we have

$$\operatorname{sat}_t(n, \mathfrak{R}(K_r)) \leq \max\left\{\frac{c_r}{\log t} n \log n, 2(r-2)n\right\}.$$

In particular, this shows (by taking $t(n)$ to be at least linear in n) that $\operatorname{sat}(n, \mathfrak{R}(K_r)) = \Theta(n)$, for any $r \geq 3$.

Finally, we shall remark that we did not rule out the existence of a 'sharp threshold' for some connected graph H, i.e., a $t \geq e(H)$ such that $\operatorname{sat}_{t+1}(n, \mathfrak{R}(H)) = o(\operatorname{sat}_t(n, \mathfrak{R}(H)))$ as a $n \to \infty$. However, if such graph exists it must belong to $\mathcal{B}$, by Theorem 2. Note also that the set of connected graphs for which we have *not* determined the correct growth rate of their t-rainbow saturation numbers for large enough t consists exactly of the aforementioned K_r's with a rotated edge for odd $r \geq 5$.

1.1 Organisation and notation

In Chapter 2, we prove lower bounds for the t-rainbow saturation number of two classes of graphs, namely graphs where every edge belongs to a triangle and graphs which contain a conical vertex, allowing us to establish the correctness of Conjecture 1. In Chapter 3, we shall prove Theorem 1 when restricted to the class of connected graphs, as well as the main parts of the proof of Theorem 2 and Theorem 4. We split the argument in the following way. First, in Secection 3.1, we show item 4 of Theorem 2 and in Section 3.2, we prove Theorem 1 assuming the graph is connected. Secondly, in Section 3.3, we establish item 5. In Section 3.4, we shall give upper bounds (depending on t), for the t-rainbow saturation numbers of complete graphs. We also show that, when the palette of colours is infinite, the rainbow saturation numbers of complete graphs are linear. In Chapter 4, we complete the proof of Theorem 1, showing it also holds for disconnected graphs without isolated vertices. In Section 3.5, we deduce from the results proved in previous sections Theorem 2 and Theorem 4. Finally, in Section 4.1 we make some remarks and propose some conjectures and questions that we would like to be investigated.

Our notation is mostly standard. For a graph G we define $e(G)$ to be the number of edges in G. For $S \subseteq V(G)$, we denote by $e(S)$ the number of edges with both endpoints in S, and, for $S, T \subseteq V(G)$, we denote by $e(S, T)$ the number of edges with one endpoint in S and the other in T. A *non-edge* of G is an edge of $\overline{G}$. Moreover, we say a non-edge in a graph G is $\mathfrak{R}(H)$-saturated if adding e in any colour from the palette of colours understood by the context creates a rainbow copy of H. Also, if v is a vertex in an edge-coloured graph, we say informally that v *sees* a given colour if it is incident with an edge of that colour. For any positive integer k, we define the *k-star* to be the graph $K_{1,k}$.

CHAPTER 2

LOWER BOUNDS ON RAINBOW SATURATION NUMBERS

In this chapter, we show that if a graph possesses certain properties then its t-rainbow saturation numbers grow at least as fast as $\Omega(n \log n)$. Before doing so, we will need the following trivial lower bound for the rainbow saturation numbers of a connected graph on at least three vertices.

Lemma 5. *If H is a connected graph on at least three vertices then* $\mathrm{sat}\,(n, \mathfrak{R}(H)) \geq \frac{n-1}{2}$.

Proof. It is easy to check that if G is an $\mathfrak{R}(H)$-saturated graph then it has at most one isolated vertex, hence $e(G) \geq \frac{n-1}{2}$. Indeed, observe first that, since H is connected and has at least three vertices, every edge in H has an endpoint with degree at least 2. Therefore, if there are two isolated vertices in G, say x and y, then adding the edge xy to G with any colour must create a copy of H, hence either x or y must have degree at least 1, which gives a contradiction. □

The following theorem improves a result of Barrus et al. [3] and confirms Conjecture 1 above.

Theorem 6. *Let H be a graph in which every edge lies in a triangle, then if $t \geq e(H)$,*

$$\mathrm{sat}_t\,(n, \mathfrak{R}(H)) \geq \left(\frac{1}{4t} + o(1)\right) n \log n.$$

Proof. For each positive integer n, let $(G, c) = (G_n, c_n)$ be a $\mathfrak{R}(H)$-saturated t-edge-coloured graph on n vertices and $m = m(n)$ edges. Note that, by Lemma 5, $m \geq \frac{n-1}{2}$. Moreover, G has diameter at most 2, and, if $n \geq 3$, then we must have $d(v) \geq 2$ for all $v \in V(G)$. Indeed, suppose v is a vertex of degree 1, and suppose the unique edge incident with v is blue. Because every edge of H lies in a triangle, adding a new blue edge incident with v must create a rainbow triangle, but this triangle must include both the original blue edge and the new one.

For every colour $i \in \{1, 2, \ldots, t\}$ and every vertex v, let $d_i(v)$ be the degree of v in the subgraph spanned by the i-coloured edges and m_i be the total number of i-coloured edges. Now, pick a colour, say 1, and, for each vertex v and each pair $i < j$ of colours different from 1, consider the complete bipartite graph $B_v^{i,j}$ with parts S_v^i and S_v^j, where, for any colour k,

$$S_v^k = \{u \in V(G) : uv \text{ is a } k\text{-coloured edge in } G\}.$$

Since the addition of a new edge to G in colour 1 must create a rainbow triangle, every non-edge of G belongs to at least one of these bipartite graphs. Let

$$\left\{X_v^{i,j} \sim \text{Bernoulli}\left(\frac{1}{2}\right) : v \in V(G), i < j, i, j \neq 1\right\}$$

be a set of independent random variables and, for each $v \in V(G)$ and every pair of colours $i < j, i, j \neq 1$, set

$$T_v^{i,j} = \begin{cases} S_v^i & \text{if } X_v^{i,j} = 0 \\ S_v^j & \text{if } X_v^{i,j} = 1. \end{cases}$$

Now let $U = V(G) \setminus \bigcup \{T_v^{i,j} : v \in V(G), i, j \in [t], 1 \notin \{i, j\}\}$. Notice that, if uw is a non-edge, then at least one of u and w is not in U. U is therefore a clique, so, because G has m edges,

$$|U| \leq \sqrt{2m + \frac{1}{4}} + \frac{1}{2} \leq \sqrt{2m} + 1. \tag{2.1}$$

We also have the following lower bound on the expected size of U, which follows from the fact that 2^{-x} is a convex function.

$$\mathbb{E}[|U|] = \sum_{v \in V(G)} 2^{-(t-2)(d(v)-d_1(v))} \geq n \cdot 2^{-2(t-2)\frac{(m-m_1)}{n}}. \tag{2.2}$$

Combining these inequalities, we get

$$n \cdot 2^{-2(t-2)\frac{(m-m_1)}{n}} \leq \sqrt{2m} + 1. \tag{2.3}$$

Since this holds for every colour, we can take the average over all colours to obtain

$$n \cdot 2^{-2\frac{(t-1)(t-2)}{tn}m} \leq \sqrt{2m} + 1. \tag{2.4}$$

Let γ be a constant such that $m < (\gamma + o(1))n \log n$. Then $m = n^{1+o(1)}$ and

$$\sqrt{2m} + 1 = m^{\frac{1}{2}+o(1)} \geq n \cdot 2^{-2\frac{(t-1)(t-2)}{tn}m} \geq n^{1-2\gamma\frac{(t-1)(t-2)}{t}+o(1)}, \tag{2.5}$$

which implies that

$$n^{1-2\gamma\frac{(t-1)(t-2)}{t}+o(1)} \leq m^{\frac{1}{2}+o(1)} = n^{\frac{1}{2}+o(1)}. \tag{2.6}$$

It follows that $1 - 2\gamma\frac{(t-1)(t-2)}{t} + o(1) \leq \frac{1}{2} + o(1)$. By taking the limit as $n \to \infty$, we can eliminate the $o(1)$ terms, so

$$\gamma \geq \frac{t}{4(t-1)(t-2)} \geq \frac{1}{4t}. \tag{2.7}$$

$\square$

Using a similar argument, we can show that every graph with a conical vertex also has t-rainbow saturation numbers at least $\Omega(n \log n)$.

Theorem 7. *If H is a graph with a conical vertex and $|H| \geq 3$, then, for any $t \geq e(H)$,*

$$\operatorname{sat}_t(n, \mathfrak{R}(H)) \geq \left(\frac{1}{4t^2} + o(1)\right) n \log n.$$

Proof. Let H be a graph which is not a star containing a conical vertex v. For every positive integer n, let $(G, c) = (G_n, c_n)$ be an $\mathfrak{R}(H)$-saturated t-edge-coloured graph. As G has at most one isolated vertex, we can find a set $S \subset V(G)$ of size at least $\frac{n-1}{t}$ such that every vertex in S sees the same colour, say colour 1. Now we claim that, for every non-edge xy with $x, y \in S$, there must exist a rainbow path of length 2 between x and y using colours in $\{2, 3, \ldots, t\}$. Suppose by way of contradiction that this is not the case. When $e = xy$ is added and coloured 1, we must create a rainbow copy H' of H, which implies one of the endpoints of e (say x) must play the role of v and the other (say y) plays the role of a leaf in H. The latter must hold by the assumption that there is no rainbow path of length 2 between x and y. However, in this case, there would already exist a rainbow copy of H in G, namely $H' \setminus \{y\} \cup \{z\}$, where z is a neighbour of x with the edge xz coloured 1. We may now apply the same technique used in the proof of Theorem 6. Let m be the number of edges of G.

As before, for each vertex $x \in G$ and each pair $i < j$ of colours other than 1, we consider the complete bipartite graph $B_x^{i,j}$ with parts S_x^i and S_x^j, where, for any colour k, $S_v^k = \{u \in S : uv \text{ is a } k\text{-coloured edge in } G\}$. Since every non-edge

between vertices of S is joined by a rainbow path in colours other than 1, each of them is covered by at least one of these bipartite graphs. Let
$\left\{X_x^{i,j} \sim \text{Bernoulli}\left(\frac{1}{2}\right) : x \in V(G), i < j, i, j \neq 1\right\}$ be a set of independent random variables and, for each $x \in V(G)$ and each pair of colours $i < j, i, j \neq 1$, set

$$T_v^{i,j} = \begin{cases} S_v^i & \text{if } X_v^{i,j} = 0 \\ S_v^j & \text{if } X_v^{i,j} = 1. \end{cases}$$

Let $S \setminus \bigcup \{T_v^{i,j} | v \in V(G), i, j \in [t], i, j \neq 1\}$. If uw is a non-edge, then at least one of u and w is not in U. Hence U is a clique, so $|U| \leq m^{\frac{1}{2}+o(1)}$. We also have

$$\begin{aligned} \mathbb{E}[|U|] = \sum_{u \in S} 2^{-(t-2)(d(u)-d_1(u))} \geq \sum_{u \in S} 2^{-(t-2)(d(u)-1)} \geq \\ |S| \cdot 2^{-(t-2)\frac{2e(S)+e(S,V(G)\setminus S)}{|S|}-1} \geq \frac{n-1}{t} \cdot 2^{-2t(t-2)\frac{m}{n-1}-1}, \end{aligned} \tag{2.8}$$

where, as before, the second inequality holds by the convexity of 2^{-x}. Suppose γ is a constant for which $m < (\gamma + o(1))(n-1)\log(n-1)$. Then

$$\begin{aligned} (n-1)^{\frac{1}{2}+o(1)} = m^{\frac{1}{2}+o(1)} \geq \frac{n-1}{t} \cdot 2^{-2t(t-2)\frac{m}{n-1}-1} \geq \\ \frac{n-1}{2t} \cdot 2^{-2t(t-2)\gamma \log(n-1)} = (n-1)^{1-2t(t-2)\gamma+o(1)}, \end{aligned} \tag{2.9}$$

which implies that $\gamma \geq \frac{1}{4t(t-2)}$. Therefore

$$m \geq \left(\frac{1}{4t(t-2)} + o(1)\right)(n-1)\log(n-1) \geq \left(\frac{1}{4t^2} + o(1)\right) n \log n. \tag{2.10}$$

$\square$

CHAPTER 3

UPPER BOUNDS FOR CONNECTED GRAPHS

Throughout this chapter we will always assume that the graph H in the expression $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right)$ is connected and has at least three vertices. The aim of this section is to provide constructions of rainbow saturated graphs which are, in some cases, optimal up to multiplicative constants.

First, we show that if H has a cycle then $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) \leq O(n \log n)$, for any $t \geq e(H)$. Next, for any graph H with a non-pendant edge not contained in any triangle, we give constructions of t-coloured graphs on n vertices and with $\Theta(n)$ edges which are $\mathfrak{R}(H)$-saturated. Observe that if H is not a star then either H contains a cycle or H is a tree which has a non-pendant edge, hence by the aforementioned results $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) \leq O(n \log n)$ for any $t \geq e(H)$. This answers a question from [3] for connected graphs, showing that stars are the only connected graphs with quadratic rainbow saturation numbers. We also provide constructions of $\mathfrak{R}(K_r)$-saturated graphs on t colours, when t is a function of n.

3.1 Graphs with a non-pendant edge not in a triangle

In this section, we show that if H is a graph with a non-pendant edge not contained in a triangle then for any integers $t \geq e(H)$, $n \geq 1$ we have $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) \leq c_H n$, where c_H is a constant depending only on H.

Let H be a connected graph on $p \geq 3$ vertices and m edges and let $e = xy \in E(H)$ be an edge which is not contained in a triangle. For $n \geq |H| \cdot e(H)$, we shall construct a graph $G = G_{n,H,e}$ on n vertices together with an edge colouring $c = c_{n,H,e} : E(G) \to [m]$ such that the vast majority of the non-edges of (G, c) are $\mathfrak{R}(H)$-saturated and, if H satisfies some additional conditions, (G, c) is $\mathfrak{R}(H)$-free. Observe that our coloured graph (G, c) uses exactly $m = e(G)$ colours, therefore any rainbow copy of H in G must use all of these colours.

First, let $\{e_1, \ldots, e_m = e\}$ and $\{v_1, \ldots, v_{p-1} = x, v_p = y\}$ be enumerations of the edges and vertices of H, respectively. For every $i \in [m]$, let H_i be a copy of $H \setminus \{x, y\}$ with the vertex set $V_i = \left\{v_1^i, \ldots, v_{p-2}^i\right\}$, where v_j^i in H_i corresponds to v_j in H.

Now, define a graph $G = K \cup L$ where $G[K] = H_1 \cup \cdots \cup H_m$ is a disjoint union of H_i's and L is an independent set of size $n - |K|$. Moreover, for every $u \in L$, u is joined with $v_j^i \in K$ if and only if either xv_j or yv_j is an edge in H.

Having defined G, let us define an edge colouring c of G. Let w_1w_2 be an edge in G. Since L is independent we may assume that $w_1 = v_j^i$, for some $i \in [m]$ and $j \in [p-2]$. Consider now two cases depending on which part w_2 belongs to:

1. if $w_2 \in K$, then $w_2 = v_k^i$ for some $k \in [p-2]$, we let s be such that $e_s = v_jv_k$;

2. if $w_2 \in L$, we let s be such that $e_s = xv_j$ or $e_s = yv_j$.

It follows from the fact that e is not in a triangle that s is well-defined. We then define $c(w_1w_2) = s$ if $s \neq i$ and $c(w_1w_2) = m$ otherwise.
First, we shall show that every non-edge in L is $\mathfrak{R}(H)$-saturated.

Proposition 8. *Every non-edge in L is $\mathfrak{R}(H)$-saturated.*

Proof. Take any non-edge w_1w_2 in L and any colour $i \in [m]$. It is easy to check that adding the i-coloured edge w_1w_2 to the graph creates a rainbow copy of H with vertex set $V_i \cup \{w_1, w_2\}$. □

Now we shall describe the properties H must have if there exists a rainbow copy of H in (G, c).

Lemma 9. *Let W be a rainbow copy of H in (G, c). Then, all the following must hold.*

1. *If v_iv_j is an edge of H, for some $i,j \in [p-2]$, then there is k such that $v_i^k v_j^k$ is an edge in W.*

2. *There is exactly one $i \in [p-2]$ such that there exist distinct k, k' with $v_i^k, v_i^{k'} \in W$ (we shall say that i is* not unique in *W).*

3. *There is exactly one vertex in W, say z, such that $z \in L$.*

4. *If $v_i^k \in W$ and v_i is adjacent to x or y in H then v_i^k is adjacent to z in W.*

5. $d_W(z) = d_H(x) + d_H(y) - 1$.

6. *If $v_i^k v_j^k \in E(W)$ and $v_i^{k'} v_j^{k'} \in E(W)$ then $k = k'$.*

Proof. For every $k \in [m]$, we let $f_k \in E[W]$ be the edge of W of colour k. Observe, that for every $k \in [m-1]$, the only k-coloured edges in (G, c) are exactly those edges which are 'copies' of e_k, in other words,

(a) if $e_k = v_iv_j$, for $i, j \in [p-2]$ then $f_k = v_i^{k'} v_j^{k'}$ for some $k' \neq k$;

(b) and if $e_k = v_iv_j$, for $i \in [p-2]$, $j \in \{p-1, p\}$, then $f_k = v_i^{k'} z$, for some $z \in L$, $k' \neq k$.

Note that since H is connected and W must intersect at least two distinct H_i's, it follows that $|W \cap L| \geq 1$. Moreover, it follows from 3.1 and 3.1 that for every $i \in [p-2]$, there exists some $k' \in [m]$ such that $v_i^{k'} \in W$. Hence, Item 1 holds.

To see Items 2 and 3, observe first that if there are two different indices $i \neq i' \in [p-2]$ for which there exists two copies of v_i, v_j in W then $|W| \geq (p-2) + 2 + 1 = p+1$, which is a contradiction. Therefore, there is at most one index which is not unique.

To finish the proof of Items 2 and 3, it is enough to show that $|W \cap K| \geq p-1$. Let us consider where the edge f_m, of colour m appears in W. If $f_m \in G[K]$, then $f_m = v_i^k v_j^k$ for some i, j, k such that $v_iv_j = e_k$. Since we know by 3.1, that

$f_k = v_i^{k'} v_j^{k'}$ for some $k' \neq k$ we have that both i and j are not unique in W, which cannot happen as we have seen. Therefore, we may assume that $f_m = zv_i^k$ for some $z \in L$ and i, k. By construction v_i is adjacent to either x or y. Without loss of generality, we can assume that v_i is adjacent to x, and again by construction, $e_k = v_i x$. Since $f_k = wv_i^{k'}$, for some $w \in L$ and $k' \neq k$, we have that i is not unique in W. Hence, $|W \cap K| = p - 1$ and $|W \cap L| = 1$ and $w = z$.

Now, to prove Item 4, suppose $v_i^k \in W$. Notice that we already showed that if i is not unique in W then z is adjacent to v_i^k in W. Therefore, we may assume that i is unique in W. Since v_i is adjacent to either x or y, without loss of generality, we may assume that v_i adjacent to x, and therefore we have that $v_i x = e_\ell$ for some ℓ. Hence, as observed before, $f_\ell = wv_i^{k'}$ for some $w \in L$ and $k' \in [m]$. Since there is only one vertex in L, namely z, and i is unique in W we have that $w = z$ and $k' = k$ hence $f_\ell = zv_i^k$ is an edge in W.

Next, to show Item 5, note that since z is the only vertex in $W \cap L$, it must be incident with f_m and $d_H(x) - 1 + d_H(y) - 1$ edges of other colours. Hence, $d_W(z) = d_H(x) + d_H(y) - 1$.

Finally, if Item 6 does not hold then both i and j are not unique in W, which contradicts Item 2. $\square$

Proposition 10. *If H has an edge e which is in a cycle but not in a triangle, then there is no rainbow copy of H in $(G, c) = (G_{n,H,e}, c_{n,H,e})$.*

Proof. Suppose for contradiction that W is a rainbow copy of H in (G, c). Let g be the length of a longest cycle in H which uses e. We shall show that there is a natural correspondence between the g–cycles in W and the g–cycles in H not using the edge e, thus yielding a contradiction, since the number of g–cycles in W is then strictly smaller than the number of g–cycles in H.

Let C be a g–cycle in W. We shall find a corresponding g–cycle K_C in H. If C does not use vertices from L, i.e., $C = v_{k_1}^i \dots v_{k_g}^i v_{k_1}^i$, with $k_1, \dots, k_g \leq p - 2$, then let

$K_C = v_{k_1} \ldots v_{k_g} v_{k_1}$. Note that by construction K_C is a g–cycle in H.

Otherwise, by Item 3 in Lemma 9, C uses exactly one vertex from L, i.e., $C = uv^i_{k_1} \ldots v^i_{k_{g-1}} u$ with $u \in L$ and $k_1, \ldots, k_{g-1} \le p-2$. In that case let $K_C = wv_{k_1} \ldots v_{k_{g-1}} w$, where $w = x$ if v_{k_1} is a neighbour of x in H, or $w = y$ otherwise.

We claim that K_C is a g-cycle in H. Indeed, observe first that by construction $v_{k_1} \ldots v_{k_{g-1}}$ is a path in H. Note also that v_{k_1} and $v_{k_{g-1}}$ both have exactly one neighbour in $\{x, y\}$. Therefore, if v_{k_1} and $v_{k_{g-1}}$ are both adjacent to the same vertex $w \in \{x, y\}$ then K_C is indeed a g–cycle. We can therefore assume, without loss of generality, that k_1 is adjacent to x and k_{g-1} is adjacent to y. We note that $k_1, \ldots, k_{g_1}, x, y$ is then a $(g+1)$–cycle in H using the edge $e = xy$, which contradicts the assumption that g is the size of a longest cycle in H using the edge e.

It is easy to check now that if $K_C = K_{C'}$, for some g–cycle C' (different from C) in W, then we obtain a contradiction to Item 6 of Lemma 9. Finally, there is no g–cycle C in W such that K_C is a g–cycle in H using the edge e, thus we obtain a contradiction. □

Recall that that an edge in a graph is called a *bridge* if its removal increases the number of connected components.

Proposition 11. *If H has a non-pendant bridge then there is an edge $e \in H$ such that there is no rainbow copy of H in $(G_{n,H,e}, c_{n,H,e})$.*

Proof. If there is an edge e' in H which is in a cycle but not in a triangle then the result follows from Proposition 10, by taking $(G_{n,H,e'}, c_{n,H,e'})$. Hence, we may assume that every edge in H which is not in a triangle is a bridge. Let $e = xy$, with $d(x) \ge d(y)$, be a non-pendant bridge in H for which $d(x)$ is maximised. By the assumption e is well defined.

Suppose for contradiction that W is a rainbow copy of H in (G, c). We will show that the number of non-pendant bridges in W is strictly smaller than the number of non-pendant bridges in H, thus obtaining a contradiction.

Observe first that we cannot have a non-pendant bridge in W incident with any vertex $z \in L$ as then, by Item 5 of Lemma 9, we would have

$$d(z) \geq d(x) + d(y) - 1 \geq d(x) + 1, \tag{3.1}$$

which would contradict the maximality of $d(x)$. Therefore, if there is a non-pendant bridge in W, then it must be within K.

Let $b = v_i^k v_j^k$ be a non-pendant bridge in W, for some i, j, k. We shall show that $e_b = v_i v_j$ must be a non-pendant bridge in H. By assumption every edge in H which is not in a triangle is a bridge hence $v_i v_j$ is contained in a triangle, say in $v_i v_j v_\ell$ for some $\ell \in [p] \setminus \{i, j\}$.

Observe that if $v_i v_j x$ or $v_i v_j y$ is a triangle in H then, by Item 4 of Lemma 9, $v_i^k v_j^k z$ is a triangle in W; this contradicts the assumption that $v_i^k v_j^k$ is a bridge. Therefore we can assume that $v_i v_j v_\ell$ is a triangle with $\ell \leq p - 2$. Since $v_i^k v_j^k$ is a bridge in W it follows that the edge cannot belong to any triangle in W. Therefore either $v_i^k v_\ell^k$ or $v_j^k v_\ell^k$ is not an edge in W. Without loss of generality we can assume that $v_i^k v_\ell^k$ is not an edge in W. Hence, by Item 1 of Lemma 9, we must have that, for some $k' \neq k$, $v_i^{k'} v_\ell^{k'}$ is an edge in W. By the same item, there also must exist k'' such that $v_j^{k''} v_\ell^{k''}$ is an edge in W. But then there are two indices i and ℓ which are not unique in W, contradicting Item 2 of Lemma 9. Therefore, we have that $e_b = v_i v_j$ is indeed a bridge in H.

Note that by Item 6 of Lemma 9 we have that $e_b \neq e_{b'}$ for distinct non-pendant bridges b, b' in W. Hence we found a correspondence between the non-pendant bridges in W and the non-pendant bridges in $H \setminus \{e\}$, which gives a contradiction as then the number of non-pendant bridges in W is strictly smaller than the number of non-pendant bridges in H. $\square$

Theorem 12. *If H has a non-pendant edge not contained in a triangle then for any integers $t \geq e(H)$ and n we have*

$$\text{sat}_t(n, \mathfrak{R}(H)) \leq c_H \cdot n,$$

where $c_H = e(H) \cdot (|H| - 2)$.

Proof. When $n \leq e(H) \cdot (|H| - 2)$, the result follows easily by considering a monochromatic K_n. We may then assume that $n > e(H) \cdot (|H| - 2)$. Consider an edge in H as in the statement of Proposition 10 or 11. Then there is no rainbow copy of H in $(G = G_{n,H,e}, c_{n,H,e})$ and every non-edge in L is $\mathfrak{R}(H)$-saturated. If there are non-edges in G which are not $\mathfrak{R}(H)$-saturated for some colour i, we can simply add those edges to G and colour them with an appropriate colour, obtaining (G', c'). Note that

$$\begin{aligned} e(G') \leq |L||K| + \binom{|K|}{2} \leq (n - |K|)|K| + |K|^2 = \\ n|K| \leq n \cdot e(H) \cdot (|H| - 2). \end{aligned} \tag{3.2}$$

$\square$

3.2 Graphs with cycles

The construction presented in this section will be very similar to the one in Section 3.1. Let H be a graph on p vertices with a cycle. Observe that if H is triangle-free then there is an edge in H which is in a cycle but not in a triangle, hence by a result from previous section we have that $\text{sat}_t(n, \mathfrak{R}(H)) = O(n)$. Therefore, we can assume that H has a triangle. Let $e = xy$ be an edge of H which is contained in a triangle.

As before, for n large enough we shall construct a graph $G = G^r_{n,H,e}$ on n vertices together with an edge colouring $c = c^r_{n,H,e} : E(G) \to [t]$ such that the vast majority of the non-edges of (G, c) are $\mathfrak{R}(H)$-saturated and (G, c) is $\mathfrak{R}(H)$-free.

Let $\{e_1, \ldots, e_m = e\}$ and $\{v_1, \ldots, v_{p-1} = x, v_p = y\}$ be enumerations of the edges and vertices of H, respectively. For all $i \in [m]$ and $j \in [h]$, where $h = \lceil \log(n^2m + 1) \rceil$, let $H_{i,j}$ be a copy of $H \setminus \{x, y\}$ with the vertex set $V_{i,j} = \{v^{i,j}_1, \ldots, v^{i,j}_{p-2}\}$, where $v^{i,j}_l$ in $H_{i,j}$ corresponds to v_l in H.

Now we define a graph $G = K \cup L$, where $G[K] = \bigcup_{i,j} H_{i,j}$ is a disjoint union of $H_{i,j}$'s and L is an independent set of size $n - |K|$. Moreover, for every $u \in L$ and $H_{i,j}$ we shall toss a coin and based on the result decide how to join the vertices in $H_{i,j}$ with u. More precisely, for $u \in L$, $i \in [m]$ and $j \in [h]$, let $X_{u,i,j}$ be a random variable such that $\mathbb{P}\{X_{u,i,j} = x\} = \mathbb{P}\{X_{u,i,j} = y\} = \frac{1}{2}$, and let all the $X_{u,i,j}$'s be independent. Now join u with $v^{i,j}_k \in H_{i,j}$ if and only if $v_k X_{u,i,j} \in E(H)$.

Having defined G, let us define the edge colouring c. Let w_1w_2 be an edge in G. Since there are no edges in L, we can assume that $w_1 = v^{i,j}_k$ for some i, j, and k. Consider two cases depending on w_2.

1. If $w_2 \in K$ and $w_2 = v^i_{k'}$ for some k', then let s be such that $e_s = v_k v_{k'}$.

2. If $w_2 \in L$ then let s be such that $e_s = v_k X_{w_2,i,j}$.

Now $c(w_1w_2) = s$ if $s \neq i$ and $c(w_1w_2) = m$ otherwise.

Proposition 13. *With positive probability every non-edge in L is $\mathfrak{R}(H)$-saturated.*

Proof. Let $f = uv$ be a non-edge in L and $i \in [m]$ some colour. Notice, that if f is i-coloured and there is some j for which $X_{u,i,j} \neq X_{v,i,j}$, then we can find a rainbow copy of H in $\{u, v, H_{i,j}\}$. Call the pair (uv, i) *bad* if $X_{u,i,j} = X_{v,i,j}$ for every $j \in [h]$. The probability that (uv, i) is bad is equal to

$$\mathbb{P}\{X_{u,i,j} = X_{v,i,j} \text{ for every } j\} = 2^{-h}. \tag{3.3}$$

Since we have $\binom{|L|}{2} \leq n^2$ non-edges in L and m colours the expected number of bad pairs is

$$\mathbb{E}\left[\#\text{bad pairs}\right] \leq 2^{-h} n^2 m \leq \frac{n^2 m}{n^2 m + 1} < 1. \tag{3.4}$$

Therefore with positive probability there is no bad pair, hence with positive probability every non-edge in L is $\mathfrak{R}(H)$-saturated. □

Proposition 14. *There is no rainbow copy of H in (G, c).*

Proof. Suppose W is a rainbow copy of H in (G, c). We shall show that there is a natural correspondence between the triangles in W and the triangles in H not using the edge xy, thus obtaining a contradiction, since the number of triangles in W is then strictly smaller than the number of triangles in H.

Let T be a triangle in W. We shall find a corresponding triangle K_T in H. If T does not uses vertices from L, i.e., $T = \left\{v_{k_1}^{i,j}, v_{k_2}^{i,j}, v_{k_3}^{i,j}\right\}$, with $k_1, k_2, k_3 \leq p - 2$, then let $K_T = \{v_{k_1}, v_{k_2}, v_{k_3}\}$. Note that by construction K_T is a triangle in H.

Otherwise, since L is independent, T uses exactly one vertex from L, i.e., $T = \left\{v_{k_1}^{i,j}, v_{k_2}^{i,j}, u\right\}$ with $u \in L$ and $k_1, k_2 \leq p - 2$. In that case let $K_T = \{v_{k_1}, v_{k_2}, X_{u,i,j}\}$. Again, by construction K_T is a triangle in H.

It is easy to check now that if $K_T = K_{T'}$ for some distinct triangles T and T' in W then at least one colour appears twice in $E(T) \cup E(T')$, which is a contradiction. Finally, there is no triangle T in W such that K_T is a triangle in H using the edge xy. This proves that there is no rainbow copy of H in G. □

Using those two propositions we are ready to prove the main theorem of this section.

Theorem 15. *If H contains a cycle, then*

$$\operatorname{sat}_t\left(n, \mathfrak{R}(H)\right) \leq (1 + o_H(1)) c_H \cdot n \log n,$$

where $c_H = 2e(H)(|H| - 2)$.

Proof. By Theorem 12 from the previous section we may assume that H contains a triangle. Let e be an edge in H contained in a triangle. For n large enough, it follows from Propositions 13 and 14, that there is (G, c), with vertex partition $K \cup L$ (where $|K| = e(H) \cdot |H| \cdot h$), such that every non-edge in L is $\mathfrak{R}(H)$-saturated and there is no rainbow copy of H in (G, c). If there are any non-edges which are not $\mathfrak{R}(H)$-saturated we can just add those edges with appropriate colours to G obtaining (G', c'). Therefore (G', c') is $\mathfrak{R}(H)$-saturated and the number of edges in G' is at most $(n - |K|) \cdot |K| + |K|^2 = n \cdot |K| = (1 + o_H(1))c_H \cdot n \log n$, where $c_H = 2e(H)(|H| - 2)$. $\square$

Theorem 1 restricted to the class of connected graphs follows easily as a corollary of the previous theorem and Theorem 12.

Corollary 16. *Let H be a connected graph on at least three vertices which is not a star. Then, for every $t \geq e(H)$,*

$$\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) = O(n \log n).$$

Proof. If H contains a cycle then we are done by Theorem 15. If not, then H is a tree containing a non-pendant edge and the result follows from Theorem 12. $\square$

3.3 Graphs with leaves

In this section we are concerned with connected graphs with at least one leaf. In [3], Barrus, Ferrara, Vandenbussche, and Wenger showed that, with few exceptions, if H is a connected graph with a leaf, then for $t \geq \binom{|H|-1}{2}$, $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) = \Theta(n)$.

Theorem 17 (Barrus, Ferrara, Vandenbussche, and Wenger [3])**.** *Let H be a graph on at least five vertices with a leaf whose neighbour is not a conical vertex and such* *that the rest of the vertices do not induce a clique. Then, for any* $t \geq \binom{|H|-1}{2}$, *we have* $\mathrm{sat}_t\,(n, \mathfrak{R}(H)) = \Theta(n)$.

To prove similar bounds for the remaining connected graphs containing a leaf we shall introduce some terminology. We let $H_{k,\ell}$ be the graph obtained by taking a K_k (where $k \geq 3$) and adding two new vertices x and y, where x adjacent to some ℓ vertices of the clique and yx is a pendant edge. We shall call x the *middle* vertex and y the *leaf* vertex. Note that all such graphs are isomorphic however we choose the ℓ neighbours of x in K_k. Also, observe that the graph K_k *with a rotated edge* is just $H_{k-1,k-2}$.

The following proposition shows that for any $\ell \leq k-2$, the t-rainbow saturation number of $H_{k,\ell}$ is linear in n when the number of colours is sufficiently large.

Theorem 18. *For any* $2 \leq \ell \leq k-2$ *and* $t \geq k(k-1)$ *we have that*

$$\mathrm{sat}_t\,(n, \mathfrak{R}(H_{k,\ell})) = O(n).$$

Proof. Let $G = K \cup L$ where $G[K]$ is a disjoint union of two cliques of size k, say C_1, C_2, and L is independent set on $n-2k$ vertices. Now, fix $\ell+1$ vertices C_1 and $\ell+1$ vertices of C_2 and join each vertex in L to all of those vertices.

Let $A, B \subseteq [k(k-1)]$, with $|A| = |B| = \frac{k(k-1)}{2}$ be a disjoint union of colours and $A' \subseteq A$, $B' \subseteq B$ be any subsets of size $\ell+1$. We shall describe the colouring of the edges of G. First, colour the edges of C_1 using distinct colours from A, and colour the edges of C_2 using distinct colours from B. Now, for every vertex $v \in L$ colour the edges incident with v with distinct colours from B' if the edges are incident with C_1 and distinct colours from A' if the edges are incident with C_2. Note that in this colouring each vertex in L is incident with $2(\ell+1)$ edge of different colours.

We claim that there is no rainbow copy of $H_{k,\ell}$ in G. Suppose for contradiction that W is a rainbow copy of $H_{k,\ell}$ in G. First let us find a copy of k–clique C in W.

Up to symmetry there are two cases: either C uses all the vertices from C_1 or it uses $k-1$ vertices from C_1 and one vertex from L. In the former case the middle vertex must be in L and the leaf vertex must be in C_2. Which is a contradiction since C uses all colours of A and the edge between the middle and leaf vertices uses a colour from $A' \subset A$, therefore W is not rainbow. In the other case, when C uses a vertex from L, say z, note that $\ell = k-2$ and therefore the edges between z and the rest of the clique C use all of the colours from B'. Observe now that the middle vertex cannot be in C_2 since it has to be adjacent to at least two vertices of the clique C (we assumed that $\ell \geq 2$). Also, the middle vertex cannot be in L since it has to be adjacent to at least one vertex from $C_1 \cap C$, hence must be incident with an edge of colour from B' but all the colours of B' have already been used by the edges incident with z. Therefore, the middle vertex z must belong to $C_1 \setminus C$ and the leaf must be in L. This is impossible as z is not joined to any vertex of L, which is a contradiction.

Now we shall show that every non-edge in L is $\mathfrak{R}(H)$-saturated for any colour $i \in [t]$. By symmetry, we can assume that $i \in B$. (If $i \notin A \cup B$ the same argument holds). It is easy to check now that adding the edge xy, with $x, y \in L$, and colouring it with colour i, we create a rainbow copy of $H_{k,\ell}$ using all the vertices from C_1 and x, y, where x and y play the roles of the middle and leaf vertices, respectively. $\square$

The following theorem shows that, when $r \geq 4$ is even, the t-rainbow saturation of K_r with a rotated edge is linear.

Theorem 19. *Let $r \geq 4$ be even and H be K_r with a rotated edge. Then, for any $t \geq \binom{r}{2}$, $\mathrm{sat}_t(n, \mathfrak{R}(H)) = O(n)$.*

Proof. Assume $t = \binom{r}{2}$. We first define a graph Γ with vertex set $[r]^{\frac{r}{2}}$ and an edge between each pair of vertices that differ in exactly one component. Now we will define an edge colouring of Γ.

We identify the elements of $[t]$ with the edges of K_r (with vertex set $[r]$). It is well known that K_r has a proper edge colouring with $r-1$ colours if r is even. Fix one such colouring c. The edges of any given colour i form a matching with $\frac{r}{2}$ edges, and every vertex is incident with exactly one edge of colour i. For each $i \in [r-1]$, choose an arbitrary bijection g_i from $[\frac{r}{2}]$ to the set of edges of colour i. For each vertex x of Γ, let $S(x)$ be the sum of the components of x modulo r. We define the edge colouring of of Γ as follows: If x and y are two vertices of Γ that differ in the k^{th} component, colour the edge xy by $g_{c(e)}\left(k + g_{c(e)}^{-1}(e)\right)$, where $e = \{S(x), S(y)\}$. We claim that every clique in Γ is rainbow and that every vertex is incident with exactly one edge of each colour. For the first claim, observe that the restriction of S to a maximal clique is a bijection from the vertices of that clique to those of our K_r, and the function $e \mapsto g_{c(e)}\left(k + g_{c(e)}^{-1}(e)\right)$, where k is the component on which all the elements of the clique differ, permutes the edges of K_r. For the second claim, let f be any edge of our K_r and let $i = c(f)$ be its colour. Given a vertex x of Γ, let v be the unique vertex of K_r such that $\{v, S(x)\}$ is coloured i. Notice that x has exactly $\frac{r}{2}$ neighbours y such that $S(y) = v$, and each of these neighbours differs from x in a different component, hence each edge xy is a coloured with a different i-coloured edge of K_r, hence x sees the colour f. Therefore every vertex of Γ sees every colour. But every vertex of Γ has degree $\binom{r}{2}$, so it must be incident with exactly one edge of each colour.

To show that Γ is $\mathfrak{R}(H)$-free, we first observe that every clique in Γ is a subset of a maximal clique. Hence if there is a rainbow copy of H in Γ, the "missing" edge of this copy must have the same colour as the pendant edge, contradicting the fact that the colouring of Γ is proper.

Now, for any n, let G be a graph on n vertices consisting of the disjoint union of $\left\lfloor \frac{n}{r^{\frac{r}{2}}} \right\rfloor$ copies of Γ and a monochromatic clique on the leftover vertices. G is $\mathfrak{R}(H)$-free because each of its components is. Suppose we add to G a new edge e in any colour i. One endpoint x of this new edge must be in a copy of Γ. Since x is incident with an edge of colour i and this edge is in a rainbow copy of K_r, removing this edge and adding e creates a rainbow H. G is therefore an $\mathfrak{R}(H)$-saturated graph with at most $\frac{1}{2}\binom{r}{2}r^{\frac{r}{2}}\left\lfloor\frac{n}{r^{\frac{r}{2}}}\right\rfloor+\binom{r^{\frac{r}{2}}-1}{2}$ edges. $\square$

3.4 Complete graphs

Theorem 20. *For any $r \geq 3$ there exists a positive constant c_r (depending only on r) such that the following holds. For any n and $t = t(n) \geq \binom{r}{2}$,*

$$\operatorname{sat}_t(n, \mathfrak{R}(K_r)) \leq \max\left\{\frac{c_r}{\log t} n \log n, 2(r-2)n\right\}.$$

Proof. First, it is clear we may assume n is sufficiently large, by taking c_r large enough. Note that if $t \leq r^7$, by Theorem 15 we have

$$\operatorname{sat}_t(n, \mathfrak{R}(K_r)) \leq 2\binom{r}{2} rn \log n \leq \frac{r^3 \log r^7}{\log t} n \log n = \frac{7r^3 \log r}{\log t} n \log n, \tag{3.5}$$

for n sufficiently large, depending only on r. We may then assume that $t \geq r^7$. Let ℓ be a positive integer (to be specified later) and G be the union of 2ℓ disjoint $(r-2)$–cliques together with an independent set M of size $n - 2(r-2)\ell$, where each edge with one endpoint in M and the other in one of the cliques is present, and there are no edges between two distinct cliques. Observe that G does not contain any copies of K_r, because any such copy would need to use at least two vertices from M.

Let A, B an equipartition of the integers $\{1, 2, \ldots, t\}$ (thus, A, B partition $[t]$ and $||A| - |B|| \leq 1$). Now, we shall arbitrarily colour the edges of the first ℓ $(r-2)$–cliques with the colours from A and the edges of the remaining ℓ $(r-2)$–cliques with the colours from B, such that in each clique no colour appears twice. For each $(r-2)$–clique K, let C_K be the set colours used by the edges of K.

Moreover, for each vertex $x \in M$ and each clique K, we shall take a subset $B_{x,K} \subseteq A \setminus C_K$, if $C_K \subseteq A$, or $B_{x,K} \subseteq B \setminus C_K$ otherwise, of size $r-2$ uniformly at random (and independently for every choice of x and K) and colour each edge from x to K with a different element of $B_{x,K}$.

Our aim is to prove that with positive probability the addition of any coloured edge between two vertices in M will form a rainbow copy of K_r. To do so, let us compute the probability that some edge $e = xy$, with both endpoints in M, coloured c creates a rainbow copy of K_r. By symmetry, we can assume that $c \in B$. Let $t' = |A| - \binom{r-2}{2}$. Suppose K' is a rainbow copy of K_{r-2} such that $C_{K'} \subseteq A$.

First, we need the following easy claim.

Claim 1. *For positive integers s, u with $s \geq 2u-1$, the following inequality holds:*

$$\frac{\binom{s-u}{u}}{\binom{s}{u}} \geq 1 - \frac{u^2}{s-u+1}.$$

Proof. Note first, that since $s \geq 2u-1$ we have $\frac{u}{s-u+1} \leq 1$. Hence

$$\begin{aligned} \frac{\binom{s-u}{u}}{\binom{s}{u}} = \frac{(s-u)!(s-u)!}{s!(s-2u)!} &= \frac{(s-2u+1)\cdot(s-2u+2)\cdots(s-u)}{(s-u+1)\cdot(s-u+2)\cdots s} \\ &= \frac{s-2u+1}{s-u+1}\cdot\frac{s-2u+2}{s-u+2}\cdots\frac{s-u}{s} \\ &= \left(1-\frac{u}{s-u+1}\right)\left(1-\frac{u}{s-u+2}\right)\cdots\left(1-\frac{u}{s}\right) \\ &\geq \left(1-\frac{u}{s-u+1}\right)^u \geq 1-\frac{u^2}{s-u+1}, \end{aligned} \tag{3.6}$$

where the last inequality follows from Bernoulli's inequality: $(1-x)^p \geq 1-px$ for $p \geq 1$ and $x \in [0,1]$. $\square$

Observe that by construction $c \notin (C_{K'} \cup B_{x,K'} \cup B_{y,K'})$. Hence as long as $B_{x,K'}$ and $B_{y,K'}$ are disjoint we are done, i.e., there is a rainbow copy of K_r in $\{x,y\} \cup K'$. Let us bound the probability that $B_{x,K'}$ and $B_{y,K'}$ are disjoint. To do that, we apply

Claim 1 with $s = t'$ and $u = r - 2$:

$$\mathbb{P}\{B_{x,K'} \cap B_{y,K'} = \emptyset\} = \frac{\binom{t'-(r-2)}{r-2}}{\binom{t'}{r}} \geq 1 - \frac{(r-2)^2}{t'-r+3}. \tag{3.7}$$

Hence, since $t' \geq t/2 - 1$ and $t \geq r^7$, we have

$$\begin{aligned} \mathbb{P}\{\{x,y\} \cup K' \text{ is not rainbow } K_r\} &= 1 - \mathbb{P}\{B_{x,K'} \cap B_{y,K'} = \emptyset\} \\ &\leq \frac{(r-2)^2}{t'-r+3} \leq \frac{1}{\sqrt{t}}. \end{aligned} \tag{3.8}$$

Note, there are ℓ rainbow copies of K_{r-2} which only use colours from A, so we deduce that

$$\mathbb{P}\{\text{adding } e \text{ in colour } c \text{ does not create a rainbow } K_r\} \leq t^{-\ell/2}. \tag{3.9}$$

Therefore, the probability that a given edge with both endpoints in M is *bad*, i.e. the addition of e in some colour does not form a rainbow copy of K_r is at most

$$e(G) \cdot t^{-\ell/2}.$$

This holds because if we colour e in some colour not appearing in the edges of the graph, then we clearly form a rainbow copy of K_r. Hence, taking $\ell = \max\left\{\left\lceil \frac{10\log n}{\log(t)} \right\rceil, 1\right\}$, we get

$$\mathbb{P}\{\text{ some edge is } bad\} \leq e(G)\binom{M}{2} t^{-\ell/2} \leq n^4 2^{-5\log n} \leq \frac{1}{n} < 1. \tag{3.10}$$

We have thus proved there exists a colouring of G for which no edge with both endpoints in M is *bad*. If there are still some unsaturated non-edges in G, just add them one by one with appropriate colours to G. Let $N = V(G) \setminus M$. We are now done as

$$
\begin{aligned}
e(G) &\leq |N|(n-|N|) + \binom{|N|}{2} \leq |N|n - |N|^2 + |N|^2 \leq |N|n \\
&\leq 2\ell(r-2)n.
\end{aligned}
\tag{3.11}
$$

So if $\ell = 1$ then $e(G) \leq 2(r-2)n$ and if $\ell > 1$ then $e(G) \leq \frac{20(r-2)}{\log t} n \log n$. In order for the graph to be well-defined we must take n big enough (depending on r only) so that $2(r-2)\ell \leq n$. $\square$

Observe that as long as $t(n) \geq \Omega(n)$ we have $\mathrm{sat}_t\,(n, \mathfrak{R}(K_r)) = \Theta(n)$.

Corollary 21. *For any $r \geq 3$ we have*

$$
\mathrm{sat}\,(n, \mathfrak{R}(K_r)) \leq 2(r-2)n.
$$

Proof. When $n \leq 2(r-2)$ then the results follows trivially by considering a monochromatic K_n. We can therefore assume that $n \geq 2(r-2)$. Observe that when there is no restriction on the number of colours then in our construction we can assign each edge a different colour. In that case we can take $\ell = 1$, which corresponds to a disjoint union of an independent set A and two $(r-2)$–cliques B and C, such that all the edges between A and $B \cup C$ are present, and possibly some edges between B and C. The number of edges is then at most $2(r-2)n$. $\square$

We conjecture that this bound is best possible up to an additive constant.

The following construction gives a better upper bound for the rainbow saturation numbers of a triangle, at least when t is not too large compared to n.

Theorem 22. *For any $t \geq 3$ with $t \equiv 1$ or $3 \pmod 6$,*

$$
\mathrm{sat}_t\,(n, \mathfrak{R}(K_3)) \leq \frac{3}{\log \binom{t}{2}} n \log n + 3n.
$$

In particular, $\mathrm{sat}_3(n, \mathfrak{R}(K_3)) \leq \frac{3}{\log 3} n \log n + 3n$.

Proof. Let S be a Steiner triple system of order t, i.e., a set of three-element subsets of $[t]$ such that every pair of elements of $[t]$ is contained in exactly one element of S. We call the elements of t *points* and the elements of S *lines*. It can be shown (see, e.g., Kirkman [33]) that such a system exists if and only if $t \equiv 1$ or $3 \pmod 6$ and that any such system has exactly $\frac{t(t-1)}{6}$ lines.[1] We define a binary operation $\star : [t]^2 \to [t]$ as follows:

$$a \star b = \begin{cases} a & \text{if } a = b \\ c, \text{ where } c \text{ is the unique point such that } abc \text{ is a line} & \text{if } a \neq b. \end{cases}$$

This operation has the property that, for every fixed a in $[t]$, the map $b \mapsto a \star b$ permutes the elements of each line containing a. We also define a *flag* of S to be an ordered pair (ℓ, p) where ℓ is a line of S and p is a point on ℓ, and denote by F the set of all flags of S. The number of flags is $3|S| = \binom{t}{2}$. For each line ℓ, we choose an arbitrary ordering of the points on ℓ and, for any $i \in [3]$, we let $\ell^{(i)}$ denote the i^{th} point of ℓ.

Given n, let k be the smallest natural number such that $\binom{t}{2}^k + 3k \geq n$. Clearly, $k \leq \frac{1}{\log \binom{t}{2}} \log n + 1$. Let G be the complete bipartite graph with parts $V \subseteq F^k$ and $K = [k] \times [3]$, with $|V| = n - 3k$. We define a t-edge-colouring c of G, using the points of our Steiner system as colours, as follows: for each $f \in V$ and $(i, j) \in K$, let $c(\{f, (i, j)\}) = p \star \ell^{(j)}$, where (ℓ, p) is the i^{th} component of f. To show that adding an edge between two vertices in V creates a rainbow triangle, it suffices to show that every pair of such vertices is joined by either two disjoint rainbow paths of length

[1] An easy divisibility argument shows that, if a Steiner triple systems of order t exists, then $t \equiv 1$ or $3 \pmod 6$. Some examples of Steiner triple systems include the projective spaces over $\mathbb{F}_2$ (e.g., the Fano plane, which is the smallest nontrivial Steiner triple system) and affine spaces over $\mathbb{F}_3$. Bollobás [6] gives another construction for systems of prime order.

two using disjoint sets of colours or three such paths that each use a different pair of colours from a set of three. Suppose f and f' are k-tuples of flags that differ in the i^{th} component, say $f_i = (\ell, p)$ and $f'_i = (\ell', p')$. First, consider the possibility that $\ell = \ell'$ and $p \neq p'$. In this case, for every $j \in [3]$, $p \star \ell^{(j)} \neq p' \star \ell^{(j)}$, and neither is equal to $(p \star p') \star \ell^{(j)}$. Thus each path f–(i, j)–f' is a rainbow path of length two using a distinct pair of colours from ℓ. Next, if $\ell \neq \ell'$, then each edge $\{f, (i, j)\}$ is coloured with a different point from ℓ and each edge $f', (i, j)$ is coloured with a different point from ℓ' for $j \in [3]$. Since ℓ and ℓ' have at most one point in common, at most one path f–(i, j)–f' is monochromatic. If this is the case, then the other two such paths are rainbow with disjoint sets of colours. Otherwise, all such paths are rainbow, and at most one pair of them have a colour in common, so there is a pair that uses disjoint sets of colours.

It is possible that adding an edge between two vertices in K in some colour does not create a rainbow triangle; there are at most $\binom{|K|}{2}$ such edges. We can add these coloured edges to (G, c) to form an $\mathfrak{R}(K_3)$-saturated t-edge-coloured graph (G', c') with at most

$$|V||K| + \binom{|K|}{2} \leq (n - |K|)|K| + |K|^2 = n|K| \leq \frac{3}{\log \binom{t}{2}} n \log n + 3n \qquad (3.12)$$

edges. □

When $t = 3$, the coefficient of the $n \log n$ term in the upper bound is $\frac{3}{\log 3}$, while for large values of t it is approximately $\frac{1.5}{\log t}$.

3.5 Deducing the main results for connected graphs

We are now ready to deduce Theorems 2 and 4.

***Proof of Theorem** 2.* First, note that Item 1 is a result appearing in [3] and Item 5 is just a restatement of Theorem 19. Now, the lower bounds in Items 2 and 3

follow from Theorems 6 and 7, respectively, and the upper bound in each is a consequence of Theorem 15, since in either case H must contain a cycle.
In Item 4 the lower bound follows from Lemma 5 and the upper bound follows from Theorem 12. □

Proof of Theorem 4. Observe first that if H is a connected graph on at most four vertices which contains a leaf and no conical vertex, then H must be a path on four vertices, hence by Item 4 of Theorem 2 its t-rainbow saturation numbers are linear. We may therefore assume that $|H| \geq 5$. Let xy be a pendant edge of H. If $H \setminus \{x, y\}$ is not a clique then we are done by Theorem 17. Hence, we may then assume $H = H_{k,\ell}$ for some $k \geq 3$ and $\ell \leq k-1$. Suppose $\ell \leq k-2$, then result follows by Theorem 18. Hence, we may assume $\ell = k-1$ in which case k must be odd, by assumption, and therefore H is a K_{k+1} with a rotated edge, so we are done by Theorem 19. □

CHAPTER 4

UPPER BOUNDS FOR DISCONNECTED GRAPHS

In this section, we shall show that the rainbow saturation number of a disconnected graph can be bounded above by the rainbow saturation number of one of its connected components, up to additive $O(n)$ term. Moreover, we shall show that if H is a disconnected graph with no isolated vertices, then the t-rainbow saturation number of H is at most $O(n \log n)$ answering a question from [3] for disconnected graphs. Throughout the section, we assume, for simplicity of exposition, that H has no isolated vertices.

For a sequence of graphs $H_1, \ldots, H_k$ we say that H_i is *maximal*, for some $i \in [k]$, if H_i is not isomorphic to any proper subgraph of H_j for any $j \in [k]$. Observe that every finite sequence has a maximal element; for example, we can take one with the largest total number of vertices and edges.

Proposition 23. *Let H be a graph with connected components $H_1, \ldots, H_k$ and let H_i be a maximal component. Then, for every $t \geq e(H)$, we have*

$$\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) \leq \mathrm{sat}_t\left(n, \mathfrak{R}(H_i)\right) + O(n).$$

Proof. Without loss of generality we may assume that $i = 1$ and $H_1 \cong H_2 \cong \ldots \cong H_\ell$ (for some $\ell \in [k]$), and that no other component is isomorphic to H_1. Let $H' = H_{\ell+1} \cup H_{\ell+2} \cup \ldots \cup H_k$.

Let $t' = e(H) \leq t$ and consider the following graph G on n vertices. First add vertex-disjoint copies of all possible rainbow copies of H' for every subset of size $|e(H')|$ in $[t']$. Write V_1 for the set of vertices spanned by these copies. Second, consider the following coloured graph $H_1^\star$: for every set A of colours of size $e(H_1)$ inside $[t']$, we add a rainbow of copy of H_1 with colours in A, where all rainbow copies share exactly one vertex. Now we add $\ell - 1$ vertex disjoint copies of $H_1^\star$ to G

and define V_2 to be the set of vertices spanned by these copies. In the set $V(G) \setminus (V_1 \cup V_2)$, consisting of the remaining vertices, we add a $\mathfrak{R}(H_1)$-saturated graph on t colours. It is easy to check that every non-edge in $V(G) \setminus (V_1 \cup V_2)$ is $\mathfrak{R}(H)$-saturated. Finally, if there are any non-edges which are not $\mathfrak{R}(H)$-saturated, we add those edges to G in some colour that does not create a rainbow H. Clearly, there are at most $O(n)$ such edges.

Let us show G does not contain a rainbow copy of H. Suppose by way of contradiction that it does. We shall obtain a contradiction by showing that the number of vertex disjoint rainbow copies of H_1 in G is strictly smaller ℓ. Note that H_1 cannot be a subgraph of $G[V_1]$ as, by construction, H_1 is not isomorphic to any connected component of $G[V_1]$ and, by maximality, H_1 cannot be a subgraph of any connected component of $G[V_1]$. Observe as well that each copy of $H_1^{\star}$ contains at most one rainbow copy of H_1. Finally, by construction, $V(G) \setminus V_1 \cup V_2$ does not contain a rainbow copy of H_1. Therefore there are at most $\ell - 1$ vertex disjoint rainbow copies of H_1.

Let $p = |V_1 \cup V_2|$. Observe that $p = \Theta(1)$ as n goes to infinity. Therefore the number of edges in G is at most

$$\begin{aligned} \binom{p}{2} + p(n-p) + \mathrm{sat}_t\left(n-p, \mathfrak{R}(H_1)\right) \leq \\ pn + \mathrm{sat}_t\left(n, \mathfrak{R}(H_1)\right) = \mathrm{sat}_t\left(n, \mathfrak{R}(H_1)\right) + O(n). \end{aligned} \tag{4.1}$$

□

We have the following immediate corollary.

Corollary 24. *Let H be a graph containing at least one component which is not a star and let H' be a maximal component among the components of H which are not* *stars. Then, for every $t \geq e(H)$, we have*

$$\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) \leq \mathrm{sat}_t\left(n, \mathfrak{R}(H')\right) + O(n) \leq O(n \log n).$$

Proof. Observe that H' cannot be a subgraph of a star, hence by Proposition 23 and Corollary 16, we have that

$$\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) \leq \mathrm{sat}_t\left(n, \mathfrak{R}(H')\right) + O(n) \leq O(n \log n).$$

$\square$

We have thus shown that if a disconnected graph contains a component which is not a star then its t-rainbow saturation numbers are is subquadratic. Since stars have t-rainbow saturation numbers which are quadratic in n, one might suspect that the same should hold for disconnected graphs where each component is a star. The following proposition shows that this is not the case.

Proposition 25. *Let $H = S_1 \cup S_2 \cup \cdots \cup S_k$ be a graph with more than one component, each of which is a star. Then for every $t \geq e(H)$ we have*

$$\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) \leq O(n).$$

Proof. Suppose $|S_1| \leq |S_2| \leq \cdots \leq |S_k|$. First we shall show the case when $k = 2$. Let $a = |S_1| - 1$ and $b = |S_2| - 1$. Let $G = K \cup L$ where $G[K]$ is a complete graph of size $a + b - 1$ and L is an independent set of size $n - |K|$. Let $K = \{x_1, \ldots, x_{a+b-1}\}$. First we join every vertex $x_i \in K$ with every vertex $y \in L$ and give the edge colour i. Next we shall describe the colouring of the edges inside K. Let $x_i, x_j \in K$ where $i \leq j$. If $i \leq a$ and $j \geq a$ then assign $a + b$ as the colour of $x_i x_j$, otherwise assign j as the colour of $x_i x_j$.

We claim that there is no rainbow copy of $S_1 \cup S_2$ in G. To see that, observe first that every rainbow copy of S_i in G uses at least $|S_i|-1$ vertices of K. Indeed, suppose for contradiction that it is not the case and that there is a rainbow copy of S_i which uses fewer than $|S_i|-1$ vertices of K. Then it must use at least two vertices, say x, y, of L. It follows from independence of L that the centre z of that rainbow copy must be in K. We obtain a contradiction by noticing that zx and zy have the same colour. Therefore if there is a rainbow copy of $S_1 \cup S_2$ then it has to use at least $a+b$ vertices of K, which is a contradiction since there are only $a+b-1$ such vertices.

Next we shall show that every non-edge is $\mathfrak{R}(H)$-saturated. Consider any non-edge xy in L and any colour $c \in [t]$.

If $c \leq a$ then we find a copy of S_1 in $\{x, y, x_1, \ldots, x_a\} \setminus \{x_c\}$ with x being the centre and a copy of S_2 in $\{x_c, x_{a+1}, \ldots, x_{a+b-1}, z\}$ with x_{a+1} as the centre, for any $z \in L \setminus \{x, y\}$. Observe that those two copies are vertex disjoint and the copy of S_1 uses only colours from $[a]$ and the copy of S_2 uses colours from $[a+1, a+b]$. Hence we have a rainbow copy of $S_1 \cup S_2$.

If $c \in [a+1, a+b-1]$ then we find a copy of S_2 in $\{x, y, x_a, \ldots, x_{a+b-1}\} \setminus \{x_c\}$ with x being the centre and a copy of S_1 in $\{x_1, \cdots, x_{a-1}, x_c, z\}$ with x_1 as the centre, for any $z \in L \setminus \{x, y\}$. Observe that those two copies are vertex disjoint and the copy of S_1 uses only colours from $[a-1] \cup \{a+b\}$ and the copy of S_2 uses colours from $[a, a+b-1]$. Hence we have a rainbow copy of $S_1 \cup S_2$.

In the remaining case when $c \geq a+b$, it is easy to check that we can find a rainbow copy of $S_1 \cup S_2$ where both of the centres are in L.

Observe that we have

$$e(G) \leq (n-|K|)|K| + |K|^2 = |K|n = (a+b-1)n = (|S_1|+|S_2|-3)n. \qquad (4.2)$$

Now, suppose $k \geq 3$. We let $t^\star = e(H)$. Moreover, let $G = G' \cup G''$ where G' is an $\mathfrak{R}(S_1 \cup S_2)$-saturated graph on $n' = n - (k-2)(t^\star + 1)$ vertices with $\mathrm{sat}_{t^\star}(n', \mathfrak{R}(S_1 \cup S_2))$ edges and G'' is the vertex-disjoint union of $k-2$ rainbow copies of $t^\star$-stars. It is easy to check that there is no rainbow copy of H in G. Indeed, by assumption there can not be two vertex-disjoint rainbow copies of distinct components of H appearing in G'. Note as well that there can only be at most $k-2$ vertex-disjoint stars in G'', hence in total there are at most $k-1$ disjoint rainbow components of H in G. Finally, it is clear that the addition of any coloured non-edge inside G' creates a rainbow copy of H. Now, we keep adding edges to G (with both endpoints in G'' or with one endpoint in G' and one in G'') until G is saturated. The case $k = 2$ shows that $\mathrm{sat}_{t^\star}(n, \mathfrak{R}((S_1 \cup S_2))) \le (|S_1| + |S_2| - 3)n$, hence the number of edges in G is at most

$$n - |G''|)|G''| + |G''|^2 + e(G') \le n|G''| + (|S_1| + |S_2| - 3)n \le O(n). \tag{4.3}$$

□

We now have the following corollary from Propositions 23 and 25.

Corollary 26. *Let H be a disconnected graph. Then for every $t \ge e(H)$ we have*

$$\mathrm{sat}_t(n, \mathfrak{R}(H)) \le O(n \log n) = o(n^2).$$

□

4.1 Concluding remarks and open problems

We have shown that for any $t \ge \binom{r}{2}$, $\mathrm{sat}_t(n, \mathfrak{R}(K_r)) = \Theta(n \log n)$ when $n \to \infty$, i.e., there exist constants $c_1 = c_1(t, r)$ and $c_2 = c_2(t, r)$ such that $c_1 n \log n \le \mathrm{sat}_t(n, \mathfrak{R}(K_r)) \le c_2 n \log n$. There is still an enormous gap between our lower and upper bounds. We originally conjectured that the true value is closer to the upper bound:

Conjecture 2. *For every $r \geq 3$, there exists a constant $c(r) > 0$ such that for every $t \geq \binom{r}{2}$,*

$$\mathrm{sat}_t\left(n, \mathfrak{R}(K_r)\right) \geq \frac{c(r)}{\log t} n \log n$$

for all $n \geq n_0(t)$.

Korándi [37] resolved Conjecture 2 by proving the following theorem.

Theorem 27 (Korándi [37])**.** *For every $r \geq 3$, and any $t \geq \binom{r}{2}$,*

$$\mathrm{sat}_t\left(n, \mathfrak{R}(K_r)\right) \geq \frac{t(1+o(1))}{(t-r+2)\log(t-r+2)} n \log n$$

as $n \to \infty$, with equality for $r = 3$. □

This theorem together with Theorem 20 gives

$$\mathrm{sat}_t\left(n, \mathfrak{R}(K_r)\right) = \Theta_r\left(\frac{n \log n}{\log t}\right). \tag{4.4}$$

Now, when H is an even clique with a rotated edge, we know that $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right)$ is always $\Theta(n)$ for $t \geq e(H)$. However, for odd cliques with rotated edges, we do not even know the asymptotic behaviour of $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right)$ for large values of t.

Question 1. *If H is a copy of K_r with a rotated edge (as shown in Figure 4.1) for some* odd *$r \geq 5$ and $t \geq \binom{r}{2}$, what is the asymptotic growth rate of* $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right)$*?*

The following conjecture together with Theorem 2 and Question 1 would completely classify the possible rates of growth of $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right)$ for all connected graphs H and every constant $t \geq e(H)$.

Conjecture 3. *Let H be a connected graph (other than an odd clique with a rotated edge) with an edge not in a triangle and no conical vertex. Then, for every $t \geq e(H)$, $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) = O(n)$.*

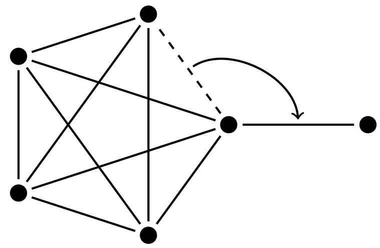

Figure 4.1: K_5 with a rotated edge. The dashed line represents the removed edge.

Note that we can confirm this conjecture when the number of available colours is at least $\binom{|H|-1}{2}$. Indeed, either H is in one of the classes defined in Theorem 2, in which case we are done, or H has a leaf and is not a clique with a rotated edge, hence by Theorem 4 we have $\mathrm{sat}_t\left(n, \mathfrak{R}(H)\right) = \Theta(n)$.

One different direction would be to allow the palette of colours to be infinite. We have only considered this question for complete graphs and showed that $\mathrm{sat}\left(n, \mathfrak{R}(K_r)\right) \leq 2(r-2)n$ for any $r \geq 3$.

Recall that the construction in Corollary 21 is a disjoint union of an independent set A and two $(r-2)$–cliques B and C, such that all the edges between A and $B \cup C$ are present and all the edges in B, C and between A and $B \cup C$ receive different colours. We conjecture that, whenever $n \geq 2(r-2)$, the above construction is best possible up to the configuration of the edges between B and C.

Conjecture 4. *For any integer $r \geq 3$, there exists a constant C_r depending only on r such that, for any $n \geq 2(r-2)$,*

$$\mathrm{sat}\left(n, \mathfrak{R}(K_r)\right) = 2(r-2)n + C_r.$$

Finally, we conjecture that, like the ordinary saturation numbers, the rainbow saturation numbers of any graph are at most linear in n.

Conjecture 5. *For any graph H,* $\mathrm{sat}\left(n, \mathfrak{R}(H)\right) = O(n)$.

Part II

Dimension Theory of Posets

CHAPTER 1

INTRODUCTION TO POSETS AND DIMENSION THEORY

In Part II of this dissertation, we study some questions about the combinatorics of finite posets.

Given a set X, a *partial order* on X is a binary relation $\leq$ on X that is *reflexive* ($x \leq x$ for all $x \in X$), *transitive* (for all x, y, and $z \in X$, if $x \leq y \leq z$, then $x \leq z$), and *antisymmetric* (for all x and $y \in X$, if $x \leq y \leq x$, then $x = y$). A *partially ordered set*, also known as a *poset*, is an ordered pair $P = (X, \leq)$, where X is a nonempty set (called the *ground set* of P) and $\leq$ is a partial order on X. We often identify a poset with its ground set, especially when the partial order is obvious in context, and write, for example, $x \in P$ to mean $x \in X$. We also write $x < y$ as an abbreviation for $x \leq y$ and $x \neq y$. We also write $x \geq y$ to mean $y \leq x$ and write $x > y$ to mean $y < x$.

Given a poset $P = (X, \leq)$, a *suborder* of P is a poset $(Y, \leq|_Y)$, where $Y \subseteq X$ and $\leq|_Y$ is the relation $\leq$ restricted to Y.

Two elements x and y of a poset are called *comparable* if either $x \leq y$ or $y \leq x$, and *incomparable* otherwise. A poset for which all elements are pairwise comparable is called a *linear order* or a *chain* (especially when regarded as a suborder of another poset) and a poset in which all elements are pairwise incomparable is called an *antichain.*

Let $P = (x, \leq)$ be a poset. The *comparability graph* of P is the graph with vertex set X and edge set $\{\{x, y\} : x \leq y\}$. Given two elements x and y of X, y is said to *cover* x if $x < y$ and, for every $z \in X$, if $x \leq z \leq y$, then $z = x$ or $z = y$. The *cover graph* of P is the graph with vertex set X and edge set $\{xy : y \text{ covers } x\}$. An *order diagram* of P is a drawing of the cover graph in the plane with the property that, whenever b covers a, the y-coordinate of b is strictly greater than that of a, and the edge ab goes monotonically upward from a to b.

Given two posets $P = (S, \leq_P)$, $Q = (T, \leq_Q)$, a poset embedding from P into Q is a map $\varphi : S \to T$ such that $\varphi(a) \leq_Q \varphi(b)$ if and only if $a \leq_P b$. The expression $P \hookrightarrow Q$ represents an embedding from P into Q or the existence of such an embedding, depending on context.

Given a family of posets $P_i = (S_i, \leq_{P_i})$, $i \in I$, the product poset $P = \prod_{i \in I} P_i$ is the unique order on the product set $S = \prod_{i \in I} S_i$ such that $a \leq_P b$ if and only if $a_i \leq_{P_i} b_i$ for all $i \in I$. Given two posets P and Q, the product of P and Q is denoted by $P \times Q$. For any $n \in \mathbb{N}$, P^n denotes the product of n copies of P. The empty product P^0 is the trivial one-element poset.

For any $n \in \mathbb{N}$, we denote by boldface $\mathbf{n}$ a chain with n elements and by A_n an antichain with n elements. The n-dimensional Boolean lattice, also known as the hypercube or simply the cube, is defined as $(\mathcal{P}[n], \subseteq)$ and denoted by $\mathcal{Q}^n$. We sometimes identify this poset with $\mathbf{2}^n$, to which it is isomorphic. For any integer ℓ, the ℓ^{th} layer of $\mathcal{Q}^n$ is defined as the set of all ℓ-element subsets of $[n]$ (i.e., $[n]^{(\ell)}$). For every subset $A \subseteq [n]$, $\mathcal{Q}^n_A$ denotes the suborder of $\mathcal{Q}^n$ consisting of the union of the a^{th} layer for each $a \in A$, i.e., $\left(\bigcup_{a \in A} [n]^{(a)}, \subseteq \right)$. In this context, we sometimes write sets as lists without brackets, e.g., $\mathcal{Q}^n_{1,2}$ instead of $\mathcal{Q}^n_{\{1,2\}}$.

The order dimension of a poset, also known as the Dushnik-Miller dimension or simply the dimension, was first defined by Dushnik and Miller in [43] and has been the subject of extensive research, e.g, by Hiraguchi in [29] and Trotter in [54]. Given a poset P, a *linear extension* of P is a linear order L on the ground set of P such that, for all x and $y \in P$, if $x \leq_P y$, then $x \leq_L y$. A *realiser* of P is a set $\mathcal{L}$ of linear extensions of P such that, for every ordered pair $(x, y) \in P$ with $x \not\geq y$, there is a linear extension $L \in \mathcal{L}$ such that $x <_L y$. The dimension of P, denoted by $\dim(P)$, is defined as the minumum cardinality of a realiser of P. For example, the dimension of a nontrivial chain is 1 and the dimension of a nontrivial antichain is 2. Note that the dimension of the trivial poset $\mathbf{1}$ is 0, not 1 as stated in some sources.

Equivalently (see [55]), $\dim(P)$ is the smallest cardinal n such that P can be embedded into the product of n chains as a suborder. It follows that dimension is monotone, i.e., if $P \hookrightarrow Q$, then $\dim(P) \leq \dim(Q)$, and subadditive, i.e., $\dim(P \times Q) \leq \dim(P) + \dim(Q)$ for all posets P and Q. Another corollary of this equivalence is that the hypercube $\mathbf{2}^n$ has dimension n.

Ore [45] defined the dimension of a poset P as the smallest n such that $P \hookrightarrow \mathbb{R}^n$. When P is countable, this definition is equivalent to Dushnik and Miller's definition; this is a corollary of the well-known fact that every countable linear order embeds into $\mathbb{Q}$.

The local dimension of a poset is a variant of the Dushnik-Miller dimension. This concept was introduced by T. Ueckerdt [58] at the *Order and Geometry* workshop and studied by Kim, Martin, Masařík, Shull, Smith, Uzzell, and Wang in [32]. A *partial linear extension* of a poset P is a linear extension of a suborder of P. A *local realiser* of P is a set $\mathcal{L}$ of partial linear extensions of P such that, for every pair $(x, y) \in P^2$ with $x \ngeq y$, there is a partial linear extension $L \in \mathcal{L}$ such that $x <_L y'$. Given a local realiser $\mathcal{L}$ of a poset P, the multiplicity of an element $x \in P$ in $\mathcal{L}$ is defined as the number of partial linear extensions in $\mathcal{L}$ of which x is an element, and is denoted by $\mu_{\mathcal{L}}(x)$. The *local dimension* of P, denoted $\operatorname{ldim}(P)$, is the minimun over all local realisers $\mathcal{L}$ of $\max\{\mu_{\mathcal{L}}(x) : x \in P\}$. It is clear that for any poset P, $\operatorname{ldim}(P) \leq \dim(P)$, however, as observed by Ueckerdt, the gap between the two can be arbitrarily large. To see this, consider the standard example S_n, which is a height 2 poset with elements $\{a_1, a_2, \ldots, a_n\}$ and $\{b_1, b_2, \ldots, b_n\}$, where $a_i < b_j$ if $i \neq j$ and there are no other nontrivial relations. Let L be a linear extension of S_n such that $b_i <_L a_i$. Then, for every $j \neq i$, $a_j <_L b_i <_L a_i <_L b_j$. Any realiser $\mathcal{L}$ of S_n must contain, for each $i \in [n]$, an extension L such that $b_i <_L a_i$, but each extension can contain at most one relation of this form, so $\dim(S_n)$ is at least (in fact exactly) n. However, we can construct a local realiser of S_n with maximum multiplicity at most 3 by taking the partial extensions $a_1, a_2, \dots a_n$, $a_n, a_{n-1}, \dots a_1$, $b_1, b_2, \dots b_n$, $b_n, b_{n-1}, \dots b_1$, and, for each $i \in [n]$, $b_i a_i$. Hence $\operatorname{ldim}(S_n) \le 3$.

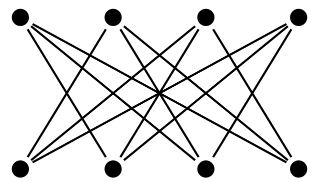

Figure 1.1: Order diagram of the standard example S_4.

Like dimension, local dimension is monotone and subadditive. See [32] for the proofs.

Another dimension variant, called t-dimension, was introduced by Novák [44]. For a poset P and $t \in \mathbb{N}$ with $t \ge 2$, the t-dimension of P, denoted $\dim_t(P)$, is the smallest cardinal d such that $P \hookrightarrow \mathbf{t}^d$. It follows immediately from the definition that t-dimension is monotone and subadditive for every t. By the pigeonhole principle, for any poset P with cardinality n, $\dim_t P \ge \lceil \log_t n \rceil$, and this bound is sharp for all n.

The most interesting case is $t = 2$, as $\dim_2(P)$ is equal to the smallest d such that $P \hookrightarrow \mathcal{Q}^d$. For example, Sperner's theorem states that

$$\dim_2(A_n) = \min\left\{m : \binom{m}{\lfloor m/2 \rfloor} \ge n\right\} = \log n + \frac{1}{2}\log\log n + O(1).$$

We also have $\dim_2(\mathbf{n}) = n - 1$. Clearly $\dim_2(P) \le |P|$ for every poset P, as the map $a \mapsto \{x \le a\}$ is a monotone function from P to the Boolean lattice of dimension $|P|$). This bound is sharp for $n \ge 2$, because $\dim_2(\mathbf{n{-}1} + \mathbf{1}) = n$; see exercise 10.2.6 in [55].

CHAPTER 2

THE DIMENSION OF THE DIVISIBILITY ORDER

All of the results in this chapter are joint work with Victor Souza. An abridged version of this chapter [41] has been submitted for publication to the *Journal of Combinatorial Theory, Series A*.

2.1 Introduction

It is a basic fact that the divisibility relation defines a partial order on $\mathbb{N}$. For any subset $S \subseteq \mathbb{N}$, denote by $\mathcal{D}_S$ the divisibility poset restricted to the set S. Properties of the divisibility order have been studied, for example, by Cameron and Erdős [12]. Surprisingly, the dimension of the divisibility order, as far as we know, has not been considered in the literature. Since the dimension of $\mathcal{D}_{\mathbb{N}}$ is infinite, we are usually concerned with the case where S is finite. Indeed, we are primarily interested in the case $S = [n] := \{1, \ldots, n\}$. In our main result, we determine the growth of $\dim\left(\mathcal{D}_{[n]}\right)$ as n goes to infinity up to a $\ln \ln n$ factor.

Theorem 28. *The dimension of $\mathcal{D}_{[n]}$, the divisibility order on $[n]$, satisfies, as $n \to \infty$,*

$$\left(\tfrac{1}{16} - o(1)\right)\frac{(\ln n)^2}{(\ln \ln n)^2} \leq \dim\left(\mathcal{D}_{[n]}\right) \leq \left(4 + o(1)\right)\frac{(\ln n)^2}{\ln \ln n}. \tag{2.1}$$

Unlike with other natural suborders of $\mathcal{D}_{\mathbb{N}}$, such as the set of divisors of a given natural number, the dimension of $\mathcal{D}_{[n]}$ doesn't seem to reduce to a well-known number-theoretic function. For example, the poset of divisors of n (which is the interval $[1, n]$ with respect to the divisibility order) is just a product of $\omega(n)$ chains and so has dimension $\omega(n)$, where $\omega(n)$ is the number of distinct prime factors of n. But the set $[n]$ is an interval in the usual order on the integers, and it displays a nontrivial interaction with the divisibility order when regarding the dimension.

We prove Theorem 28 by embedding a suborder of the hypercube into $\mathcal{D}_{[n]}$, then embedding $\mathcal{D}_{[n]}$ into a product of simple posets and showing that each of them has small dimension. We observe that this same idea works, with small modifications, in a variety of circumstances. For example, we have an analogue of Theorem 28 for 2-dimension.

Theorem 29. *The 2-dimension of of $\mathcal{D}_{[n]}$ satisfies, as $n \to \infty$,*

$$\left(\tfrac{1}{16} - o(1)\right)\frac{(\ln n)^2}{(\ln\ln n)^2} \leq \dim_2\left(\mathcal{D}_{[n]}\right) \leq \left(\tfrac{4}{3}e\pi^2 + o(1)\right)\frac{(\ln n)^2}{\ln\ln n}. \tag{2.2}$$

We also consider other natural choices for subsets of $\mathbb{N}$ to bound the dimension. Some sets like $(n^\alpha, n]$ and $a[n]+b = \{ak+b : k \in [n]\}$ behave similarly to $[n]$ with respect to the dimension. On the other hand, in Section 2.5 we shall see that the dimension of the divisibility order on the set $(\alpha n, n]$ behaves quite differently. In fact, if $\alpha \geq 1/2$, then $\mathcal{D}_{(\alpha n, n]}$ is an antichain and thus has dimension 2. Using a result of Scott and Wood [47] on posets with bounded degree, we show that $\mathcal{D}_{(\alpha n,n]}$ has bounded dimension, and that, as $\alpha \to 0$,

$$\sup_{n\in\mathbb{N}} \dim\left(\mathcal{D}_{(\alpha n,n]}\right) \leq \tfrac{1}{\alpha}\left(\ln(\tfrac{1}{\alpha})\right)^{1+o(1)}. \tag{2.3}$$

In Section 2.5, we prove an analogue of a result by Füredi and Kahn [24] for 2-dimension and use it to show that $\dim_2\left(\mathcal{D}_{(\alpha n,n]}\right) = \Theta_\alpha(\ln n)$ as $n \to \infty$, and that the same holds for t-dimension for any $t \geq 2$.

In Section 2.6, we consider the fractional dimension of $\mathcal{D}_{[n]}$, a linear programming relaxation of the Dushnik-Miller dimension, introduced by Brightwell and Scheinerman [11]. In contrast to the dimension and 2-dimension, we can recover the correct asymptotic behaviour for the fractional dimension.

Theorem 30. *The fractional dimension of of $\mathcal{D}_{[n]}$ satisfies, as $n \to \infty$,*

$$\dim^{\star}\left(\mathcal{D}_{[n]}\right) \sim \frac{\ln n}{\ln \ln n}. \tag{2.4}$$

Finally, in Section 2.7, we observe that our techniques work for other divisibility posets as well, such as the set of ideals in a number field and the set of monic polynomials over a finite field.

While combinatorial properties of the divisibility poset have been studied before, results often are not stated in the language of partial orders. For example, Cameron and Erdős [12] called antichains in $\mathcal{D}_{[n]}$ *primitive sets*, and conjectured that the number of primitive subsets of $\mathcal{D}_{[n]}$ is $(\alpha + o(1))^n$ for some constant α. This conjecture was recently proven by Angelo [1]. Continuing this work, Liu, Pach, and Palincza [42] proved that the number of maximum-size primitive subsets of $[n]$ is $(\beta + o(1))^n$ for some constant β, and gave algorithms for computing both $\alpha \approx 1.57$ and $\beta \approx 1.318$. They also showed that the number of strong antichains in $\mathcal{D}_{[2,n]}$ is $2^{\pi(n)} \cdot e^{(1+o(1))\sqrt{n}}$, where a strong antichain in a poset P is a subset of P such that no two elements have a common lower bound in P. We hope this note motivates further work on the combinatorial aspects of the divisibility order.

2.2 Suborders of the hypercube

Our proof strategy for Theorem 28 consists of comparing the dimension of $\mathcal{D}_{[n]}$ with the dimension of suborders of the hypercube. In this section, we review the theory of Dushnik [16] that describes the dimension of suborders of the hypercube with another combinatorial object: suitable sets of permutations.

We write $\mathcal{Q}^n$ for the n-dimensional hypercube, that is, the subset lattice of $[n]$. For any set $A \subseteq [n]$, $\mathcal{Q}^n_A$ denotes the suborder of $\mathcal{Q}^n$ consisting of the subsets $X \subseteq [n]$ with $|X| \in A$. We write $\mathcal{Q}^n_{a,b}$ instead of $\mathcal{Q}^n_{\{a,b\}}$ for simplicity.

The poset of multisets of numbers in $[n]$, ordered by inclusion with multiplicity, is denoted $\mathcal{M}^n$. For any $A \subseteq \mathbb{N}$, we denote by $\mathcal{M}^n_A$ the suborder of $\mathcal{M}^n$ of multisets whose cardinalities with multiplicity are in A, and by $\widetilde{\mathcal{M}}^n_A$ the suborder of $\mathcal{M}^n$ consisting of all finite multisets whose ground sets have cardinalities in A, ignoring multiplicity. Note that all the posets mentioned are finite, with the exception of $\widetilde{\mathcal{M}}^n_A$. Usually, we take $A = [0, k]$ or $A = \{1, k\}$.

We will now prove a slightly stronger version of a lemma by Dushnik [16], which characterises the dimension of these posets. To state this lemma, we need a few more definitions.

A *pointed k-subset of* $[n]$ is an ordered pair (A, a) with $a \in A$, $A \subset [n]$ and $|A| = k$. A set S of permutations of $[n]$ is called *k-suitable* if, for every pointed k-subset (A, a) of $[n]$, there is a $\sigma \in S$ such that $b \leq_\sigma a$ for every $b \in A$. We say that such a σ *covers* the pointed set (A, a).

For any pair $1 \leq k \leq n \in \mathbb{N}$, $N(n, k)$ is defined as the minimum cardinality of a k-suitable set of permutations of $[n]$. It is clear that $N(n, 1) = 1$ and that $N(n, 2) = 2$. We also have $N(n, k) \geq k$, since each permutation covers only one of the k pointed sets on a given ground set. Because every k-suitable set with $2 \leq k \leq n$ is also $(k-1)$-suitable and the restriction of a k-suitable set of permutations of $[n]$ with $1 \leq k \leq n-1$ to $[n-1]$ is still k-suitable, $N(n, k)$ is monotone increasing in both arguments. Later, we will provide upper and lower bounds for $N(n, k)$.

Lemma 31. *For every n and every $k \leq n - 1$,*

$$\dim\left(\mathcal{Q}^n_{1,k}\right) = \dim\left(\mathcal{Q}^n_{[0,k]}\right) = \dim\left(\mathcal{M}^n_{[0,k]}\right) = \dim\left(\widetilde{\mathcal{M}}^n_{[0,k]}\right) = N(n, k+1).$$

Proof. We show this by proving the following sequence of inequalities:

$$\begin{aligned} N(n,k+1) \leq \dim\left(\mathcal{Q}^n_{1,k}\right) \leq \dim\left(\mathcal{Q}^n_{[0,k]}\right) \leq \\ \dim\left(\mathcal{M}^n_{[0,k]}\right) \leq \dim\left(\widetilde{\mathcal{M}}^n_{[0,k]}\right) \leq N(n,k+1). \end{aligned} \tag{2.5}$$

To show that $N(n,k+1) \leq \dim\left(\mathcal{Q}^n_{1,k}\right)$, observe that every realiser $\mathcal{L}$ of $\mathcal{Q}^n_{1,k}$ induces a $(k+1)$-suitable set of permutations of the one-element subsets of $[n]$ in the following way. For every $L \in \mathcal{L}$, let σ_L permutation of $[n]$ induced by the restriction of L to $[n]^{(1)}$. Now, for every pointed $(k+1)$-set $(\{a_1,\ldots,a_{k+1}\},a_{k+1})$, there is an $L \in \mathcal{L}$ such that $\{a_1,\ldots,a_k\} \leq_L \{a_{k+1}\}$. By transitivity, $\{a_i\} \leq_L \{a_{k+1}\}$ and hence $a_i \leq_{\sigma_L} a_{k+1}$ for all $i \in [k]$, so $\{\sigma_L : L \in \mathcal{L}\}$ is $(k+1)$-suitable.

The inequalities $\dim\left(\mathcal{Q}^n_{1,k}\right) \leq \dim\left(\mathcal{Q}^n_{[0,k]}\right) \leq \dim\left(\mathcal{M}^n_{[0,k]}\right) \leq \dim\left(\widetilde{\mathcal{M}}^n_{[0,k]}\right)$ hold because each poset embeds into the next.

Now to prove that $\dim\left(\widetilde{\mathcal{M}}^n_{[0,k]}\right) \leq N(n,k+1)$ we just have to show how to extend a $(k+1)$-suitable set of permutations to a realiser of $\widetilde{\mathcal{M}}^n_{[0,k]}$ with the same cardinality.

Let S be a $(k+1)$-suitable set of permutations of $[n]$. For each $\sigma \in S$, let L_σ be the colexicographic order on $\widetilde{\mathcal{M}}^n_{[0,k]}$ with respect to σ. In other words, if A and B are two distinct finite multisets of numbers in $[n]$ whose ground sets have cardinality at most k and x is the σ-greatest element of $A \cup B$ whose multiplicity in A differs from its multiplicity in B, then $A <_{L_\sigma} B$ if x has greater multiplicity in B than in A and $B <_{L_\sigma} A$ if x has greater multiplicity in A.

If $A \subset B$, then $A <_{L_\sigma} B$ for every $\sigma \in S$, so L_σ is a linear order that extends the order on $\widetilde{\mathcal{M}}^n_{[0,k]}$. If A and B are incomparable in $\widetilde{\mathcal{M}}^n_{[0,k]}$, then there exists an $x \in B$ whose multiplicity in B is greater than its multiplicity in A. Since S is ℓ-suitable for every $1 \leq \ell \leq k+1$, we can find a $\sigma \in S$ that covers $(X \cup \{x\}, x)$, where X is the underlying set of A. Hence $A <_{L_\sigma} B$. Similarly, there exists a $y \in A$ whose multiplicity in A is greater than its multiplicity in B, so we can find a $\tau \in S$

such that $B <_{L_\tau} A$. Therefore $\{L_\sigma : \sigma \in S\}$ is a realiser of $\widetilde{\mathcal{M}}^n_{[0,k]}$. $\square$

The following theorem by Dushnik [16] gives the exact value of $N(n,k)$, when k is at least of order $2\sqrt{n}$. Note that, by Lemma 31, we also obtain the exact dimension of $\mathcal{Q}^n_{1,k}$ and related posets.

Theorem 32. *For any j and k with $2 \leq j \leq \sqrt{n}$ and*

$$\left\lfloor \frac{n}{j} \right\rfloor + j - 1 \leq k \leq \left\lfloor \frac{n}{j-1} \right\rfloor + j - 3,$$

we have

$$N(n,k) = n - j + 1$$

. In particular, if $2\sqrt{n} - 1 \leq k < n$, then $N(n,k) \geq n - \sqrt{n}$. $\square$

Spencer proved in [50] that, for all fixed $k \geq 3$, $N(n,k) = \Theta_k(\ln \ln n)$ as $n \to \infty$. However, the implicit constants grow exponentially in k.

The following bound, which was proved in a slightly stronger form by Füredi and Kahn [24], is more useful when $\ln \ln n \ll k \ll \sqrt{n}$, which is the relevant magnitude for the proof of the upper bound in Theorem 28.

Lemma 33. *For all $1 \leq k \leq n$, $N(n,k) \leq \lceil k^2 \ln n \rceil$.*

Proof. The proof is probabilistic. Fix a natural number s and choose s permutations of $[n]$ independently and uniformly at random. The probability that a given pointed k-subset isn't covered by any of these permutations is $(1 - 1/k)^s < e^{-s/k}$. Since the total number of pointed k-subsets of $[n]$ is $k\binom{n}{k} \leq n^k$, the expected number of pointed k-subsets not covered is less than $n^k e^{-s/k} \leq 1$ when $s \geq k^2 \ln n$, so $N(n,k) \leq \lceil k^2 \ln n \rceil$. $\square$

2.3 The dimension of the divisibility order on $[n]$

In this section, we provide a proof of Theorem 28. Additionally, we give similar lower and upper bounds on the dimension of $\mathcal{D}_S$ for other interesting subsets of $\mathbb{N}$. The following principle will be useful to give upper bounds on the dimension and the 2-dimension of $\mathcal{D}_{[n]}$.

Lemma 34. *Let $P_1, \ldots, P_k$ be a partition of the primes in $[n]$ and let Q_i be the set of numbers in $[n]$ that can be written as a (possibly empty) product of powers of primes in P_i. Then*

$$\mathcal{D}_{[n]} \hookrightarrow \mathcal{D}_{Q_1} \times \ldots \times \mathcal{D}_{Q_k}.$$

Proof. As $P_1, \ldots, P_k$ is a partition the primes in $[n]$, any number $a \in [n]$ can be factored uniquely as $a = q_1 \ldots q_k$, where $q_i \in Q_i$. Thus, the mapping $a \mapsto (q_1, \ldots, q_k)$ is well defined and we claim that it the poset embedding we need. Indeed, if $a = q_1 \ldots q_k$ and $b = r_1 \ldots r_k$, with $q_i, r_i \in Q_i$, then $a|b$ if and only if $q_i|r_i$ for all i. $\square$

Denote by p_k the k^{th} prime number and by $\pi(x)$ the number of prime numbers less than or equal to x. We only use standard estimates for these functions: $p_k = (1 + o(1))k \ln k$ and $\pi(x) = (1 + o(1))\frac{x}{\ln x}$. Now, we have all the ingredients we need to prove Theorem 28.

Proof of Theorem 28. First we prove the lower bound. Observe that, if k is an integer such that every product of at most $2\sqrt{k}$ distinct elements of $\{p_1, p_2, \ldots, p_k\}$ is in $[n]$, then we have an embedding $\mathcal{Q}^k_{[0,\lfloor 2\sqrt{k}\rfloor]} \hookrightarrow \mathcal{D}_{[n]}$. The image of the embedding is the set of all products of at most $2\sqrt{k}$ of the first k primes, and is contained in $\mathcal{D}_{[n]}$ if $p_{k-2\sqrt{k}+1} \ldots p_k \leq n$. It follows by Theorem 32 and Lemma 31 that

$$\dim\left(\mathcal{D}_{[n]}\right) \geq \dim\left(\mathcal{Q}^k_{[0,2\sqrt{k}]}\right) \geq k - \sqrt{k}. \tag{2.6}$$

This condition for this embedding to exist is satisfied if $p_k^{2\sqrt{k}} \le n$. Now fix $\alpha < 1/16$ and let $k = \left\lfloor \alpha \left(\frac{\ln n}{\ln\ln n}\right)^2 \right\rfloor$. Using the estimate $p_k = k^{1+o(1)}$, we obtain

$$p_k^{2\sqrt{k}} = k^{(2+o(1))\sqrt{k}} \le \left(\frac{\ln n}{\ln\ln n}\right)^{(4\sqrt{\alpha}+o(1))\frac{\ln n}{\ln\ln n}} \ll \left(\frac{\ln n}{\ln\ln n}\right)^{\frac{\ln n}{\ln\ln n}} < n, \tag{2.7}$$

whenever n is sufficiently large. Letting α approach $1/16$ from below, we obtain

$$\dim\left(\mathcal{D}_{[n]}\right) \ge (1-o(1))k = \left(\tfrac{1}{16} - o(1)\right)\frac{(\ln n)^2}{(\ln\ln n)^2}. \tag{2.8}$$

To prove the upper bound, set $\varepsilon = \varepsilon(n) > 0$ to be chosen later. Let S be the set of all elements of $[n]$ that can be factored into primes less than $(\varepsilon \ln n)^2$ and let R be the set of all elements whose prime factors are all at least $(\varepsilon \ln n)^2$. By Lemma 34, we have an embedding $\mathcal{D}_{[n]} \hookrightarrow \mathcal{D}_S \times \mathcal{D}_R$, so $\dim\left(\mathcal{D}_{[n]}\right) \le \dim(\mathcal{D}_S) + \dim(\mathcal{D}_R)$. The poset $\mathcal{D}_S$ can then be embedded in the product of $\pi((\varepsilon\ln n)^2)$ chains (namely the powers of p for each small prime p) and so

$$\dim(\mathcal{D}_S) \le \pi((\varepsilon \ln n)^2) = \left(\tfrac{\varepsilon^2}{2} + o(1)\right)\frac{(\ln n)^2}{\ln\ln n}. \tag{2.9}$$

We further partition the large primes. Let $L = \left\lfloor \log_2\left(\frac{\ln n}{\ln\ln n + \ln\varepsilon}\right)\right\rfloor$ and, for each $0 \le i < L$, let $\theta_i = n^{2^{-i}}$, and R_i be the set of numbers in $[n]$ whose prime factors all lie in the interval $\left(\theta_{i+1}, \theta_i\right]$. Lemma 34 now implies that $\dim(\mathcal{D}_R) \le \sum_{i=0}^{L-1} \dim(\mathcal{D}_{R_i})$. For every i, the prime factors of each element of R_i form a multiset of elements of $\left[\lfloor\theta_i\rfloor\right]$ of cardinality strictly less than 2^{i+1}. For all $a, b \in \mathbb{N}$, a divides b if and only if the multiset of prime factors of a is a submultiset of the multiset of prime factors of b, so $\mathcal{D}_{R_i} \hookrightarrow \mathcal{M}_{[0,2^{i+1}-1]}^{\lfloor\theta_i\rfloor}$. By Lemma 31 and Lemma 33, we have

$$\dim(\mathcal{D}_{R_i}) \le \dim\left(\mathcal{M}_{[0,2^{i+1}-1]}^{\lfloor\theta_i\rfloor}\right) = N\left(\lfloor n^{2^{-i}}\rfloor, 2^{i+1}\right) \le 4\cdot 2^i \ln n + 1. \tag{2.10}$$

Therefore, we observe that

$$\begin{aligned}\dim(\mathcal{D}_R) &\le \sum_{i=0}^{L-1} \dim(\mathcal{D}_{R_i}) \le \sum_{i=0}^{L-1} 4\cdot 2^i \ln n + L \le 4\cdot 2^L \ln n + L \\ &\le \frac{4(\ln n)^2}{\ln\ln n + \ln\varepsilon} + O_\varepsilon(\ln\ln n) = (4+o_\varepsilon(1))\frac{(\ln n)^2}{\ln\ln n}.\end{aligned} \tag{2.11}$$

Next, we combine inequalities (2.9) and (2.11) to obtain

$$\dim\left(\mathcal{D}_{[n]}\right) \le \dim(\mathcal{D}_S) + \dim(\mathcal{D}_R) \le \left(\tfrac{\varepsilon^2}{2} + 4 + o_\varepsilon(1)\right)\frac{(\ln n)^2}{\ln\ln n}. \tag{2.12}$$

Now we just let $\varepsilon = \varepsilon(n)$ approach 0 sufficiently slowly as $n\to\infty$, and the result follows. □

With some modifications, we can adapt our proof to other settings. We begin by looking at $(n^\alpha, n]$.

Corollary 35. *For any fixed $\alpha \in (0,1)$, as $n\to\infty$,*

$$\left(\tfrac{(1-\alpha)^2}{16} - o_\alpha(1)\right)\frac{(\ln n)^2}{(\ln\ln n)^2} \le \dim\left(\mathcal{D}_{(n^\alpha,n]}\right) \le \left(4+o(1)\right)\frac{(\ln n)^2}{\ln\ln n}.$$

Proof. These bounds follow from the fact that $\mathcal{D}_{[\lfloor n^{1-\alpha}\rfloor]} \hookrightarrow \mathcal{D}_{(n^\alpha,n]} \hookrightarrow \mathcal{D}_{[n]}$. The first embedding $\mathcal{D}_{[\lfloor n^{1-\alpha}\rfloor]} \hookrightarrow \mathcal{D}_{(n^\alpha,n]}$ is the map $x\mapsto \lceil n^\alpha\rceil x$ and the second is just the inclusion map. □

Also of interest is the arithmetic progression $a[n]+b = \{ak+b : k\in[n]\}$.

Corollary 36. *For any fixed a and b, as $n\to\infty$,*

$$\left(\tfrac{1}{16} - o_{a,b}(1)\right)\frac{(\ln n)^2}{(\ln\ln n)^2} \le \dim\left(\mathcal{D}_{a[n]+b}\right) \le \left(4+o_{a,b}(1)\right)\frac{(\ln n)^2}{\ln\ln n}.$$

Proof. Since the divisibility poset is dilation-invariant, we may assume a and b are coprime. Since $\mathcal{D}_{a[n]+b} \hookrightarrow \mathcal{D}_{[an+b]}$, the upper bound from Theorem 28 holds. For the lower bound, recall that the multiplicative group $(\mathbb{Z}/a\,\mathbb{Z})^\times$ has order $\varphi(a)$, and so $b^\ell \equiv b \pmod{a}$ whenever $\ell \equiv 1 \pmod{\varphi(a)}$. This implies that $\mathcal{Q}^k_{1,\ell} \hookrightarrow \mathcal{D}_{a[n]+b}$, where $\ell = \varphi(a)\left\lceil \frac{2\sqrt{k}-1}{\varphi(a)} \right\rceil + 1$, i.e., $2\sqrt{k}$ rounded up to the nearest integer congruent to 1 modulo $\varphi(a)$, as long as k is not too big. If we denote by $p_{a,b,m}$ the m^{th} prime congruent to $b \pmod{a}$, then, by the prime number theorem for arithmetic progressions, we have $p_{a,b,m} \sim \varphi(a) m \ln m = m^{1+o_{a,b}(1)}$. In the spirit of Theorem 28, we have an embedding if $p^\ell_{a,b,k} \leq an+b$. Since $p^\ell_{a,b,k} = k^{(2+o_{a,b}(1))\sqrt{k}}$, a lower bound of the same form as in Theorem 28 holds asymptotically for $\dim\left(\mathcal{D}_{a[n]+b}\right)$. □

As the last result in this section, we observe that the dimension of $\mathcal{D}_{[n]}$ is supported on the set of squarefree elements. Let $\mathcal{S}$ be the set of squarefree integers. We say that a set $A \subseteq \mathbb{N}$ is closed under taking divisors if, for all $a \in A$ and $d \in \mathbb{N}$, $d \mid a$ implies $d \in A$. The next result shows that the set of squarefree numbers has full dimension inside a set closed under taking divisors. In particular, $\dim\left(\mathcal{D}_{[n]}\right) = \dim\left(\mathcal{D}_{[n]\cap\mathcal{S}}\right)$.

Theorem 37. *If $A \subseteq \mathbb{N}$ is closed under taking divisors, then $\dim\left(\mathcal{D}_A\right) = \dim\left(\mathcal{D}_{A\cap\mathcal{S}}\right)$.*

Proof. Let $\{L_1, \ldots, L_d\}$ be a realiser of $\mathcal{D}_{A\cap\mathcal{S}}$. For each $a \in A$, we define a squarefree factorisation of a as follows. First, let $a_1 = \operatorname{rad}(a)$, where $\operatorname{rad}(a)$ is the greatest squarefree factor of a. Next, for each $i \geq 1$, let $a_{i+1} = \operatorname{rad}\left(\frac{a}{a_1 \ldots a_i}\right)$. For every $j \in [d]$, let R_j be a linear extension of $\mathcal{D}_A$ defined in the following way: consider the mapping $\phi(a) = \{a_i\}$, we say that $a \leq_{R_j} b$ if $\phi(a) \leq \phi(b)$ in the lexicographic order on the space of sequences induced by L_j. We claim that $\{R_1, \ldots, R_d\}$ is a realiser of $\mathcal{D}_A$.

Indeed, let $a, b \in A$, and $\phi(a) = \{a_i\}$, $\phi(b) = \{b_i\}$ and $k = \min\{k : a_k \neq b_k\}$. If $a|b$, then $a_k \mid b_k$, so R_j is a linear extension of $\mathcal{D}_A$. If a and b are incomparable, then there is a $j \in [d]$ such that $a_k \leq_{L_j} b_k$, and thus $a \leq_{R_j} b$. Therefore,

$\dim(\mathcal{D}_A) \le \dim(\mathcal{D}_{A\cap\mathcal{S}})$. □

2.4 The 2-dimension of the divisibility order on $[n]$

The purpose of this section is to prove Theorem 29.

Kierstead [31] showed that the 2-dimension of two-layer suborders of $\mathcal{Q}^n$ can be described in terms of suitable sets of subsets of $[n]$.

For any $2 \le k \le n$, $N_2(n,k)$ is defined as the minimum cardinality of a set S of subsets of $[n]$ such that, for any pointed k-subset (A,a) of $[n]$, there exists a set $B \in S$ such that $A \cap B = \{a\}$. By analogy with $N(n,k)$, we call such a set a k-suitable set of subsets. The following partial analogue to Lemma 31 is essentially due to Kierstead [31].

Lemma 38. *For every n and every $k \le n+1$,*

$$\dim_2\left(\mathcal{Q}^n_{1,k}\right) = \dim_2\left(\mathcal{Q}^n_{[0,k]}\right) = N_2(n,k+1).$$

Proof. To show this, we prove the following sequence of inequalities:

$$N_2(n,k+1) \le \dim_2\left(\mathcal{Q}^n_{1,k}\right) \le \dim_2\left(\mathcal{Q}^n_{[0,k]}\right) \le N_2(n,k+1). \tag{2.13}$$

To show that $N_2(n,k+1) \le \dim_2\left(\mathcal{Q}^n_{1,k}\right)$, let $f : \mathcal{Q}^n_{1,k} \hookrightarrow \mathcal{Q}^d$ be an embedding. For each $i \in [d]$, let X_i be the set of all $j \in [n]$ such that $i \in f\big(\{j\}\big)$. We claim that $\{X_i : i \in [d]\}$ is $(k+1)$-suitable. Indeed, let (A,a) be a pointed $(k+1)$-subset of $[n]$. Since $f\big(\{a\}\big) \not\subseteq f\big(A \setminus \{a\}\big)$, there is an $i \in [d]$ such that $i \in f\big(\{a\}\big)$ but $i \notin f\big(A \setminus \{a\}\big)$. It follows that $i \notin f\big(\{b\}\big)$ for any $b \in A \setminus \{a\}$, so $X_i \cap A = \{a\}$.

The second inequality follows by monotonicity from the fact that $\mathcal{Q}^n_{1,k} \hookrightarrow \mathcal{Q}^n_{[0,k]}$.

To show that $\dim_2\left(\mathcal{Q}^n_{[0,k]}\right) \le N_2(n,k+1)$, let $\{X_1, X_2, \ldots, X_d\}$ be a $(k+1)$-suitable set of subsets of $[n]$. Define a map $f : \mathcal{Q}^n_{[0,k]} \to \mathcal{Q}^d$,

$f(A) = \{i \in [d] : A \cap X_i \neq \emptyset\}$. We claim that such map is an embedding. Indeed, if $A \subseteq B \subseteq [n]$, then $f(A) \subseteq f(B)$. Now, let $A \not\subseteq B$, where $a \in A$, but $a \notin B$. Since the family $\{X_i\}$ is $(k+1)$-suitable, there is i with $(B \cup \{a\}) \cap X_i = \{a\}$, therefore $X_i \cap B = \emptyset$, so $i \notin f(B)$, whereas $i \in f(A)$. In other words, $A \not\subseteq B$ implies $f(A) \not\subseteq f(B)$. Thus f is an embedding and $\dim_2\left(\mathcal{Q}^n_{[0,k]}\right) \leq d$. □

An analogue of Lemma 33 can be proved via the first moment method by taking random subsets of $[n]$ with each element having probability $\frac{1}{k}$ of being chosen. This leads to a theorem of Kierstead [31].

Theorem 39. *For all* $2 \leq k \leq n$, $N_2(n,k) \leq \lceil ek^2 \ln n \rceil$. □

Because the 2-dimension of a poset depends in part on its cardinality, we can't ignore the non-squarefree elements of $[n]$. This means that the analogue of Theorem 37 for 2-dimension cannot hold. The following lemma will help us deal with non-squarefree elements.

Lemma 40. *For all* $k \leq n-1$, $\dim_2\left(\mathcal{M}^n_{[0,k]}\right) < e(\frac{\pi^2}{6}k^2 + 2k\ln k + 3k)\ln n + k$.

Proof. For each $A \in \mathcal{M}^n_{[0,k]}$ and each $i \in [k]$, let A^i be the set of all elements of A of multiplicity at least i. Observe that $i|A^i| \leq |A| \leq k$, so $|A^i| \leq k/i$. For any two multisets A and B in $\mathcal{M}^n_{[0,k]}$, $A \subseteq B$ if and only if $A^i \subseteq B^i$ for every $i \in [k]$. Hence the map $A \mapsto (A^1, \ldots, A^k)$ is a poset embedding $\mathcal{M}^n_{[0,k]} \hookrightarrow \prod_{i=1}^{k} \mathcal{Q}^n_{[0,\lfloor k/i \rfloor]}$. Therefore, by Theorem 39, we have

$$\begin{aligned}
\dim_2\left(\mathcal{M}^n_{[0,k]}\right) &\leq \sum_{i=1}^{k} \dim_2\left(\mathcal{Q}^n_{[0,\lfloor k/i \rfloor]}\right) \leq \sum_{i=1}^{k} N_2(n, \lfloor k/i \rfloor + 1) \\
&< e \ln n \sum_{i=1}^{k} \left(\frac{k^2}{i^2} + 2\frac{k}{i} + 1\right) + k \\
&\leq e\big(\tfrac{\pi^2}{6}k^2 + 2k(\ln k + 1) + k\big) \ln n + k.
\end{aligned} \tag{2.14}$$

□

Since $\dim\left(\mathcal{D}_{[n]}\right) \le \dim_2\left(\mathcal{D}_{[n]}\right)$, Theorem 28 already provides a lower bound for $\dim_2\left(\mathcal{D}_{[n]}\right)$. Therefore, in order to prove Theorem 29, only the proof of the upper bound is required.

Proof of Theorem 29. The proof is essentially the same as that the upper bound of Theorem 28, so we will omit some of the details. Fix $\varepsilon > 0$. Let S be the set of all elements of $[n]$ whose prime factors are all at most $\varepsilon \ln n$ and R be the set of all elements whose prime factors are all greater than $\varepsilon \ln n$.

The poset $\mathcal{D}_S$ can be embedded into the product of $\pi(\varepsilon \ln n)$ chains, each of size at most $1 + \log_2 n$. Since the 2-dimension of a chain of length ℓ is $\ell - 1$, we have

$$\dim_2\left(\mathcal{D}_S\right) \le \pi(\varepsilon \ln n)\log_2 n = \left(\frac{\varepsilon}{\ln 2} + o(1)\right)\frac{(\ln n)^2}{\ln\ln n}. \tag{2.15}$$

Let $L = \left\lceil \log_2\left(\frac{\ln n}{\ln\ln n + \ln\varepsilon}\right)\right\rceil$. For each i from 0 to $L-1$, let $\theta_i = n^{2^{-i}}$ and R_i be the set of elements of $[n]$ whose prime factors all lie in the interval $\left(\theta_{i+1}, \theta_i\right]$. Just as before, we have embeddings $\mathcal{D}_R \hookrightarrow \prod_{i=1}^{L-1} \mathcal{D}_{R_i}$ and $\mathcal{D}_{R_i} \hookrightarrow \mathcal{M}^{\theta_i}_{[0,2^{i+1}-1]}$. By Lemma 40, we have

$$\dim_2\left(\mathcal{D}_{R_i}\right) \le \dim_2\left(\mathcal{M}^{\theta_i}_{[0,2^{i+1}-1]}\right) \le \left(\frac{2e\pi^2}{3} + o(1)\right)2^i \ln n. \tag{2.16}$$

Therefore, we obtain the following bound:

$$\begin{aligned}\dim_2\left(\mathcal{D}_R\right) &\le \sum_{i=0}^{L-1} \dim_2\left(\mathcal{D}_{R_i}\right) \le \frac{2e\pi^2}{3}2^L \ln n + o(2^L \ln n) \\ &\le \left(\frac{4e\pi^2}{3} + o(1)\right)\frac{(\ln n)^2}{\ln\ln n + \ln\varepsilon}.\end{aligned} \tag{2.17}$$

Finally, we have

$$\dim_2\left(\mathcal{D}_{[n]}\right) \le \dim_2\left(\mathcal{D}_S\right) + \dim_2\left(\mathcal{D}_R\right) \le \left(\frac{\varepsilon}{\ln 2} + \tfrac{4}{3}e\pi^2 + o_\varepsilon(1)\right)\frac{(\ln n)^2}{\ln\ln n} \tag{2.18}$$

for every fixed $\varepsilon > 0$. Again we can let $\varepsilon = \varepsilon(n)$ approach 0 sufficiently slowly as not to interfere with the $o_\varepsilon(1)$ term, and the result follows. $\square$

We note that the analogues of Corollaries 35 and 36 hold for 2-dimension as well.

Corollary 41. *For any fixed $\alpha \in (0,1)$ $a, b \in \mathbb{N}$, as $n \to \infty$,*

$$\left(\tfrac{(1-\alpha)^2}{16} - o_\alpha(1)\right)\frac{(\ln n)^2}{(\ln\ln n)^2} \leq \dim_2\left(\mathcal{D}_{(n^\alpha, n]}\right) \leq \left(\tfrac{4}{3}e\pi^2 + o(1)\right)\frac{(\ln n)^2}{\ln\ln n},$$

$$\left(\tfrac{1}{16} - o_{a,b}(1)\right)\frac{(\ln n)^2}{(\ln\ln n)^2} \leq \dim_2\left(\mathcal{D}_{a[n]+b}\right) \leq \left(\tfrac{4}{3}e\pi^2 + o_{a,b}(1)\right)\frac{(\ln n)^2}{\ln\ln n}. \qquad \square$$

The proofs are nearly identical to the ones for dimension, so we omit them.

2.5 Using Bounded Degree

In previous sections, we have already considered the dimension of the divisibility order on sets other than $[n]$, such as $(n^\alpha, n]$ or $a[n]+b$. The proof of Theorem 28 can be adapted to those cases after some small modifications. In this section, we will study the dimension of $(\alpha n, n]$, whose dimension behaves in a different manner. Indeed, $\mathcal{D}_{(\alpha n, n]}$ is antichain when $\alpha > 1/2$, for instance, so has dimension at most 2.

The comparability graph of a poset P is the graph with vertex set P where two elements are connected if they are comparable in P. A theorem by Füredi and Kahn [24] states that a poset whose comparability graph has maximum degree Δ has dimension less than $50\Delta(\ln\Delta)^2$. This bound was recently improved by Scott and Wood [47], who showed that the maximum dimension of a poset of maximum degree Δ is $\Delta(\ln\Delta)^{1+o(1)}$ as $\Delta \to \infty$.

The comparability graph of $\mathcal{D}_{(\alpha n, n]}$ has maximum degree at most $1/\alpha + 1$. Indeed, let $x \in (\alpha n, n]$ with $x = \beta n$ for some $\beta \in (\alpha, 1]$. The number of elements that divide x is at most β/α and the number divisible by x is at most $1/\beta$, so the degree of x in the comparability graph is at most $1/\beta + \beta/\alpha \leq 1 + 1/\alpha$. Therefore, as $\alpha \to \infty$, we have

$$\sup_{n\in\mathbb{N}} \dim\big(\mathcal{D}_{(\alpha n,n]}\big) \leq \tfrac{1}{\alpha}\big(\ln(\tfrac{1}{\alpha})\big)^{1+o(1)}. \tag{2.19}$$

We note that the t-dimension of $\mathcal{D}_{(\alpha n,n]}$ behaves very differently from the ordinary dimension, since this poset has unbounded cardinality and hence unbounded t-dimension for every $t \geq 2$.

For a poset P and $x \in P$, we define the outdegree of x as $\big|\{y \in P : y > x\}\big|$ and the indegree of x as $\big|\{y \in P : y < x\}\big|$. Another theorem by Füredi and Kahn [24] says that a poset of cardinality n and maximum outdegree υ has dimension at most $\lceil 2(\upsilon+2)\ln n\rceil$. The following lemma gives similar bounds for 2-dimension.

Lemma 42. *Let P be a poset of cardinality n, maximum outdegree υ, and maximum indegree δ. Then we have the following bounds:*

$$\dim_2(P) \leq \lceil 2e(\upsilon+2)\ln n\rceil \tag{2.20}$$

$$\dim_2(P) \leq \Big\lceil e(\upsilon+2)\big(\ln n + \ln(\upsilon+2) + \ln(\delta+2) + 1\big)\Big\rceil. \tag{2.21}$$

Proof. Let P be a poset of cardinality n and maximum outdegree υ. We are going to construct an embedding from P into $\mathcal{Q}^d$ randomly, for d sufficiently large. For each $x \in P$, let A_x be an independent random subset of $[d]$, where each element is selected independently with probability $p = 1 - \frac{1}{\upsilon+2}$. We define a map $f : P \to \mathcal{Q}^d$, $f(x) = \bigcap_{y\geq x} A_y$. Our goal is to show that, if d is large enough, then with positive probability f is a poset embedding.

Note that f is monotone by construction. It is an embedding if and only if, for every pair $(x, y) \in P^2$ with $x \not\leq y$, we have $f(x) \not\subseteq f(y)$. For each such pair (x, y), let $E_{x,y}$ be the event that $f(x) \subseteq A_y$. Since $f(y) \subseteq A_y$, if none of the events $E_{x,y}$ occurs, then f is a poset embedding. For each $i \in [d]$, we have

$\mathsf{P}(i \in f(x), i \notin A_y) \geq p^{\upsilon+1}(1-p)$, so

$$\mathsf{P}\left(E_{x,y}\right) \leq (1-p^{\upsilon+1}(1-p))^d \leq \exp\big(-p^{\upsilon+1}(1-p)d\big) \leq \exp\big(-\tfrac{d}{e(\upsilon+2)}\big).$$

To prove (2.20), choose $d \geq 2e(\upsilon+2)\ln n$. The expected number of events $E_{x,y}$ that occur is at most $(n^2-n)n^{-2} < 1$, so with positive probability none of them occurs.

To prove (2.21), we use the following form of the Lovász local lemma:

Lemma 43 (Lovász Local Lemma, Theorem 1.5 in [51]). *Suppose* $0 < p < 1$ *and let* $A_1, A_2, \ldots, A_k$ *be events in a probability space such that* $\mathsf{P}(A_i) \leq p$. *Let* G *be a graph with vertex set* $[k]$ *such that, for all* $i \neq j \in [k]$, A_i *and* A_j *are independent unless* $ij \in E(G)$, *and suppose* G *has maximum degree* Δ. *If* $ep(\Delta+1) \leq 1$, *then* $\mathsf{P}\big(\bigcap_{i=1}^{k} \overline{A_i}\big) > 0$. □

The event $E_{x,y}$ is independent from $E_{z,w}$ if the sets $\{y\} \cup \{u : u \geq x\}$ and $\{w\} \cup \{u : u \geq z\}$ are disjoint. If they are not disjoint, then either $w = y$, or $z \leq y$, or $w \geq x$, or x and z have a common upper bound. For fixed x and y, the number of choices for (z, w) such that these sets intersect (not counting (x, y) itself) is therefore at most $n + (\delta+1)n + (\upsilon+1) + (\upsilon+1)(\delta+1) - 1 = (\upsilon+2)(\delta+2) - 1$.

Hence the total number of events $E_{z,w}$ dependent on $E_{x,y}$ is at most $(\upsilon+2)(\delta+2)n - 1$. If we choose $d \geq e(\upsilon+2)(\ln n + \ln(\upsilon+2) + \ln(\delta+2) + 1)$, then $e(\upsilon+2)(\delta+2)^2 n e^{-\frac{1}{e(\upsilon+2)}d} \leq 1$, and by the Lovász Local Lemma, the probability that none of the events $E_{x,y}$ occurs is positive. □

Using this result, we can bound the t-dimension of $\mathcal{D}_{(\alpha n, n]}$ for any fixed t and α. This poset has at least $(1-\alpha)n - 1$ elements, so its t-dimension is at least

$$\log_t\big((1-\alpha)n - 1\big) = \frac{\ln n}{\ln t} - O_{\alpha,t}(1). \tag{2.22}$$

We can apply Lemma 42 to obtain an upper bound. The maximum degree of $\mathcal{D}_{(\alpha n,n]}$ is at most $1/\alpha$, and its cardinality is at most $(1-\alpha)n$, so by Inequality (2.21)

$$\dim_t\left(\mathcal{D}_{(\alpha n,n]}\right) \leq \dim_2\left(\mathcal{D}_{(\alpha n,n]}\right) \leq (e+o_\alpha(1))\left(\tfrac{1}{\alpha}\right)\ln n. \tag{2.23}$$

This, together with the trivial lower bound (2.22), implies that

$$\dim_t\left(\mathcal{D}_{(\alpha n,n]}\right) = \Theta_{\alpha,t}(\ln n) \tag{2.24}$$

as $n \to \infty$.

2.6 Fractional Dimension

The fractional dimension of a poset is the linear programming relaxation of Dushnik-Miller dimension. A fractional realiser of a poset P is a function that assigns a nonnegative real weight to every linear extension of P so that, for every pair (x, y) of incomparable elements of P, the total weight of the extensions L for which $x <_L y$ is at least 1. The fractional dimension of P, or $\dim^\star(P)$, is the minimum total weight of a fractional realiser of P. Like the other variants, fractional dimension is monotonic and subadditive. Since a realiser is just a fractional realiser with $\{0,1\}$-valued weights, $\dim^\star(P) \leq \dim(P)$ for every poset P.

Brightwell and Scheinerman [11] proved that $\dim^\star(\mathcal{Q}^n) = n$.

Lemma 44 (Cf. Theorem 14 in [11])**.** *For all* $1 \leq k \leq n-1 \in \mathbb{N}$*,*

$$\dim^\star\left(\widetilde{\mathcal{M}}^n_{[0,k+1]}\right) \leq k+1.$$

Proof. Extend each permutation of $[n]$ to a linear extension of $\dim^\star\left(\widetilde{\mathcal{M}}^n_{[0,k+1]}\right)$ using colexicographic order. Assign each of these extensions weight $\frac{k+1}{n!}$. Consider two elements $X \not\geq Y$ of $\dim^\star\left(\widetilde{\mathcal{M}}^n_{[0,k+1]}\right)$. We need to show that there are at least

$\frac{n!}{k+1}$ extensions L such that $X <_L Y$. But this is easy, just take any $y \in Y$ whose multiplicity in Y is greater than its multiplicity in X. The probability that a randomly chosen permutation places y after all the elements of X is exactly $\frac{1}{|X|+1} \geq \frac{1}{k+1}$, and each such permutation induces an extension L such that $X <_L Y$, so the total weight of all such extensions is at least 1. □

Using this lemma, we can determine the asymptotic behaviour of $\dim^\star\left(\mathcal{D}_{[n]}\right)$ exactly.

Proof of Theorem 30. Embed $\mathcal{Q}^k \hookrightarrow \mathcal{D}_{[n]} \hookrightarrow \widetilde{\mathcal{M}}^{\pi(n)}_{[0,k+1]}$, where

$$k = \max\left\{\ell : \prod_{i=1}^{\ell} p_i \leq n\right\} \sim \frac{\ln n}{\ln \ln n}. \tag{2.25}$$

□

2.7 Other Divisibility Posets

In this section, we observe that our technique can be generalised to bound the dimension of the divisibility poset in other environments. Generally speaking, the ingredients that we need in order to make our proof work are unique factorisation, a norm to truncate the environment, and an asymptotic estimate of the number of irreducibles with norm less than a given number.

2.7.1 Arithmetic Semigroups

As it turns out, the most general environment for our results is that of an arithmetical semigroup, a notion that is central to abstract analytic number theory. A multiplicative arithmetical semigroup is a free commutative monoid G with identity element 1 and free generating set $\mathcal{P} \subset G$ (whose elements are called the *primes* of G) together with a norm function $\mathcal{N} : G \to \mathbb{R}$ satisfying the following axioms:

(a) $\mathcal{N}(1) = 1$, and $\mathcal{N}(p) > 1$ for all $p \in \mathcal{P}$,

(b) $\mathcal{N}(ab) = \mathcal{N}(a)\mathcal{N}(b)$ for all $a, b \in G$,

(c) For each $x \in \mathbb{R}$, the set $G(x) := \{a \in G : \mathcal{N}(a) \leq x\}$ is finite.

(c′) For each $x \in \mathbb{R}$, the set $\mathcal{P}(x) := \{p \in \mathcal{P} : \mathcal{N}(p) \leq x\}$ is finite.

Note that conditions (c) and (c′) are equivalent given the other axioms. We define $N_G(x) := |G(x)|$ and $\pi_G(x) := |\mathcal{P}(x)|$. The prototypical example of a multiplicative arithmetical semigroup is the set of natural numbers $\mathbb{N}$ under multiplication, but there are many more general examples. See the book by Knopfmacher [36] for a further reference on the theory of arithmetical semigroups. Before enumerating some examples that are of particular interest, we restrict ourselves to a class of arithmetical semigroups that satisfy the so called Axiom A. An arithmetical semigroup is said to satisfy Axiom A with exponent δ if there exist real constants $A > 0$ and $\nu \in [0, \delta)$ such that, as $x \to \infty$,

$$N_G(x) = Ax^\delta + O(x^\nu). \tag{2.26}$$

The following abstract prime number theorem holds for all arithmetical semigroups that satisfy this axiom:

$$\pi_G(x) = (1 + o(1))\frac{x^\delta}{\delta \ln x}, \tag{2.27}$$

as $x \to \infty$; see [36]. The natural numbers clearly satisfy Axiom A with exponent 1. More generally, given a number field K, the set of nonzero ideals of the ring of integers $\mathcal{O}_K$ of K is an arithmetical semigroup under multiplication. The norm of an ideal I is the cardinality of $\mathcal{O}_K/I$. We have unique factorisation of ideals since $\mathcal{O}_K$ is a Dedekind domain, and Landau [40] proved this semigroup satisfies Axiom A with exponent 1 by showing that

$$N_K(x) = C_K x + O(x^{1-\frac{2}{n+1}}), \tag{2.28}$$

as $x \to \infty$, where C_K is an explicit constant given by the Class Field Formula, and n is the degree of K. We have then the Landau prime ideal theorem [40]:

$$\pi_K(x) = (1 + o(1))\frac{x}{\ln x}. \tag{2.29}$$

Other examples have less resemblance with the naturals. For example,, we can take the set $\mathcal{G}$ of isomorphism classes of finite abelian groups. By the structure theorem for finitely generated abelian groups, every finite abelian group can be written unquely as a direct sum of cyclic groups of prime power order. Erdős and Szekeres [18] proved that $\mathcal{G}$ satisfies Axiom A with exponent 1 by showing that, as $x \to \infty$,

$$N_{\mathcal{G}}(x) = \Big(\prod_{k\geq 2}\zeta(k)\Big)x + O(x^{1/2}). \tag{2.30}$$

We now show that Theorems 28 and 29 hold for arithmetical semigroups that satisfy Axiom A.

Theorem 45. *Let G be an arithmetical semigroup that satisfies Axiom A with exponent δ and let $\mathcal{D}_{G(n)}$ be the divisibility poset of elements in G with norm less than or equal to n. Then the following bounds hold as $n \to \infty$:*

$$\big(\tfrac{\delta^2}{16} - o(1)\big)\frac{(\ln n)^2}{(\ln\ln n)^2} \leq \dim\big(\mathcal{D}_{G(n)}\big) \leq \big(4\delta^2 + o(1)\big)\frac{(\ln n)^2}{\ln\ln n}, \tag{2.31}$$

$$\big(\tfrac{\delta^2}{16} - o(1)\big)\frac{(\ln n)^2}{(\ln\ln n)^2} \leq \dim_2\big(\mathcal{D}_{G(n)}\big) \leq \big(\tfrac{4}{3}e\pi^2\delta^2 + o(1)\big)\frac{(\ln n)^2}{\ln\ln n}. \tag{2.32}$$

Proof. It follows from (c′) that $\mathcal{P}$ is countable. Label the primes of G $q_1, q_2, \ldots$ so that $\mathcal{N}(q_k)$ is monotone increasing. It follows from the abstract prime number theorem 2.27 that $\mathcal{N}(q_k) \sim (k \ln k)^{1/\delta} = k^{1/\delta + o(1)}$.

As in the proof of Theorem 28, if k is small enough that $\mathcal{N}(q_k)^{2\sqrt{k}} \leq n$, then we have the embedding $\mathcal{Q}^{\pi_G(k)}_{[0,2\sqrt{\pi_G(k)}]} \hookrightarrow \mathcal{D}_{G(n)}$.

It follows that $\dim\left(\mathcal{D}_{G(n)}\right) \geq k - \sqrt{k}$. If we take $c < \frac{1}{16}$ and $k = \left\lfloor c\left(\frac{\delta \ln n}{\ln\ln n}\right)^2\right\rfloor$, this condition is satisfied for all large n. If we let c approach $\frac{1}{16}$ we obtain the lower bounds in 2.31 and 2.32.

For the upper bound in 2.31, set $\varepsilon = \varepsilon(n)$. Let S be the set of all elements of $G(n)$ that factor into primes of norm less than $(\varepsilon \ln n)^{2/\delta}$ and let R be the set of all elements that factor into primes of norm greater than or equal to $(\varepsilon \ln n)^{2/\delta}$. Then $\dim(\mathcal{D}_S) \leq \pi_G((\varepsilon \ln n)^{2/\delta}) = (\frac{\varepsilon^2 \delta}{2} + o(1))\frac{(\ln n)^2}{\ln \ln n}$.

We further partition the large primes. Let $L = \left\lfloor \log_2\left(\delta \frac{\ln n}{\ln\ln n + \ln \varepsilon}\right)\right\rfloor$ and, for each $0 \leq i < L$, let $\theta_i = n^{2^{-i}}$, and R_i be the set of elements of $G(n)$ whose prime factors all have norms in the interval $\left(\theta_{i+1}, \theta_i\right]$. The analogue of Lemma 34 implies that $\dim(\mathcal{D}_R) \leq \sum\limits_{i=0}^{L-1} \dim(\mathcal{D}_{R_i})$. Just as before, $\mathcal{D}_{R_i} \hookrightarrow \mathcal{M}^{N_G(\theta_i)}_{[0,2^{i+1}-1]}$, so by Lemmas 31 and 33,

$$\dim(\mathcal{D}_{R_i}) \leq \dim\left(\mathcal{M}^{N_G(\theta_i)}_{[0,2^{i+1}-1]}\right) = N\left(N_G\left(n^{2^{-i}}\right), 2^{i+1}\right) \leq (4\delta + o(1)) \cdot 2^i \ln n, \quad (2.33)$$

and so

$$\dim(\mathcal{D}_R) \leq \sum_{i=0}^{L-1} \dim(\mathcal{D}_{R_i}) \leq (4\delta + o(1)) \cdot 2^L \ln n = (4\delta^2 + o(1))\frac{(\ln n)^2}{\ln\ln n}. \quad (2.34)$$

Finally, we combine these two inequalities to obtain

$$\dim\left(\mathcal{D}_{G(n)}\right) \leq \dim(\mathcal{D}_S) + \dim(\mathcal{D}_R) \leq \left(\tfrac{\varepsilon^2\delta}{2} + 4\delta^2 + o_\varepsilon(1)\right)\frac{(\ln n)^2}{\ln\ln n}, \quad (2.35)$$

and let $\varepsilon \to 0$ as slowly as necessary.

Now we prove the upper bound in 2.32. Fix $\varepsilon > 0$. Let S be the set of all elements of $G(n)$ whose prime factors all have norm at most $(\varepsilon \ln n)^{1/\delta}$ and R be the set of all elements whose prime factors are all greater than $(\varepsilon \ln n)^{1/\delta}$.

Let $r = \min\{\mathcal{N}(p) : p \in \mathcal{P}\}$, which exists by condition (c′) and is strictly greater than 1 by condition (a). The poset $\mathcal{D}_S$ can be embedded into the product of $\pi_G\left((\varepsilon \ln n)^{1/\delta}\right)$ chains, with at most $1 + \log_r n$ elements. Since the 2-dimension of each of these chains is at most $\log_r n$,

$$\dim_2(\mathcal{D}_S) \leq \pi_G\left((\varepsilon \ln n)^{1/\delta}\right) \log_r n = \Big(\frac{\varepsilon}{\ln r} + o(1)\Big)\frac{(\ln n)^2}{\ln \ln n}, \tag{2.36}$$

which follows from 2.27.

Now let $L = \left\lceil \log_2\left(\delta \frac{\ln n}{\ln\ln n + \ln \varepsilon}\right)\right\rceil$. For each i from 0 to $L-1$, let $\theta_i = n^{2^{-i}}$ and R_i be the set of elements of G_n whose prime factors all have norms that lie in the interval $\left(\theta_{i+1}, \theta_i\right]$. We have embeddings $\mathcal{D}_R \hookrightarrow \prod_{i=1}^{L-1} \mathcal{D}_{R_i}$ and $\mathcal{D}_{R_i} \hookrightarrow \mathcal{M}^{N_G(\theta_i)}_{[0,2^{i+1}-1]}$. By Lemma 40 and Axiom A (2.26),

$$\begin{aligned}\dim_2(\mathcal{D}_{R_i}) \leq \dim_2\left(\mathcal{M}^{N_G(\theta_i)}_{[0,2^{i+1}-1]}\right) \leq \Big(\frac{2e\pi^2}{3} + o(1)\Big)2^{2i}\ln\left(N_G(\theta_i)\right) = \\ \Big(\delta\frac{2e\pi^2}{3} + o(1)\Big)2^{2i}\ln\theta_i = \Big(\delta\frac{2e\pi^2}{3} + o(1)\Big)2^{i}\ln n.\end{aligned} \tag{2.37}$$

Therefore,

$$\begin{aligned}\dim_2(\mathcal{D}_R) &\leq \sum_{i=0}^{L-1}\dim_2(\mathcal{D}_{R_i}) \leq \delta\frac{2e\pi^2}{3}2^L\ln n + o(2^L \ln n) \\ &\leq \Big(\frac{4e\pi^2}{3}\delta^2 + o(1)\Big)\frac{(\ln n)^2}{\ln\ln n + \ln\varepsilon}.\end{aligned} \tag{2.38}$$

Finally, we have

$$\dim_2(\mathcal{D}_{G_n}) \leq \dim_2(\mathcal{D}_S) + \dim_2(\mathcal{D}_R) \leq \Big(\frac{\varepsilon}{\ln r} + \tfrac{4}{3}e\pi^2\delta^2 + o_\varepsilon(1)\Big)\frac{(\ln n)^2}{\ln\ln n} \tag{2.39}$$

for every fixed $\varepsilon > 0$. As before, let $\varepsilon = \varepsilon(n)$ approach 0 slowly to obtain 2.32.

□

Corollary 46. *For any number field K, upper and lower bounds of the same form as in Theorems 28 and 29 hold for the dimension and 2-dimension of the divisibility poset of ideals in K with norm at most n. The same is true of poset of isomorphism classes of abelian groups of order less than n, ordered by the direct summand relation.*

2.7.2 Polynomials over Finite Fields

Another interesting environment where we can prove an analogue of Theorem 28 is the divisibility poset of monic polynomials over $\mathbb{F}_q$, truncated by the degree. While equipping $F_q[t]$ with the norm $\mathcal{N}(p) = q^{\deg p}$ makes it into an arithmetic semigroup, it fails to satisfy Axiom A, since $N_{\mathbb{F}_q[t]}(x)$ is, up to a constant term, equal to $\frac{q}{q-1}$ times the smallest power of q less than or equal to x. This function is $\Theta(x)$, so, if it did satisfy Axiom A, it would do so with exponent 1. Then the error term would also be $\Theta(x)$, but Axiom A requires the error term to be $O(x^{\nu})$ for some $\nu < 1$.

In the theory of arithmetical semigroups, the notion of an *additive arithmetical semigroup* is more appropriate for handling semigroups with similar properties to $\mathbb{F}_q[t]$. A weaker version of Axiom A for such semigroups, called Axiom $A^\sharp$, is required, along with some other technical conditions, to obtain an abstract prime number theorem. See Knopfmacher and Zhang [35].

In order to avoid these additional technicalities, we will show that Theorem 28 generalizes to $\mathbb{F}_q[t]$, but essentially the same proof would work for the class of additive arithmetic semigroups that have an abstract prime number theorem.

Let $\pi_q(n)$ be the number of irreducible monic polynomials in $\mathbb{F}_q[x]$ of degree at most n and $\pi'_q(n)$ be the number of such irreducibles with degree exactly n. The following prime number theorem for polynomials is due to Gauss [25]. While the prime number theorem for integers is extremely difficult to prove without complex analysis, the one for polynomials has a simple proof that uses only elementary properties of finite fields.

Theorem 47. *For every prime power q and every $n \in \mathbb{N}$,*

$$\pi_q(n) \sim \frac{q}{q-1} \cdot \frac{q^n}{n} \tag{2.40}$$

as $n \to \infty$.

Proof. Every monic irreducible irreducible polynomial in $\mathbb{F}_q[X]$ of degree n is the minimal polynomial of exactly n distinct elements of $\mathbb{F}_{q^n}$. It follows that $\pi'_q(n) \leq \frac{q^n}{n}$.

Given an element $\alpha \in \mathbb{F}_{q^n}$, the minimal polynomial of α over $\mathbb{F}_q$ has degree n if and only if α is not contained in any proper subfields of $\mathbb{F}_{q^n}$. The union of all proper subfields of $\mathbb{F}_{q^n}$ is the same as the union of all maximal subfields, and each maximal subfield is of the form $\mathbb{F}_{q^{n/p}}$, where p is a prime factor of n. Since $n/p \leq n/2$ and n has at most $\log_2 n$ prime factors, $\pi'_q(n) \geq \frac{1}{n}(q^n - \log_2 n\sqrt{q^n})$.

Finally, to obtain 2.40, we sum over π'_q:

$$\pi_q(n) = \sum_{k=0}^{n} \pi'_q(k) = \sum_{k=0}^{n} \frac{1}{k}\left(q^k - O(\log_2 k\sqrt{q^k})\right) \sim \frac{q}{q-1} \cdot \frac{q^n}{n}. \tag{2.41}$$

$\square$

Now we can prove the analogues of Theorems 28 and 29 for polynomials.

Theorem 48. *Let $P_q(n)$ be the set of monic polynomials over $\mathbb{F}_q$ of degree n or less. As $n \to \infty$, we have the following bounds:*

$$\left(\frac{q-1}{16q^2} - o_q(1)\right)\frac{n^2}{(\log_q n)^2} \leq \dim\left(\mathcal{D}_{P_q(n)}\right) \leq \left(4\ln q + o_q(1)\right)\frac{n^2}{\log_q n} \tag{2.42}$$

$$\left(\frac{q-1}{16q^2} - o_q(1)\right) \frac{n^2}{(\log_q n)^2} \le \dim_2\left(\mathcal{D}_{P_q(n)}\right) \le \left(\frac{4\pi^2}{3}\ln q + o_q(1)\right)\frac{n^2}{\log_q n}. \tag{2.43}$$

Proof. Observe that, if k is small enough that $2k\sqrt{\pi_q(k)} \le n$, then there exists an embedding $\mathcal{Q}^{\pi_q(k)}_{[0,2\sqrt{\pi_q(k)}]} \hookrightarrow \mathcal{D}_{P_q(n)}$ whose image consists of all monic irreducibles of degree at most k and all products of at most $2\sqrt{\pi_q(k)}$ such polynomials. It follows that $\mathcal{D}_{P_q(n)}$ has dimension and 2-dimension at least $\pi_q(k) - \sqrt{\pi_q(k)}$. We can choose $k = \left\lfloor 2\log_q n - \log_q \log_q n - \log_q \frac{8q}{q-1} - o(1)\right\rfloor$ to obtain the lower bounds.

For the upper bound in 2.42, fix $\varepsilon > 0$ and let S be the set of all polynomials in $P_q(n)$ that factor into irreducibles with degree at most $2(\log_q n + \log_q \varepsilon)$ and R be the set of all polynomials that factor into irreducibles of degree greater than $2(\log_q n + \log_q \varepsilon)$. Clearly,

$$\dim(S) \le \pi_q(2(\log_q n + \log_q \varepsilon)) = \left(\frac{q}{2(q-1)}\varepsilon^2 + o_\varepsilon(1)\right)\frac{n^2}{\log_q n}. \tag{2.44}$$

Let $L = \left\lfloor \log_2\left(\frac{n}{\log_q n + \log_q \varepsilon}\right)\right\rfloor$, and, for each i from 0 to $L-1$, let R_i be the set of polynomials in $P_q(n)$ whose irreducible factors all have degrees in the interval $(2^{-(i+1)}n, 2^{-i}n]$.

For each i, $\mathcal{D}_{R_i} \hookrightarrow \mathcal{M}^{\pi_q(2^{-i}n)}_{[0,2^{i+1}-1]}$, so $\dim(R_i) \le 4\cdot 2^i \cdot n\ln q + 1$, and hence

$$\dim(\mathcal{D}_R) \le \sum_{i=0}^{L-1}\dim(\mathcal{D}_{R_i}) \le 4\cdot 2^L\cdot n\ln q + L \le (4\ln q + o_\varepsilon(1))\frac{n^2}{\log_q n}. \tag{2.45}$$

Finally, we add 2.44 and 2.45 to otain

$$\dim\left(\mathcal{D}_{P_q(n)}\right) \le \dim(\mathcal{D}_S) + \dim(\mathcal{D}_R) \le \left(4\ln q + \frac{q}{2(q-1)}\varepsilon^2 + o_\varepsilon(1)\right)\frac{n^2}{\log_q n}, \tag{2.46}$$

and we let $\varepsilon = \varepsilon(n) \to 0$ slowly.

For the upper bound in 2.43, we again fix an $\varepsilon > 0$. Let S be the set of all monic polynomials whose prime factors all have degree at most $\log_q n + \log_q \varepsilon$. Since $\mathcal{D}_S$ embeds into a product of $\pi_q(\log_q n + \log_q \varepsilon) \sim (\frac{q}{q-1}\varepsilon)\frac{n}{\log_q n}$ chains, each with at most $1 + n$ elements,

$$\dim_2(\mathcal{D}_S) \leq \left(\frac{q}{q-1}\varepsilon + o_\varepsilon(1)\right)\frac{n^2}{\log_q n}. \tag{2.47}$$

Now let $L = \left\lceil \log_2\left(\frac{n}{\log_q n + \log_q \varepsilon}\right)\right\rceil$. For each i from 0 to $L-1$, let R_i be the set of all monic polynomials whose irreducible factors all have degrees that lie in the interval $(2^{-(i+1)}n, 2^{-i}n]$. By Lemma 40,

$$\dim_2(\mathcal{D}_{R_i}) \leq \dim_2\left(\mathcal{M}_{[0,2^{i+1}-1]}^{\pi_q(2^{-i}n)}\right) \leq \left(\frac{2e\pi^2}{3}\ln q + o(1)\right)2^i \cdot n, \tag{2.48}$$

and so

$$\begin{aligned}\dim_2(\mathcal{D}_R) \leq \sum_{i=1}^{L-1}\dim_2(R_i) \leq \left(\frac{2e\pi^2}{3}\ln q + o_\varepsilon(1)\right)2^L \cdot n \leq \\ \left(\frac{4e\pi^2}{3}\ln q + o_\varepsilon(1)\right)\frac{n^2}{\log_q n}.\end{aligned} \tag{2.49}$$

As before, we add 2.47 and 2.49 to get

$$\dim_2\left(\mathcal{D}_{P_q(n)}\right) \leq \dim_2(S) + \dim_2(R) \leq \left(\frac{4e\pi^2}{3}\ln q + \frac{q}{q-1}\varepsilon + o_\varepsilon(1)\right)\frac{n^2}{\log_q n}, \tag{2.50}$$

and we let $\varepsilon \to 0$ as slowly as necessary. $\square$

To make the parallel with the bounds of Theorem 28, make the substitution $m = q^n$ in equations (2.42) and (2.43). We recover the same order of bounds we had for $\mathcal{D}_{[n]}$, up to constants depending on q.

2.8 Open problems

We pose several problems in this section, of which the central one is the following.

Question 2. *What is the correct asymptotic order of growth of* $\dim\left(\mathcal{D}_{[n]}\right)$*?*

We do not make any conjecture as to whether the lower bound or the upper bound in Theorem 28 is closer to the truth. On one hand, the lower bound is sharp in the sense that no $\mathcal{Q}^k_{1,\ell}$ of higher dimension can be embedded into $\mathcal{D}_{[n]}$, but on the other, the upper bound is more technically refined, where we bound each layer appropriately. In any case, we believe that determining the correct exponent on the $\ln\ln n$ factor requires new ideas. But we conjecture that, for $\mathcal{D}_{[n]}$, dimension and 2-dimension should behave similarly.

Conjecture 6. *As* $n \to \infty$*,* $\dim_2\left(\mathcal{D}_{[n]}\right) = \Theta\big(\dim\left(\mathcal{D}_{[n]}\right)\big)$.

So far, we have seen how the dimension behaves for some specific well structured sets of integers, like $[n]$ and $a[n]+b$. How does the dimension of a typical set behave?

Problem 3. *Let* $p = p(n)$ *and let* $A \subseteq [n]$ *be a random subset where each element is chosen independently with probability* p*. How does* $\dim(\mathcal{D}_A)$ *grow with* n*?*

Although we believe this question to be of great interest, we have made no serious attempt to answer it. It would be interesting to see how other poset properties vary with p. Finally, we have shown in Section 2.5 that the dimension of $\mathcal{D}_{(\alpha n,n]}$ is bounded for all n and gave some bounds depending only on α. It would be nice to improve the bounds obtained. Also recall that we have shown that $\dim_t\left(\mathcal{D}_{(\alpha n,n]}\right) = \Theta_{\alpha,t}\big(\ln n\big)$. This suggests the following conjecture.

Conjecture 7. *For each* $0 < \alpha < 1$ *and* $t \geq 2$*, there exists a constant* $c = c(\alpha,t)$ *such that* $\dim_t\left(\mathcal{D}_{(\alpha n,n]}\right) \sim c\ln n$ *as* $n \to \infty$*.*

Our best upper bound for $c(\alpha, t)$ depends only on α, and our best lower bound depends only on t. We believe that the correct value should depend on both α and t.

CHAPTER 3

LOCAL DIMENSION AND COMPLETE BIPARTITE COVERINGS OF DIFFERENCE GRAPHS

The results in this chapter are joint work with António Girão. Theorem 51 was proved independently by Felsner and Ueckerdt [21], and also by Damásdi, Keszegh, and Nagy [15]. The three manuscripts were combined into [14], which has been submitted to the *SIAM Journal on Discrete Mathematics*.

3.1 Introduction and definitions

Difference graphs were first studied by Hammer, Peled, and Sun in [53]. They gave several equivalent characterisations of difference graphs, of which we take the following as our definition. A *difference graph* is a bipartite graph G with parts A and B such that A and B are linearly ordered sets, the neighbourhood of every element of A is an initial segment of B and similarly the neighbourhood of every element of B is an initial segment of A.

Note that if G is a difference graph with parts A and B and $x \leq y$ in A, then $\Gamma(y) \subseteq \Gamma(x)$. This condition is actually equivalent to our definition. Other definitions and characterisations of difference graphs, including an explanation of the name "difference graph", can be found in [53].

The problem of covering the edges of a graph using difference graphs is closely related to the problem of estimating the local dimension of a poset, as we shall see later. In the remaining of this chapter, we will only consider difference graphs without isolated vertices. We need to introduce now a few more definitions.

Let G be a difference graph and v a vertex of G, the *complement* of v, denoted v^c, is the highest-ranked neighbour of v.

For each $n \in \mathbb{N}$, let H_n be the bipartite graph with vertex set $A \dot\cup B$, where

$A = B = [n]$, and edge set $\{ab : a \in A, b \in B, a + b \leq n + 1\}$. Observe that H_n is indeed a difference graph where every vertex equals its double complement. Actually, every difference graph with this property is isomorphic to H_n, for some n. Indeed, suppose D is a difference graph with parts A and B with the property that every vertex is its own double complement, or equivalently that every vertex has a unique neighbourhood. Since there are at most $|B|$ possible neighbourhoods of vertices in A, $|B| \geq |A|$, and for the same reason $|A| \geq |B|$, so $|A| = |B| = n$ for some n. Now it is clear that assigning each vertex v the number $n + 1 - d(v)$ defines an isomorphism from D to H_n.

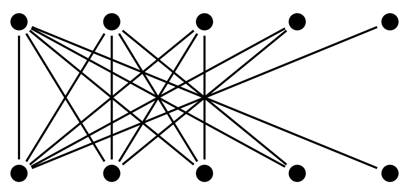

Figure 3.1: The difference graph H_5.

Let $\mathcal{H}$ be a class of graphs, which we assume is closed under isomorphism. An *$\mathcal{H}$-cover* of a graph G is a set of subgraphs of G in $\mathcal{H}$ whose union contains all the edges of G. For any $\mathcal{H}$-cover $\mathcal{C}$ of G and and $v \in V(G)$, the *$\mathcal{C}$-multiplicity* of v, denoted $\mu_{\mathcal{C}}(v)$, is the number of subgraphs in $\mathcal{C}$ that contain v. The *local $\mathcal{H}$-covering number* of G is defined as

$$\min \Big\{ \max \{\mu_{\mathcal{C}}(v) : v \in V(G)\} : \mathcal{C} \text{ is an } \mathcal{H}\text{-cover of } G \Big\}. \tag{3.1}$$

For any graph G, we denote by $\operatorname{lbc}(G)$ the local complete bipartite covering number of G and by $\operatorname{ldc}(G)$ the local difference graph covering number of G. The following proposition appearing in [32] explains the link between the local difference covering number and the local dimension of a poset.

Proposition 49. *Let P be a poset of height 2 and let G be the bipartite graph whose vertices are the elements of P and whose edges are the pairs of incomparable* *elements between the two layers. Then* $\mathrm{ldc}\,(G) \leq \mathrm{ldim}\,(P) \leq \mathrm{ldc}\,(G) + 2$. $\square$

Given a poset P, the *split* of P is a height-two poset Q with elements $\{x', x'' : x \in P\}$, where $x' < y''$ in Q if and only if $x \leq y$ in P. Barrera-Cruz, Prag, Smith, Taylor, and Trotter proved in [2] that the inequality $\mathrm{ldim}\,(Q) - 2 \leq \mathrm{ldim}\,(P) \leq 2\,\mathrm{ldim}\,(Q) - 1$ always holds. This observation together with Proposition 49 reduces the problem of determining the local dimension of a poset (up to a factor of 2) to the problem of finding the local difference graph covering number of a certain bipartite graph.

Notice that, for all $m < n$, the subgraph of H_n induced by $([m], [n] \setminus [n-m])$ is isomorphic to H_m. This fact implies that $\mathrm{lbc}\,(H_n)$ is a monotone increasing function of n. The following proposition reduces the problem of determining $\mathrm{lbc}\,(G)$ for an arbitrary difference graph G to the problem of determining $\mathrm{lbc}\,(H_n)$.

Proposition 50. *Let G be a difference graph with parts A and B. Then,* $\mathrm{lbc}\,(G) = \mathrm{lbc}\,(H_n)$, *for some* $n \leq \min\{|A|, |B|\}$.

Proof. Let $\mathcal{B}$ be a complete bipartite cover of G. If no two vertices have identical neighbourhoods we are done; otherwise, suppose u and v have the same neighbourhood. We may suppose that u and v belong to the same set of graphs in $\mathcal{B}$. If not, and $\mu_{\mathcal{B}}(u) \leq \mu_{\mathcal{B}}(v)$, we may remove v from every bipartite graph and add it to every graph containing u. Since u and v have the same neighbourhood, we still have a cover, and we have not increased the multiplicity of any vertex. Hence identifying u with v does not change the local bipartite covering number of H. If we iteratively identify pairs of vertices with the same neighbourhood, we eventually obtain a difference graph for which no two vertices have the same neighbourhood, which must be H_n for some $n \leq \min\{|A|, |B|\}$. $\square$

In [32], Kim, Martin, Masařík, Shull, Smith, Uzzell, and Wang proved that $\mathrm{lbc}\,(H_n) \leq \log n + O(1)$ and asked for the exact value of $\mathrm{lbc}\,(H_n)$, whenever $n + 1$ is a power of two. We answer this question for all values of n with the following theorem.

Theorem 51. *For all $n \in \mathbb{N}$,*

$$\operatorname{lbc}(H_n) = \min\left\{i \in \mathbb{N} : \binom{2i}{i} \geq n+1\right\} = \frac{1}{2}\log n + \frac{1}{4}\log\log n + O(1).$$

In Section 3.3, we look at the poset $\mathcal{Q}^n_{1,2}$ induced by the first two layers of the hypercube $\mathcal{Q}^n$. Spencer [50] and Füredi, Hajnal, Rödl, and Trotter [56] collectively proved that $\mathcal{Q}^n_{1,2}$ has dimension $(1+o(1))\log\log n$. We show that $\mathcal{Q}^n_{1,2}$ also has local dimension $(1+o(1))\log\log n$. We also show that the local complete bipartite covering number of the graph of non-related pairs between the first and second layer is $\Theta(\log n)$.

3.2 Proof of Theorem 51

Our proof of Theorem 51 is based on the fact that every complete bipartite cover of H_n can be simplified in a way that does not increase the maximum multiplicity.

Suppose $\mathcal{B}$ is a complete bipartite cover of H_n. We say that $\mathcal{B}$ *splits* if there is a partition of $\mathcal{B}$ into three sets $\mathcal{B}_1$, $\mathcal{B}_2$, and $\mathcal{B}_3$ and a partition of $V(H_n) = A \cup B$ into four sets $A = A_1 \cup A_2$ and $B = B_1 \cup B_2$ with the following properties:

- A_1 and B_1 are initial segments of A and B respectively;
- $|A_1| + |B_1| = n+1$;
- $\mathcal{B}_1$ is a complete bipartite cover of the difference graph $H_{|A_1|-1} = (A_1 \setminus \{\max(A_1)\}, B_2)$;
- $\mathcal{B}_2$ is a complete bipartite cover of the difference graph $H_{|B_1|-1} = (A_2, B_1 \setminus \{\max(B_1)\})$;

- $\mathcal{B}_3$ consists of the complete bipartite graph (A_1, B_1).

For example, in Figure 3.2, the black edges span a copy of $K_{4,3}$, the red edges span a copy of H_3, and the blue edges span a copy of H_2. A complete bipartite cover of H_6 that splits might consist of a complete bipartite cover of the blue edges, a complete bipartite cover of the red edges, and the complete bipartite graph spanned by the black edges.

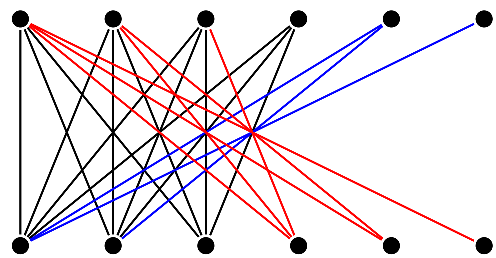

Figure 3.2: H_6 as a union of $K_{4,3}$, H_3, and H_2.

To prove Theorem 51, we need the following lemma:

Lemma 52. *For every complete bipartite cover $\mathcal{B}$ of H_n, there exists a complete bipartite cover $\mathcal{C}$ such that $\mathcal{C}$ splits and, for every $i \in V(H_n)$, $\mu_{\mathcal{C}}(i) \leq \mu_{\mathcal{B}}(i)$.*

We postpone the proof of Lemma 52 and first show that it implies Theorem 51.

Proof of Theorem 51, given Lemma 52. For any two nonnegative integers a and b, define an (a, b) covering of H_n to be a complete bipartite covering $\mathcal{B}$ such that every vertex in A has $\mathcal{B}$-multiplicity at most a and every vertex in B has $\mathcal{B}$-multiplicity at most b. Let $t_{a,b}$ to be the smallest n such that there does not exist an (a, b) covering of H_n. Then $t_{i,i}$ is the smallest n such that $\operatorname{lbc}(H_n) = i + 1$.

We want to show that $t_{a+1,b+1} = t_{a,b+1} + t_{a+1,b}$ for all a and b. First we prove the lower bound. Let $n = t_{a,b+1} + t_{a+1,b} - 1$, $A_1 = [t_{a,b+1}] \subseteq A$, $B_1 = [t_{a+1,b}] \subseteq B$, $A_2 = A \setminus A_1$, and $B_2 = B \setminus B_1$. Take an $(a, b + 1)$ covering $\mathcal{B}_1$ of $H_{t_{a,b+1}-1} = ([t_{a,b+1} - 1], B_2)$ and an $(a + 1, b)$ covering $\mathcal{B}_2$ of

$H_{t_{a+1,b}-1} = (A_2, [t_{a+1,b} - 1])$. Then $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \{(A_1, B_1)\}$ is an $(a+1, b+1)$ covering of H_n, so $t_{a+1,b+1} \geq n+1$.

Now we prove the upper bound. Let $n = t_{a,b+1} + t_{a+1,b}$ and consider a complete bipartite covering $\mathcal{B}$ of H_n. By Lemma 52, we may assume $\mathcal{B}$ splits. Either $|A_1| \geq t_{a,b+1} + 1$ or $|B_1| \geq t_{a+1,b} + 1$. If the former, then $\mathcal{B}_1$ cannot be an $(a, b+1)$ covering, so either B_2 contains a vertex of $\mathcal{B}_1$-multiplicity at least $b+2$ or A_1 contains a vertex of $\mathcal{B}_1$-multiplicity at least $a+1$, and adding (A_1, B_2) increases its $\mathcal{B}$-multiplicity to at least $a+2$. If the latter, then $\mathcal{B}_2$ cannot be an $(a+1, b)$ covering, so either A_2 contains a vertex of $\mathcal{B}_2$-multiplicity at least $a+2$ or B_1 contains a vertex of $\mathcal{B}_2$-multiplicity at least $b+1$ and hence $\mathcal{B}$-multiplicity at least $b+2$. In either case, either A contains a vertex of $\mathcal{B}$-multiplicity at least $a+2$ or B contains a vertex of $\mathcal{B}$-multiplicity at least $b+2$, so $\mathcal{B}$ is not an $(a+1, b+1)$ covering.

Given this recurrence, we can prove that $t_{a,b} = \binom{a+b}{a}$ by double induction on a and b. First observe that $t_{a,0} = t_{0,b} = 1$, so the equation holds when either $a = 0$ or $b = 0$. Also notice that the lexicographic ordering on pairs (a, b) is a well-ordering, as it is the ordering induced by the map $(a, b) \mapsto \omega \cdot a + b$ onto the set of ordinals below ω^2. If (a, b) is the lexicographically minimal pair such that $t_{a,b} \neq \binom{a+b}{a}$, then $a \geq 1$ and $b \geq 1$, and so $t_{a-1,b} = \binom{a+b-1}{a-1}$ and $t_{a,b-1} = \binom{a+b-1}{a}$. But then $t_{a,b} = t_{a-1,b} + t_{a,b-1} = \binom{a+b-1}{a-1} + \binom{a+b-1}{a} = \binom{a+b}{a}$, a contradiction. Therefore the equation holds for all pairs. In particular, $t_{i,i} = \binom{2i}{i}$, from which the theorem immediately follows.

To prove the asymptotic formula for $\operatorname{lbc}(H_n)$, suppose i is the smallest natural number such that $\binom{2i}{i} \geq n+1$. Using Stirling's approximation, we have

$$\binom{2i}{i} = \frac{(2i)!}{i!^2} = (1+o(1))\frac{\sqrt{4\pi i} \cdot (2i)^{2i} e^{-2i}}{2\pi i \cdot i^{2i} e^{-2i}} = (1+o(1))\frac{4^i}{\sqrt{\pi i}}. \tag{3.2}$$

If we write

$$i = \frac{1}{2}\log n + \frac{1}{4}\log\log n + c_n, \tag{3.3}$$

then we have

$$\binom{2i}{i} = (1+o(1))\frac{n\sqrt{\log n}\cdot 4^{c_n}}{\sqrt{(\pi/2+o(1))\log n}} = (1+o(1))n\frac{4^{c_n}}{\sqrt{\pi/2}}. \tag{3.4}$$

Since $\binom{2(i-1)}{i-1} \leq n < \binom{2i}{i}$, we must have $4^{c_n-1} \leq \sqrt{\frac{\pi}{2}} + o(1) \leq 4^{c_n}$, so c_n is bounded.

Using more precise estimates, such as those given by Robbins in [46], one can show that $\mathrm{lbc}\,(H_n)$ is equal to either $\left\lfloor \frac{1}{2}\log n + \frac{1}{4}\log\log n + \frac{1}{4}\log 2\pi \right\rfloor$ or $\left\lceil \frac{1}{2}\log n + \frac{1}{4}\log\log n + \frac{1}{4}\log 2\pi \right\rceil$ for all $n \geq 2$. The calculations are simple but tedious, so we omit them here.

$\square$

Now we shall prove Lemma 52.

Proof of Lemma 52. Let $\mathcal{B}$ be a complete bipartite cover of $H_n = (A, B)$. Recall that, for each vertex $v \in A$, the complementary vertex of v is $v^c = n + 1 - v \in B$. Similarly, if $v \in B$, $v^c = n + 1 - v \in A$. For any $(C, D) \in \mathcal{B}$, if v is the last element of C, then the last element of D is at most v^c. Since the subgraph of H_n induced by $([v], [v^c])$ is complete bipartite, we may replace all the graphs in $\mathcal{B}$ whose last element in A is v with the graph induced by the union of their vertex sets without increasing the multiplicity of any vertex, so we can assume that each edge of the form vv^c is covered exactly once and that every graph in $\mathcal{B}$ contains such an edge. We also assume throughout the proof that $\mathcal{B}$ is minimal in the sense that, if we remove any vertex from any of the subgraphs in $\mathcal{B}$, the result is no longer a cover. At any time, if $\mathcal{B}$ is not minimal, we replace it with a minimal refinement.

For each $i \in A$, let $B_i = (C_i, D_i)$ be the unique subgraph in $\mathcal{B}$ that covers ii^c.

We also define

$$a_i = \inf\{a > i : i \in C_a\} \in A \cup \{\infty\} \tag{3.5}$$

for each $i \in A$, with the convention that $\inf(\emptyset) = \infty$. Now we apply the following algorithm to turn all the C_v's and D_v's into intervals, without increasing the multiplicity of any vertex.

We proceed iteratively for each $i \in A$. Suppose $\mathcal{B}$ has the following properties:

1. For all $j < i$, D_j is an interval, and

2. for all $a \geq a_j$, $j \in C_a$ if and only if $a_j \in C_a$.

Note that both of these properties hold vacuously when $i = 1$. Let t be the left endpoint of the longest interval in D_i containing i^c. If $t = 1$, then D_1 is already an interval. In this case, we can remove i from every C_a with $a > i$ without losing any edges, so $a_i = \infty$. Otherwise, if $t \geq 2$, observe that $a_i = (t-1)^c$. This holds because the edge $\{i, t-1\}$ must be covered by some B_j with $i < j \leq (t-1)^c$; however, if $j < (t-1)^c$, then $j^c \in D_i$, and we could have removed it and still had a cover. By property 2, for all $k \in C_i$ and $a > i \geq a_k$, C_a contains i if and only if it contains k. This implies that $C_i \subset C_{(t-1)^c}$, so we can move all the elements of $D_i \setminus [t, i^c]$ from D_i to $D_{(t-1)^c}$ without losing any edges. Now, for every $a \geq a_i$, a^c is connected to both i and a_i, and $a^c \notin D_i$. If i is contained in more C_a's with $a \geq a_i$ than a_i, then we can remove i from each such C_a and add it to every C_a that contains a_i without losing any edges, and vice versa if a_i is contained in at least as many such C_a's than i.

Once we have done this for every $i \in A$, $\mathcal{B}$ has the properties that D_i is an interval for all i and that, for all $a \geq a_i$, $i \in C_a$ if and only if $a_i \in C_a$. We now apply the same algorithm with the roles of A and B reversed to turn all the C_i's into intervals, without changing the fact that every D_i is an interval.

As before, define $b_i = \inf\{b > i : i \in D_{b^c}\} \in B \cup \{\infty\}$ for each $i \in B$. Suppose $\mathcal{B}$ has the following properties:

1. For all $j < i \in B$, C_{j^c} is an interval, and

2. for all $b \geq b_j$, $j \in D_{b^c}$ if and only if $b_j \in D_{b^c}$.

Let s be the left endpoint of the longest interval in C_{i^c} containing i^c. If $s = 1$, C_{i^c} is already an interval. In this case, we can remove $[i]$ from every D_{b^c} with $b > i$ without losing any edges. Since we're removing an initial segment from each D_{b^c}, it remains an interval. If $s \geq 2$, then, by the same argument as before, $b_i = (s-1)^c$. By property 2, for all $k \in D_{i^c}$ and $b > i \geq b_k$, D_{b^c} contains i if and only if it contains k. This implies that $D_{i^c} \subset D_{s-1}$, so we can move all the elements of $C_{i^c} \setminus [s, i^c]$ from C_{i^c} to C_{s-1} without losing any edges. Now, for every $b \geq b_i$, b^c is connected to both i and b_i. For every $b \geq b_i$, if $i \in D_{b^c}$, then, because D_{b^c} is an interval, $b_i \in D_{b^c}$. If D_{b^c} contains b_i but not i, we can remove the initial segment $[b_i]$ from D_{b^c}. Because i is connected to b^c, there is a $b' > b$ such that $D_{b'^c}$ contains i and therefore the whole interval $[i, b]$, so we still have a cover.

We have thus proved that we may pass to a bipartite cover $\mathcal{B}$ in which both parts of every subgraph are intervals. Now we show that any such cover that is minimal must split. Let $(A_1, B_1) \in \mathcal{B}$ be any bipartite graph in $\mathcal{B}$ such that A_1 and B_1 are both initial segments of $[n]$, or equivalently any graph containing the edge $\{1, 1\}$. Since $\mathcal{B}$ is a cover, it must contain such a graph. By definition, $(A_1, B_1) = (C_i, D_i)$ for some i, so $|A_1| + |B_1| = n + 1$. Let $A_2 = [n] \setminus A_1$ and $B_2 = [n] \setminus B_1$. Now, for every $(C_j, D_j) \in \mathcal{B}$ with $j \neq i$, either $j < i$ or $j > i$. If the former, then $C_j \subseteq A_1 \setminus \{\max(A_1)\}$, and we can replace D_j with $D_j \cap B_2$. If the latter, then $D_j \subseteq B_1 \setminus \{\max(B_1)\}$ and we can replace C_j with $C_j \cap A_2$. Either way, all the removed edges are already in (A_1, B_1), so we still have a cover. If we do this for every j then it is clear that the resulting cover splits. $\square$

Corollary 53. *For every difference graph D with parts A and B,*

$$\mathrm{lbc}\,(D) \le \tfrac{1}{2}\log|A| + \tfrac{1}{4}\log\log|A| + 2, \tag{3.6}$$

and this is the best bound possible (up to a small constant additive term). For any graph G on n vertices,

$$\mathrm{ldc}\,(G) \le \mathrm{lbc}\,(G) \le \left\lfloor \tfrac{1}{2}\log n + \tfrac{1}{4}\log\log n + \tfrac{3}{2} \right\rfloor \mathrm{ldc}\,(G)\,. \tag{3.7}$$

Proof. The first statement follows from Theorem 51 and Proposition 50. The second statement follows from the first and the fact that every difference subgraph of a graph on n vertices has a part with at most $\frac{n}{2}$ vertices. □

3.3 The first two layers of the hypercube

Here we look at the suborder $\mathcal{Q}^n_{1,2}$ of the cube $\mathcal{Q}^n$ induced by the two layers $[n]^{(1)}$ and $[n]^{(2)}$ and the bipartite graph G_n whose vertex set is $\mathcal{Q}^n_{1,2}$ and whose edges are non-related pairs between the two layers. For simplicity's sake, we identify $[n]^{(1)}$ with $[n]$ in the obvious way.

Theorem 54. *As $n \to \infty$, $\mathrm{lbc}\,(G_n) = \Theta(\log n)$.*

To prove Theorem 54, we use the following proposition proved by Hansel in [27] (see also Bollobás and Scott [8]).

Proposition 55. *For all n, $\mathrm{lbc}\,(K_n) \ge \log n$.* □

In fact, Hansel proved a slightly stronger result, namely that for any complete bipartite covering $\mathcal{B}$ of K_n, the average $\mathcal{B}$-multiplicity of a vertex in K_n is at least $\log n$.

Proof of Theorem 54. First we shall prove the upper bound. Take a random partition $A\dot{\cup}B$ of $[n]$ where each element is chosen to be in A independently with probability $\frac{1}{3}$ and consider the complete bipartite graph $\left(A, B^{(2)}\right)$. The edge $\{a, bc\}$ is covered if and only if $a \in A$, $b, c \in B$, so the probability that any given edge is not covered is $\frac{23}{27}$. If we take at least $3 \log_{27/23} n$ such partitions independently, then the expected number of edges not covered is $3\binom{n}{3}\left(\frac{23}{27}\right)^{3\log_{27/23} n} < n^3 \cdot \left(\frac{23}{27}\right)^{3\log_{27/23} n} = 1$. This implies that $\operatorname{lbc}(G_n) \leq \lceil 3\log_{27/23} n \rceil = \left\lceil \frac{3}{\log 27/23} \log n \right\rceil$.

Now, to prove the lower bound, pick any element $x \in [n]$ and consider the subgraph F of G induced by the set $\{xy : y \in [n] \setminus x\} \cup [n] \setminus x$. It's easy to check that the homomorphism $\varphi : F \to K_{n-1}$ defined by $\varphi(y) = \varphi(xy) = y$ is a double covering map. Hence if $\mathcal{B}$ is a complete bipartite cover of G and $\mathcal{B}'$ is its restriction to F, then $\varphi(\mathcal{B}')$ is a complete bipartite cover of K_{n-1}, so by Proposition 55 there exists a vertex $z \in K_{n-1}$ with $\varphi(\mathcal{B}')$-multiplicity at least $\log(n-1)$. But $\mu_{\varphi(\mathcal{B}')}(z) \leq \mu_{\mathcal{B}'}(z) + \mu_{\mathcal{B}'}(xz)$, so one of z and xz has $\mathcal{B}'$-multiplicity (and therefore $\mathcal{B}$-multiplicity) at least $\frac{1}{2}\log(n-1)$. □

Next, we look at the dimension and local dimension of $\mathcal{Q}^n_{1,2}$. Spencer [50] proved that

$$\dim\left(\mathcal{Q}^n_{1,2}\right) \leq \log\log n + \frac{1}{2}\log\log\log n + \log(\sqrt{2}\pi) + o(1). \tag{3.8}$$

Füredi, Hajnal, Rödl, and Trotter [56] proved the corresponding lower bound, showing that

$$\dim\left(\mathcal{Q}^n_{1,2}\right) = \log\log n + \left(\frac{1}{2} + o(1)\right)\log\log\log n. \tag{3.9}$$

The next theorem implies that $\operatorname{ldim}\left(\mathcal{Q}^n_{1,2}\right) \sim \dim\left(\mathcal{Q}^n_{1,2}\right)$ as $n \to \infty$.

Theorem 56. *As $n \to \infty$, $\operatorname{ldim}\left(\mathcal{Q}^n_{1,2}\right) = \log\log n + O(\log\log\log n)$.*

Proof. The upper bound follows from 3.8. We shall prove the lower bound

$$\begin{gathered}\operatorname{ldim}\left(\mathcal{Q}^n_{1,2}\right) \geq \operatorname{ldc}(G_n) \geq \\ \log\log n - \left(1 + \tfrac{1}{\ln 2}\right)\log\log\log n - o(1).\end{gathered} \tag{3.10}$$

Let $\mathcal{D}$ be a difference graph covering of $G_n = G_{\mathcal{Q}^n_{1,2}}$. Recall that, for each $D \in \mathcal{D}$, the singletons in D are weakly ordered by reverse inclusion of their neighbourhoods. We define a sequence of difference graphs $D_i \in \mathcal{D}$ and a sequence of subsets $L_i \subseteq [n]$ as follows. Let $c < 1$ be a fixed positive real number. First, choose $D_1 \in \mathcal{D}$ such that D_1 contains at least n^c singeletons, if there is such a graph in $\mathcal{D}$. If there isn't, then each pair is contained in at least $\frac{n-2}{n^c} \gg \log\log n$ elements of $\mathcal{D}$.

Otherwise, let L_1 be the set of singletons in D_1. Now suppose L_i and $D_1, D_2, \ldots, D_i$ have already been chosen. We choose a graph $D_{i+1} \in \mathcal{D}$ such that $V(D_{i+1}) \cap L_i \geq |L_i|^c$, if such a graph exists. If so, then, by the Erdős-Szekeres theorem, there is a subset $L_{i+1} \subseteq V(D_{i+1}) \cap L_i$ such that $|L_{i+1}| \geq |L_i|^{c/2}$ and the elements of L_{i+1} appear in the same or opposite order in D_i and D_{i+1}.

Continue in this way until either $|L_i| \leq \log n$ or $|L_i| > \log n$ and there is no graph in $\mathcal{D}$ that contains $|L_i|^c$ elements of L_i. In the former case, each element of L_i appears in at least i elements of $\mathcal{D}$, and $n^{(c/2)^i} \leq \log n$, so

$$i \geq \frac{1}{1 - \log c}(\log\log n - \log\log\log n). \tag{3.11}$$

In the latter case, let a and b be the first and last elements of L_i in the order induced by D_i and look at the set of chosen difference graphs D_j that contain the pair ab. Because the ordering on L_i induced by D_j begins with either a or b for every j, none of these graphs can contain any edges from ab to L_i. Every other difference graph in $\mathcal{D}$ contains less than $|L_i|^c$ edges from L_i to ab, so there must be at least $\frac{|L_i|-2}{|L_i|^c} \geq (\log n)^{1-c} - 2(\log n)^{-c}$ such difference graphs containing ab. Now, if we take $c = 1 - \frac{\log\log\log n}{\log\log n}$, then

$$\begin{aligned}(\log n)^{1-c} - 2(\log n)^{-c} &= \\ (\log n)^{\frac{\log\log\log n}{\log\log n}} - 2(\log n)^{-1+o(1)} &= \\ 2^{\cancel{\log\log n}\cdot\frac{\log\log\log n}{\cancel{\log\log n}}} - o(1) &= \\ \log\log n - o(1).\end{aligned} \tag{3.12}$$

Using the affine approximation

$$\tfrac{1}{1-\log c} = 1 + \tfrac{1}{\ln 2}(c-1) + O\left((c-1)^2\right) \tag{3.13}$$

as $c \to 1$, we have

$$\begin{aligned}\tfrac{1}{1-\log c}(\log\log n - \log\log\log n) &= \\ \left(1 - \tfrac{1}{\ln 2}\cdot\tfrac{\log\log\log n}{\log\log n} + O\left((\tfrac{\log\log\log n}{\log\log n})^2\right)\right)(\log\log n - \log\log\log n) &= \\ \log\log n - \left(1 + \tfrac{1}{\ln 2}\right)\log\log\log n - o(1).\end{aligned} \tag{3.14}$$

Therefore,

$$\operatorname{ldc}(G_n) \geq \min\left\{\log\log n - o(1), \log\log n - \left(1 + \tfrac{1}{\ln 2}\right)\log\log\log n - o(1)\right\}, \tag{3.15}$$

and the stated lower bound follows immediately. □

3.4 Open problems

Question 4. *What is the exact value of* $\operatorname{ldim}\left(\mathcal{Q}^n_{1,2}\right)$*? In particular, what is the coefficient of the* $\log\log\log n$ *term?*

We can sample a random poset of cardinality n and dimension (at most) d by taking n points independently and uniformly at random from the unit hypercube $[0,1]^d$.

Question 5. *Let $P_{d,n}$ be a random poset of cardinality n and dimension (at most) d. What can we say about* $\operatorname{ldim}(P_{d,n})$ *when n is a function of d?*

It's obvious that when n is large and d is fixed, then with high probability $P_{d,n}$ contains all small posets of dimension d. As we will see in Chapter 4, when n is sufficiently large, at least one of them will have dimension d.

Question 6. *What is the slowest-increasing function $f(d)$ such that, if $n \gg f(d)$, then with high probability* $\operatorname{ldim}(P_{d,n}) = d$*?*

CHAPTER 4

THE LOCAL DIMENSION OF SUBORDERS OF THE HYPERCUBE

In this chapter, we study the local dimension of the hypercube and its two-layer suborders.

Kim, Martin, Masařík, Shull, Smith, Uzzell, and Wang [32] proved that $\operatorname{ldim}(\mathcal{Q}^n)$ is at least $\Omega\left(\frac{n}{\log n}\right)$, but so far the only upper bound we have for $\operatorname{ldim}(\mathcal{Q}^n)$ is the trivial bound of n.

The main result in this chapter is that the local dimension of the first and middle layers of the Boolean lattice is asymptotically $\frac{n}{\log n}$.

Theorem 57. *As $n \to \infty$, $\operatorname{ldim}\left(\mathcal{Q}^n_{1,\lfloor n/2 \rfloor}\right) = \frac{n}{\log n} + O\left(\frac{n \log\log n}{(\log n)^2}\right)$.*

4.1 Lexicographic sums

Let P be a poset with ground set X and, for each $x \in X$, let Q_x be a poset with ground set Y_x. The *lexicographic sum* of $\{Q_x\}$ over P, denoted $\sum_{x \in P} Q_x$, is a poset on the ground set $\{(x, y) : x \in X, y \in Y_x\}$ where $(x, y) \leq (z, w)$ if and only if either $x < z$ or $x = z$ and $y \leq w$. Hiraguchi [29] proved that

$$\dim\left(\sum_{x \in X} Q_x\right) = \max\left\{\dim(P), \max\{\dim(Q_x) : x \in X\}\right\}. \tag{4.1}$$

We don't have such a simple equation for local dimension, but we can prove some weaker inequalities.

Proposition 58. *For any poset P with ground set X and any family $\{Q_x\}_{x\in X}$ of posets indexed by X, we have the following inequalities:*

$$\operatorname{ldim}\left(\sum_{x\in X} Q_x\right) \geq \max\left\{\operatorname{ldim}(P), \max\{\operatorname{ldim}(Q_x) : x \in X\}\right\}, \tag{4.2}$$

$$\operatorname{ldim}\left(\sum_{x\in X} Q_x\right) \leq \max\left\{\operatorname{ldim}(P), \max\{\dim(Q_x) : x \in X\}\right\}, \tag{4.3}$$

$$\operatorname{ldim}\left(\sum_{x\in X} Q_x\right) \leq \operatorname{ldim}(P) + \max\{\operatorname{ldim}(Q_x) : x \in X\}. \tag{4.4}$$

Proof. Inequality 4.2 follows by monotonicity from the fact that $\sum_{x\in P} Q_x$ has suborders isomorphic to P and to Q_x for each $x \in X$.

To prove inequality 4.3, let $\mathcal{L}$ be a local realiser of P and, for each $x \in X$, let $\mathcal{M}_x$ be a realiser of Q_x. For convenience, we regard partial linear orders as lists rather than posets. We will construct a local realiser of $\sum_{x\in P} Q_x$ as follows. For each $x \in X$, if $|\mathcal{M}_x| \leq \mu_{\mathcal{L}}(x)$, replace each occurrence of x in the lists in $\mathcal{L}$ with $x \times M$ for some $M \in \mathcal{M}_x$, using each list in $\mathcal{M}_x$ at least once. If $\mu_{\mathcal{L}}(x) < |\mathcal{M}_x|$, replace each occurrence of x in the lists in $\mathcal{L}$ with $x \times M$ for some $M \in \mathcal{M}_x$, using a different list M for each occurrence. Then, for each unused $M \in \mathcal{M}_x$, add $x \times M$ as a new list. Let $\mathcal{N}$ be the set of all lists thus constructed. For each $x \in X$ and each $y \in Y_x$, we can see that $\mu_{\mathcal{N}}(x, y) = \max\{\mu_{\mathcal{L}}(x), |\mathcal{M}_x|\}$ by counting the number of occurrences of (x, y) in these lists. To show that $\mathcal{N}$ is a local realiser, suppose $(x, y) \ngeq (x', y')$. Then either $x = x'$ and $y \ngeq y'$ or $x \ngeq x'$. If the former, then there is some $M \in \mathcal{M}_x$ such that y occurs before y' in M, and $x \times M$ is a sublist of some list in $\mathcal{N}$. If the latter, then there is some $L \in \mathcal{L}$ in which x occurs before x'. In the corresponding element of $\mathcal{N}$, x and x' have been replaced by sublists containing (x, y) and (x', y') respectively, so (x, y) occurs before (x', y') in this list.

To prove inequality 4.4, let $\mathcal{L}$ be a local realiser of P and, for each $x \in X$, let $\mathcal{M}_x$ be a local realiser of Q_x. For each $x \in X$, let K_x be an arbitrary linear extension of Q_x. Let $\mathcal{N}$ be the set of all lists obtained by replacing, for each $x \in X$, each occurrence of x in the linear orders in $\mathcal{L}$ with $x \times K_x$ as well as all lists of the form $x \times M$ with $x \in X$ and $M \in \mathcal{M}_x$. This set is a local realiser of $\sum_{x \in X} Q_x$, and, for every $x \in X$ and $y \in Y_x$, $\mu_{\mathcal{N}}(x, y) = \mu_{\mathcal{L}}(x) + \mu_{\mathcal{M}_x}(y)$. The special case where P is an antichain was stated as an exercise by Bosek, Grytczuk, and Trotter in [9]. $\square$

4.2 Lower bounds

Wang Zhiyu [59], following a comment by Christophe Crespelle in [32], suggested that an information entropy method might help improve the bounds on $\operatorname{ldim}(\mathcal{Q}^n)$.

Let X be a discrete random variable taking values in $\{x_1, x_2, \ldots, x_n\}$ with $\mathbb{P}\{X = x_i\} = p_i$ for each i. Then the entropy of X, denoted $H(X)$, is defined by the formula

$$H(X) = -\sum_{\substack{1 \leq i \leq n \\ p_i \neq 0}} p_i \log p_i. \tag{4.5}$$

It can be shown using the strict concavity of log that $H(X) \leq \log n$, and that this bound is only attained when X is uniformly distributed. Entropy was introduced by Shannon [48], and $H(X)$ can be thought of as the average amount of information in bits obtained be observing the value of X. Shannon's fundamental theorem for a noiseless channel, also known as the source coding theorem, makes this precise. Before stating the theorem, we need a few definitions. An *alphabet* is a finite set with two or more elements, and the elements of an alphabet are called symbols. Given an alphabet Ω, a *word* over Ω is a finite sequence of symbols in Ω, and $\Omega^\star$ denotes the set of all words over Ω. Given a finite set Σ and an alphabet Ω, a *prefix-free code* is a map C from Σ to $\Omega^\star$ such that, for all $x, y \in \Sigma$, if $x \neq y$, then $C(x)$ is not an initial segment of $C(y)$.

Theorem 59. *Let X be a random variable taking values in a finite set Σ and let Ω be a finite alphabet. For every prefix-free code C, the expected length of $C(X)$ is at* *least* $\frac{H(X)}{\log|\Omega|}$*. Conversely, there exists a prefix-free code* C *such that the expected length of* $C(X)$ *is at most* $\frac{H(X)}{\log|\Omega|}+1$. $\square$

In a note added to the end of [32], Crespelle suggested a method of encoding an arbitrary poset on a fixed ground set, which we call the Crespelle code. Given a ground set X with cardinality n and a poset P on X, we encode P as a word over a $3n$-symbol alphabet Ω as follows. Let $\Omega = \{x_i, x_m, x_f : x \in X\}$. Given a local realiser $\mathcal{L}$ of P, we write each nontrivial[1] partial extension in $\mathcal{L}$ as a list of elements of X, using symbols of the form x_i at the beginning of each list and symbols of the form x_m everywhere else. Then we concatenate these lists in any order and replace the last symbol x_m with the corresponding x_f. We call the resulting word a *Crespelle codeword* for P. Note that this code is prefix-free by construction. If $\operatorname{ldim}(P) = d$, then P has a local realiser $\mathcal{L}$ for which each element of X has multiplicity at most d, and the Crespelle codeword for P constructed from $\mathcal{L}$ has length at most dn. Since each symbol in Ω is equivalent to $\log 3n$ bits, we use a total of at most $dn \log 3n$ bits. For example, here is a Crespelle codeword for the standard example S_3 (with spaces added for readability):

$$a_i b_m c_m x_m y_m z_m \ c_i b_m a_m z_m y_m x_m \ x_i a_m \ y_i b_m \ z_i c_f.$$

Kleitman and Rothschild [34] proved that the entropy of a uniformly random partial order[2] on $[n]$ is $(\frac{1}{4}+o(1))n^2$. Kim, Martin, Masařík, Shull, Smith, Uzzell, and Wang [32] proved that the maximum local dimension of a poset on n points is at between $\left(\frac{1}{4e\ln 2}-o(1)\right)\frac{n}{\log n}$ and $(4+o(1))\frac{n}{\log n}$ $n \to \infty$, so every partial order on $[n]$ has a Crespelle codeword with at most $(4+o(1))n^2$ bits. As Crespelle observed, this means that the Crespelle code is optimal up to a constant factor.

[1] Of course, if we remove all the one-element partial extensions from a local realiser, the resulting set is still a local realiser.

[2] See Kozieł and Sulkowska [39] for a method of sampling random labelled partial order on n points which converges in total variation to the uniform distribution.

We can use this fact to improve Kim et al.'s lower bound on the maximum local dimension of an n-element poset.

Theorem 60. *As $n \to \infty$, the expected local dimension of a poset chosen uniformly at random from the set of all n-element labelled posets is at least*

$$\left(\frac{1}{4} - o(1)\right)\frac{n}{\log n}.$$

Proof. Suppose $n \geq 2$ and let P be a random partial order on $[n]$, where we assign equal probability to each partial order. Assume for convenience that n is even. The total number of such partial orders is at least the number of two-layer partial orders with minimum elements $1, 2, 3, \ldots, n/2$ and maximal elements $n/2+1, n/2+2, \ldots, n$, which is equal to $2^{\frac{1}{4}n^2}$. It follows that $H(P) \geq \frac{1}{4}n^2$. It is not much harder to show that this is also true when n is odd.

Let $d = \mathbb{E}[\mathrm{ldim}\,(P)]$. Then the expected length of the shortest Crespelle codeword for P is at most dn. Hence by Theorem 59 $H(P) \leq dn \log 3n$, so

$$d \geq \frac{n}{4 \log 3n}. \tag{4.6}$$

$\square$

A similar argument shows that the expected 2-dimension of a random partial order on $[n]$ is at least $\frac{1}{4}n - o(1)$ as $n \to \infty$. First, we define a prefix-free binary code for the set of all partial orders on $[n]$ as follows. Suppose P is a partial order on $[n]$ with 2-dimension $d \geq 1$. Fix a poset embedding f from P into $\mathcal{Q}^d$. The codeword for P consists of a block of $\lceil \log n \rceil$ bits representing $d-1$ as a binary number, followed by n blocks of d bits each, where the i^{th} block is the representation of $f(i)$ as a binary string. The length of this word is $\lceil \log n \rceil + dn$ bits. For example, here is a binary codeword for the standard example S_4 (as before, with spaces added for readability):

$$011\ 0001\ 0010\ 0100\ 1000\ 1110\ 1101\ 1011\ 0111.$$

Now let P be a uniformly random partial order on $[n]$, as defined in Theorem 60. Repeating the proof of Theorem 60 using this binary code instead of the Crespelle code, we find that

$$\mathbb{E}[\dim_2(P)] \geq \frac{1}{4}n - \frac{\lceil \log n \rceil}{n} \tag{4.7}$$

for all $n \geq 2$. By essentially the same argument, we can prove that

$$\mathbb{E}[\dim_t(P)] \geq \frac{1}{4\log t}n - \frac{\lceil \log_t n \rceil}{n} \tag{4.8}$$

for every $t \geq 2$ and $n \geq 2$.

It's clear from the proofs that the same results hold (up to an additive $o(1)$ term) for a uniformly random two-level poset with minimum elements $1, 2, \ldots, \lfloor n/2 \rfloor$ and maximum elements $\lfloor n/2 \rfloor + 1, \lfloor n/2 \rfloor + 2, \ldots, n$.

The following lower bound applies to any pair of layers in the Boolean lattice.

Theorem 61. *For any $\ell, k < n$,*

$$\operatorname{ldim}\left(\mathcal{Q}^n_{\ell,k}\right) \geq \frac{\log\binom{n}{k}}{\log\binom{n}{\ell}} - \frac{\log\binom{n}{k}}{\left(\log\binom{n}{\ell}\right)^2}\left(\log\log\binom{n}{\ell} + c\right), \tag{4.9}$$

where $c \leq \log 12 < 3.585$. In particular,

$$\operatorname{ldim}\left(\mathcal{Q}^n_{1,k}\right) \geq \left(\frac{1}{\log n} - O\left(\frac{\log\log n}{(\log n)^2}\right)\right)\log\binom{n}{k}. \tag{4.10}$$

Proof. Assume $\ell < k$ and let P be a random two-level poset defined as follows. Let $A = [n]^{(\ell)}$ and $B = [m]$. For each $b \in B$, choose a random subset X_b of $[n]$ of cardinality k, and then let P be the two-level poset on $A \dot\cup B$ where each $b \in B$ is above $a \in A$ if $a \subseteq X_b$. Because the entropy of the joint distribution of mutually independent random variables is the sum of the entropies of those random variables, $H(P)$ is at least $m \log \binom{n}{k}$. Now define $d = \mathbb{E}[\mathrm{ldim}\,(P)]$. The expected length of the shortest Crespelle codeword for P is at most $d\left(\binom{n}{\ell} + m\right)$. Hence by Theorem 59 $H(P) \leq d\left(\binom{n}{\ell} + m\right)\left(\log\left(3\binom{n}{\ell} + 3m\right)\right)$, so

$$d \geq \frac{m \log \binom{n}{k}}{\left(\binom{n}{\ell} + m\right)\left(\log\left(3\binom{n}{\ell} + 3m\right)\right)}. \tag{4.11}$$

If we set $m = \left\lfloor \binom{n}{\ell}\left(\log \binom{n}{\ell} - 1\right)\right\rfloor$, then

$$d \geq \frac{\log \binom{n}{k}}{\log \binom{n}{\ell}} - \frac{\log \binom{n}{k}}{\left(\log \binom{n}{\ell}\right)^2}\left(\log\log \binom{n}{\ell} + \log 6 + \binom{n}{\ell}^{-1}\right). \tag{4.12}$$

Now assume $\mathrm{ldim}\,(P) \geq d$ (which occurs with nonzero probability) and modify P as follows to obtain a new poset P'. For each subset S of A, if there is more than one vertex in B whose neighbourhood is S, delete all but one of them. Since P is a lexicographic sum of antichains over P' and P' is not a chain, by Proposition 58, P and P' have the same local dimension. Since P' embeds into $\mathcal{Q}^n_{\ell,k}$, $\mathrm{ldim}\left(\mathcal{Q}^n_{\ell,k}\right) \geq \mathrm{ldim}\,(P') \geq d$. A similar argument works when $k < \ell$. □

This theorem has the following immediate corollary.

Corollary 62. *For any $\alpha \in [0,1]$, as $n \to \infty$,*

$$\mathrm{ldim}\left(\mathcal{Q}^n_{1,\lfloor n^\alpha \rfloor}\right) \geq (1-\alpha)n^\alpha - O\left(\frac{n^\alpha \log\log n}{(\log n)^2}\right) \tag{4.13}$$

, and the same is true for $\mathcal{Q}^n_{1,(n-\lfloor n^\alpha \rfloor)}$. In general, whenever $0 \leq \alpha < \beta \leq 1$,

$$ldim \mathcal{Q}^n_{\lfloor n^\alpha \rfloor, \lfloor n^\beta \rfloor} \geq \frac{1-\beta}{1-\alpha} \cdot n^{\beta-\alpha} - O\left(\frac{n^{\beta-2\alpha \log\log n}}{(\log n)^2}\right). \tag{4.14}$$

□

We also have the following lower bound for the first and middle layers. The bound $\operatorname{ldim}(\mathcal{Q}^n) \geq (1+o(1))\frac{n}{\log n}$, which improves one of Kim et al.'s results by a constant factor, was also proved by Stefan Felsner by different means. Felsner's proof can be found in [14].

Corollary 63. *As $n \to \infty$, $\operatorname{ldim}(\mathcal{Q}^n) \geq \operatorname{ldim}\left(\mathcal{Q}^n_{1,\lfloor n/2 \rfloor}\right) \geq \frac{n}{\log n} - O\left(\frac{n \log\log n}{(\log n)^2}\right)$.* □

4.2.1 Suborders of the divisibility lattice

Recall that, in Chapter 2, we studied the dimension of suborders of the divisibility lattice $\mathcal{D}_{\mathbb{N}} = (\mathbb{N}, |)$ and proved that

$$\operatorname{ldim}(([n], |)) \leq \dim\left(\mathcal{D}_{[n]}\right) \leq (4\ln 2 + o(1))\frac{(\log n)^2}{\log\log n}, \tag{4.15}$$

where $\mathcal{D}_{[n]} = ([n], |)$.

We can prove the following lower bound using Corollary 62.

Proposition 64. *As $n \to \infty$,*

$$\operatorname{ldim}\left(\mathcal{D}_{[n]}\right) \geq \left(\frac{1}{4} - o(1)\right)\frac{\log n}{\log\log n}.$$

Proof. Fix $c < 1$ and let $k = \left\lfloor \frac{c^2}{4}\left(\frac{\log n}{\log\log n}\right)^2 \right\rfloor$. Because

$$p_k^{\sqrt{k}} = (\log n)^{(c+o(1))\frac{\log n}{\log\log n}} = n^{c+o(1)} \ll n, \tag{4.16}$$

we can define an embedding from $\mathcal{Q}^k_{1,\lfloor\sqrt{k}\rfloor}$ into $\mathcal{D}_{[n]}$, which therefore has local dimension at least $(\frac{c}{4} - o(1))\frac{\log n}{\log\log n}$. Now taking $c \to 1$ as slowly as necessary completes the proof of the lower bound. □

4.2.2 Posets whose dimension and local dimension are equal

Recall the following lower bound for the local dimension of $\mathcal{Q}^n_{1,2}$ from Chapter 3.

Theorem 65. *As $n \to \infty$, $\operatorname{ldim}\left(\mathcal{Q}^n_{1,2}\right) \geq \log\log n - O(\log\log\log n)$.* $\square$

Spencer [50] proved that $\dim\left(\mathcal{Q}^n_{1,2}\right) \leq \log\log n + O(\log\log\log n)$, so this bound is asymptotically the best possible. Because $\mathcal{Q}^{n-k+1}_{1,2} \hookrightarrow \mathcal{Q}^n_{k,k+1}$ and $\mathcal{Q}^n_{\ell,k} \hookrightarrow \mathcal{Q}^{n+1}_{\ell,k+1}$, we also have

$$\operatorname{ldim}\left(\mathcal{Q}^n_{\ell,k}\right) \geq (1 - o_{\ell,k}(1)) \log\log n \tag{4.17}$$

as $n \to \infty$, for every fixed $\ell < k \in \mathbb{N}$.

The following theorem and corollary show non-constructively that there exist posets of arbitrarily large dimension whose dimension and local dimension are equal. This result was proved independently by Barrera-Cruz, Prag, Smith, Taylor, and Trotter [2].

Theorem 66. *For all $n \in \mathbb{N}$, $\operatorname{ldim}(\mathbb{R}^n) = n$.*

Proof. It follows from Theorem 65 and Spencer's theorem [50] that $\operatorname{ldim}\left(\mathcal{Q}^m_{1,2}\right) \sim \dim\left(\mathcal{Q}^m_{1,2}\right) \sim \log\log m$, so there exists an infinite sequence of finite posets with dimension n and local dimension $n - o(n)$. Because every finite n-dimensional poset embeds into $\mathbb{R}^n$, $\operatorname{ldim}(\mathbb{R}^n) = n - o(n)$ by monotonicity.

Now suppose that $\operatorname{ldim}\left(\mathbb{R}^k\right) \leq k - 1$ for some $k \in \mathbb{N}$. Then, for all n, $\mathbb{R}^n$ embeds into $(\mathbb{R}^k)^{\lceil n/k \rceil}$, so $\operatorname{ldim}(\mathbb{R}^n) \leq \frac{k-1}{k} n + k - 1$ by subadditivity. But this contradicts the previous claim that $\operatorname{ldim}(\mathbb{R}^n) = n - o(n)$. Therefore $\operatorname{ldim}(\mathbb{R}^n) = n$ for all n. $\square$

Corollary 67. *For every $d \in \mathbb{N}$, there exists a* finite *poset P such that $\dim(P) = \operatorname{ldim}(P) = d$.*

Proof. Let Q be an infinite poset. We wish to prove that, if every finite suborder of Q has local dimension at most d, then Q has local dimension at most d.

Let $\mathcal{K}$ be a first-order language with a constant symbol c for each $c \in Q$, ternary relation symbol $\langle \_, \_, \_ \rangle$, binary relation symbol $\_ \epsilon \_$, and unary predicate symbol $P\_$. A $\mathcal{K}$-structure is thought of as having two types of elements, called points and lists.[3] For terms r, s, and t, the intended meaning of Pr is that r is a point, the intended meaning of $r \in s$ is that s is a list containing the point r, and the intended meaning of $\langle r, s, t\rangle$ is that s is a list in which r is listed before or in the same place as t.

Let T be the the set of all the following $\mathcal{K}$-sentences:

i) $\forall \ell \; \forall x \; x \in \ell \implies Px \wedge \neg P\ell$ (i.e., if r is in s, then r is a point and s is a list)

ii) $\forall \ell \; \forall x \; \forall y \; \langle x, \ell, y\rangle \implies x \in \ell \wedge y \in \ell$ (i.e., if s is a list in which r comes before t, then s contains both r and t),

iii) $\forall \ell \; \forall x \; \forall y \; \langle x, \ell, y\rangle \wedge \langle y, \ell, x\rangle \implies x = y$ (i.e., if two points appear in the same place in a list, then they must be the same),

iv) $\forall \ell \; \forall x \; x \in \ell \implies \langle x, \ell, x\rangle$ (i.e., the order on each list is reflexive),

v) $\forall \ell \; \forall x \; \forall y \; \forall z \; \langle x, \ell, y\rangle \wedge \langle y, \ell, z\rangle \implies \langle x, \ell, z\rangle$ (i.e., the order on each list is transitive),

vi) $\forall \ell \; \forall x \; \forall y \; x \in \ell \wedge y \in \ell \implies \langle x, \ell, y\rangle \vee \langle y, \ell, x\rangle$ (i.e., the order on each list is linear),

vii) $\forall x \; \exists \ell_1 \ldots \exists \ell_d \; \forall \ell \; x \in \ell \implies \bigvee_{i=1}^{d} \ell = \ell_i$ (i.e., each point is contained in at most d lists),

viii) for each $c, c' \in Q$ with $c \leq c'$, $\forall \ell \; c \in \ell \wedge c' \in \ell \implies \langle c, \ell, c'\rangle$, (i.e., the order on each list, restricted to the elements of Q, is a partial linear extension of Q),

ix) for each $c, c' \in Q$ with $c \not\geq c'$, $\exists \ell \; \langle c, \ell, c'\rangle$, (i.e., the set of all lists, restricted to elements of Q, is a local realiser of Q),

x) for each $c, c' \in Q$ with $c \neq c'$, $\neg c = c'$.

[3] It would be more natural to do this with a two-sorted language, but unsorted first-order logic is more familiar.

Suppose every finite suborder of Q has local dimension at most d. Let S be a finite subset of T and let $C(S)$ be the set of all constant symbols that appear in the sentences of S. Since $C(S)$ is a finite subset of Q, it induces a suborder of Q that has local dimension at most d. Let $\mathcal{L}$ be a local realiser of $\big(C(S), \leq_{C(S)}\big)$ such that $\mu_{\mathcal{L}}(c) \leq d$ for each $c \in C(S)$. Now we shall define a $\mathcal{K}$-structure $\mathcal{N}$ with domain $C(S)\dot{\cup}\mathcal{L}$. We interpret the non-logical symbols as follows:

- $P^{\mathcal{N}} = C(S)$,
- $\_ \epsilon \_^{\mathcal{N}} = \big\{(c, L) : L \in \mathcal{L}, c \in |L|\big\}$,
- $\langle \_, \_, \_ \rangle^{\mathcal{N}} = \big\{(c, L, c') : L \in \mathcal{L}, c, c' \in |L|, c \leq_L c'\big\}$,
- for each constant symbol $c \in C(S)$, $c^{\mathcal{N}} = c$, and
- for each $c \notin C(S)$, choose an arbitrary interpretation $c^{\mathcal{N}} \in C(S)$.

Now $\mathcal{N}$ satisfies axioms i-vii as well as all axioms viii, ix, and x with c and $c' \in C(S)$, so $\mathcal{N} \vDash S$.

Since every finite subset of T is satisfiable, by compactness T is satisfiable. Suppose $\mathcal{M}$ is a $\mathcal{K}$-structure such that $\mathcal{M} \vDash T$. For each $\ell \in |\mathcal{M}|$ such that $\mathcal{M} \vDash \neg P\ell$, let $L_\ell = \{c \in Q : \mathcal{M} \vDash c \,\epsilon\, \ell\}$. Now define a binary relation $\leq_\ell$ on L_ℓ by $c \leq c' \iff \mathcal{M} \vDash \langle x, \ell, y\rangle$. The axioms of T imply that each $(L_\ell, \leq_\ell)$ is a partial linear extension of Q and that $\big\{(L_\ell, \leq_\ell) : \mathcal{M} \vDash \neg P\ell\big\}$ is a local realiser of Q for which each element of Q has multiplicity at most d.

Now suppose that every finite d-dimensional poset has local dimension at most $d-1$. It would follow that $\mathbb{R}^d$ has local dimension at most $d-1$, which would contradict Theorem 66.

$\square$

4.3 Upper bounds for two layers

In this section, we prove some upper bounds on the local dimension of posets of the form $\mathcal{Q}^n_{\ell,k}$. The following proposition gives good upper bounds on the local dimension of suborders of the form $\mathcal{Q}^n_{\ell,(n-k)}$ when ℓ and k are both constant (or slowly increasing functions of n).

Proposition 68. *Whenever $\ell < n-k$, $\operatorname{ldim}\left(\mathcal{Q}^n_{\ell,(n-k)}\right) \leq 2 + \max\{\ell, k\}$.*

Proof. Let π_0 be the linear extension consisting of ℓ-sets in any order, followed by the $(n-k)$-sets in any order. Then let π_1 be the ℓ-element sets in the opposite order as in π_0 followed by the $(n-k)$-element sets in the opposite order as in π_0. Then, for each $i \in [n]$, let L_i be the $(n-k)$-element sets not containing i (in any order) followed by the ℓ-element sets containing i. It's obvious that $\{\pi_0, \pi_1, L_1, L_2, \ldots, L_n\}$ is a local realiser of $\mathcal{Q}^n_{\ell,(n-k)}$ in which every ℓ-set has multiplicity $2+\ell$ and every $(n-k)$-set has multiplicity $2+k$. $\square$

This generalises Ueckerdt's [58] result that $\operatorname{ldim}(S_n) = 3$ for $n \geq 3$. By contrast, Füredi [23] proved that, for any constant $k \geq 3$, $\dim\left(\mathcal{Q}^n_{k,(n-k)}\right) = n-2$ for all sufficiently large n.

Corollary 69. *For any fixed $\alpha \in (0,1)$, $\operatorname{ldim}\left(\mathcal{Q}^n_{1,(n-\lfloor n^\alpha \rfloor)}\right) = \Theta(n^\alpha)$, with constants between $1-\alpha$ and 1.*

For pairs of layers that are close together, we have the following upper bound due to Brightwell, Kierstead, Kostochka, and Trotter [10].

Theorem 70. *For any $s, k, n \in \mathbb{N}$ with $s+k \leq n$,*

$$\operatorname{ldim}\left(\mathcal{Q}^n_{s,(s+k)}\right) \leq \dim\left(\mathcal{Q}^n_{s,(s+k)}\right) \leq (4k^2 + 18k)\lceil \ln n \rceil. \tag{4.18}$$

In the case $k = 1$,

$$\operatorname{ldim}\left(\mathcal{Q}^n_{s,(s+1)}\right) \leq \dim\left(\mathcal{Q}^n_{s,(s+1)}\right) \leq 6\lceil \log_3 n \rceil. \tag{4.19}$$

□

Kostochka [38] later improved the second bound.

Theorem 71. *For all $s \leq n$*

$$\operatorname{ldim}\left(\mathcal{Q}^n_{s,s+1}\right) \leq \dim\left(\mathcal{Q}^n_{s,s+1}\right) \leq 2\min\left\{k : 2 \cdot k! \geq n\right\} = (2 + o(1))\frac{\log n}{\log\log n}.$$

. □

Dushnik [16] showed that $\dim\left(\mathcal{Q}^n_{1,k}\right) \geq n - \sqrt{n}$ whenever $k \geq 2\sqrt{n}$. The next theorem shows that local dimension behaves very differently.

Theorem 72. *For any $n \in \mathbb{N}$ and $k \leq n$,*

$$\operatorname{ldim}\left(\mathcal{Q}^n_{1,k}\right) \leq \frac{n}{\log n} + \frac{2n\log\log n}{(\log n)^2} + 3. \tag{4.20}$$

More generally, whenever $1 \leq \ell < k \leq n$ and $\ell < \frac{n}{\log n}$,

$$\operatorname{ldim}\left(\mathcal{Q}^n_{\ell,k}\right) \leq (1 + o_\ell(1))\frac{n}{\log n}. \tag{4.21}$$

Proof. We first describe a general construction, then show two different ways the construction can be realised. Let G be an n-edge bipartite graph with parts A and B and suppose each vertex in A has degree at most Δ. We identify $\mathcal{Q}^n_{\ell,k}$ with the set of all k-subsets and ℓ-subsets of $E(G)$ and define a local realiser of $\mathcal{Q}^n_{\ell,k}$ as follows. For each $v \in A$ and each $X \subseteq \Gamma(v)$, let $L_{v,X}$ be a partial linear extension that lists all k-sets S such that $\{u \in B : vu \in S\} = X$ followed by all ℓ-sets containing an edge vu with $u \notin X$. Now let π_0 list all the ℓ-sets of edges in some order followed by all the k-sets, and let π_1 list all the ℓ-sets in the opposite order followed by all the k-sets in the opposite order. Then $\{\pi_0, \pi_1\} \cup \{L_{v,X} : v \in [A], X \subset [B]\}$ is a local realiser of $\mathcal{Q}^n_{\ell,k}$ in which every ℓ-set has multiplicity at most $2^{\Delta-1}\ell + 2$ and every k-set has multiplicity at most $|A| + 2$.

For any n, ℓ, and k with $1 \leq \ell < k \leq n$ and $\ell < \frac{n}{\log n}$, we may take G to be any n-edge bipartite graph with parts A and B with $|A| = \left\lceil \frac{n}{\log n - \log\log n - \log \ell} \right\rceil$ and $|B| = \lceil \log n - \log\log n - \log \ell \rceil$. Here we have $2^{\Delta-1} + 2\ell \leq 2^{|B|-1}\ell + 2 < \frac{n}{\log n} + 2$ and $|A| + 2 \leq \frac{n}{\log n} + \frac{2n}{(\log n)^2}(\log\log n + \log \ell) + 3$, which gives us an upper bound of

$$\operatorname{ldim}\left(\mathcal{Q}^n_{\ell,k}\right) \leq \frac{n}{\log n} + \frac{2n}{(\log n)^2}(\log\log n + \log \ell) + 3. \tag{4.22}$$

We can do slightly better when $n = 2^{m-1}m\ell$ for some $m \in N$. In this case, we may take G to be the disjoint union of ℓ copies of the m-dimensional hypercube graph. For this graph, $\Delta = m$ and $|A| = 2^{m-1}\ell$, so we have

$$\operatorname{ldim}\left(\mathcal{Q}^n_{\ell,k}\right) \leq 2^{m-1}\ell + 2 = \frac{n}{m} + 2. \tag{4.23}$$

Since $m = \log \frac{n}{\ell} - \log m + 1$, this bound is asymptotically the same as 4.22.

□

Theorem 57 follows immediately from Theorem 61 and Theorem 72.

4.4 Open problems

It's known (see, for example, Dushnik [16]) that $\dim\left(\mathcal{Q}^n_{1,k}\right)$ is monotone in k for $0 \leq k \leq n-1$, but $\operatorname{ldim}\left(\mathcal{Q}^n_{1,k}\right)$ is not, as $\operatorname{ldim}\left(\mathcal{Q}^n_{1,n/c}\right) \geq \Omega_c(n/\log n)$ while $\operatorname{ldim}\left(\mathcal{Q}^n_{1,n-k}\right) \leq k + 2$. We do not know whether or not it's unimodal. Our best upper and lower bounds are unimodal, with a single peak at $k \approx n/2$ – although the upper bound is more like a plateau.

Question 7. *How does the function* $f_n(k) := \operatorname{ldim}\left(\mathcal{Q}^n_{1,k}\right)$ *behave? Is it unimodal?*

A poset has a short Crespelle codeword if and only if it has a local realiser with small average multiplicity. It follows that a poset with small local dimension has a short Crespelle codeword. Is the converse true? In other words, how far apart can the smallest possible average multiplicity be from the smallest possible maximum multiplicity of a local realiser? Of course, this question is trivial, since adding a top element to a nontrivial poset does not change its local dimension. By repeatedly adding new top elements to a d-local-dimensional poset, we obtain d-local-dimensional posets with local realisers that have average multiplicity arbitrarily close to 1. The question becomes nontrivial (or at least not obviously trivial) if we consider both cardinality and local dimension.

Question 8. *For* $d, n \in \mathbb{N}$*, what is the minimal encoding length of a poset with local dimension* d *and cardinality* n*?*

The dimension of a lexicographic sum is determined by the dimensions of the summands and of the indexing poset, but it's not clear whether or not the same is true of local dimension. As Bosek, Gryczuk, and Trotter stated in [9], it's an open question whether or not inequality 4.4 in Proposition 58 can be improved, even when the indexing poset is an antichain.

Question 9. *Can any of the bounds in Proposition 58 be improved?*

We proved in Corollary 67 that, for every $d \in \mathbb{N}$, there exists a finite d-dimensional poset with local dimension d. However, we do not have an upper bound on the smallest possible cardinality of such a poset.

Question 10. *Given* $d \in \mathbb{N}$*, what is the smallest* n *such that there exists a* d*-dimensional poset with local dimension* n*? What is the smallest* t *such that* $\operatorname{ldim}\left(\mathbf{t}^d\right) = d$*?*

Kim et al. [32] asked whether or not $\operatorname{ldim}(\mathcal{Q}^n) = n$ for all n. We believe that this is not the case, and that our best lower bound is asymptotically correct.

Conjecture 8. *As $n \to \infty$, $\operatorname{ldim}(\mathcal{Q}^n) = \Theta\big(\frac{n}{\log n}\big)$.*

If this conjecture is true, it would imply that

$$\operatorname{ldim}(P) = O\left(\dim_2(P) / \log \dim_2(P)\right) \tag{4.24}$$

for every poset P, and hence provide a new proof of Kim et al.'s theorem that

$$\operatorname{ldim}(P) = O\big(|P| / \log |P|\big) \tag{4.25}$$

for every P, perhaps even improving it by a constant factor.

We may even propose the following stronger conjecture.

Conjecture 9. *There exists a universal constant c such that, for any $t \in \mathbb{N}$, as $n \to \infty$, $\operatorname{ldim}(\mathbf{t}^n) \sim c\frac{n}{\log_t n}$.*

One can prove a lower bound of this form using an entropy argument similar to the proof of Theorem 61. Indeed, suppose $n \ll m = n^{1+o(1)}$. Let P be a random partial order on $[(t-1)n + m]$ induced by a random function from $[(t-1)n + m]$ to $\mathbf{t}^n$ that maps $[(t-1)n]$ to the set of n-tuples with exactly one nonzero entry in lexicographic order. Each such function induces a different labelled poset. Let $d = \mathbb{E}[\operatorname{ldim}(P)]$. Then, by essentially the same argument as in the proof of Theorem 61,

$$\begin{aligned} H(P) = (1 + o(1))mn \log t \leq (m + (t-1)n)d \log(m + (t-1)n) = \\ (1 + o(1))md \log n, \end{aligned} \tag{4.26}$$

so $d \geq (1 + o(1))\frac{n}{\log_t n}$.

Even finding one cube with local dimension smaller than its dimension would be very good. If $\operatorname{ldim}\left(\mathcal{Q}^k\right) = \ell < k$, then, by subadditivity, $\operatorname{ldim}\left(\mathcal{Q}^n\right) \leq \frac{\ell}{k}n + \ell$ for all n.

Searching for good local realisers by computer does not seem feasible. Proving or disproving these conjectures will require new ideas.

CHAPTER 5

OTHER DIMENSION VARIANTS

5.1 Introduction

In this chapter, we introduce some new variants of poset dimension. We will make use of some nonstandard definitions of dimension-theoretic concepts. However, all of these definitions are easily seen to be equivalent to the standard ones given in Chapter 1.

Given a poset P, a *local realiser* of P is a set $\mathcal{L}$ of monotone partial functions from P to a chain C such that, for every x and y in P with $x \not\geq y$, there is a partial function $f \in \mathcal{L}$ such that $x, y \in \mathrm{dom}(f)$ and $f(x) < f(y)$. A local realiser of P is called a *realiser* of P if all of its elements are total functions, i.e., functions whose domains are all of P. Given an integer $t \geq 2$, a local realiser $\mathcal{L}$ of P is called a *local t–realiser* of P if the codomain of every partial function in $\mathcal{L}$ is the t-element chain $\mathbf{t}$. A t–realiser of P is a local t–realiser that is also a realiser.

The *dimension* of a poset P, denoted $\dim(P)$. is defined as the minimum cardinality of a realiser of P. For any integer $t \geq 2$, the *t–dimension* of P, denoted $\dim_t(P)$, is the minimum cardinality of a t–realiser of P. The t–dimension of P is monotone decreasing in t, and, for finite posets P, the minimum value of $\dim_t(P)$ over all values of t is equal to $\dim(P)$. The most interesting case of t–dimension is 2–dimension, as $\dim_2(P)$ is equal to the smallest d such that P embeds into the d–dimensional Boolean lattice as a suborder.

Given a local realiser $\mathcal{L}$ of P and a point $x \in P$, the multiplicity of x in $\mathcal{L}$, denoted $\mu_{\mathcal{L}}(x)$, is defined as the number of partial functions $f \in \mathcal{L}$ such that $x \in \mathrm{dom}(f)$. The local dimension of P, denoted $\mathrm{ldim}(P)$, is the minimum over all local realisers $\mathcal{L}$ of P of $\max\{\mu_{\mathcal{L}}(x) : x \in P\}$. For any integer $t \geq 2$, the local t–dimension of P, denoted $\mathrm{ldim}_t(P)$, is defined as the minimum over all local

t–realisers $\mathcal{R}$ of P of $\max\{\mu_{\mathcal{R}}(x) : x \in P\}$. Similar to the case with t–dimension, local t–dimension is monotone decreasing in t, and the minimum value of $\operatorname{ldim}_t(P)$ over all values of t is (when P is finite) equal to $\operatorname{ldim}(P)$. As with t–dimension, we will usually consider the case where $t = 2$.

The following inequalities follow immediately from the definitions, and hold for all posets P and all choices of t.

$$\operatorname{ldim}(P) \leq \operatorname{ldim}_t(P) \leq \dim_t(P), \tag{5.1}$$

$$\operatorname{ldim}(P) \leq \dim(P) \leq \dim_t(P). \tag{5.2}$$

As we will see, dimension and local t–dimension are incomparable, and there exist posets of bounded dimension and arbitrarily large local t–dimension. However, local t–dimension is bounded below by a logarithmic function of dimension:

$$\operatorname{ldim}_t(P) \geq \log_t(2\dim(P) - 1). \tag{5.3}$$

A poset parameter f is called *monotone* if, for every poset Q and every suborder P of Q, $f(P) \leq f(Q)$. It is called *subadditive* if, for all posets P and Q, $f(P \times Q) \leq f(P) + f(Q)$. Dimension, local dimension, and t–dimension are all monotone and subadditive; see [43], [32], and [55], respectively, for the proofs. We will now show that this is true for local t–dimension as well.

To prove monotonicity, let $\mathcal{R}$ be a local t–realiser of Q and P a suborder of Q. Since $\mathcal{R}|_P = \{f|_P : f \in \mathcal{R}\}$ is a local t–realiser of P whose maximum multiplicity is at most that of $\mathcal{R}$, $\operatorname{ldim}_t(P) \leq \operatorname{ldim}_t(Q)$.

For subadditivity, let P and Q be posets and let $\mathcal{R}$ and $\mathcal{S}$ be local t–realisers of P and Q respectively. Let $\mathcal{T} = \{f \circ \pi_P : f \in \mathcal{R}\} \cup \{f \circ \pi_Q : f \in \mathcal{S}\}$, where π_P and π_Q are the projection maps from $P \times Q$ onto P and Q respectively. Then $\mathcal{T}$ is a local t–realiser of $P \times Q$ and, for every (x, y), $\mu_{\mathcal{T}}(x, y) = \mu_{\mathcal{R}}(x) + \mu_{\mathcal{S}}(y)$.

5.2 Bounds on local t–dimension

Recall that $\dim_t(P)$ is equal to the smallest cardinal d such that P embeds into $\mathbf{t}^d$ as a suborder. We therefore have the trivial bound $\dim_t(P) \geq \log_t |P|$ due to the pigeonhole principle. Our first theorem shows that the same bound holds for local t–dimension. The argument we use is similar to the one used in the proofs of Theorems 6 and 7 in Part I, Chapter 2.

Theorem 73. *For every poset P with cardinality n,* $\mathrm{ldim}_t(P) \geq \log_t n$.

Proof. Let P be a poset of cardinality n and let $\mathcal{R}$ be a local t–realiser of P. For each pair of distinct elements $x, y \in P$, either $x \ngeq y$ or $y \ngeq x$. Either way, there is a partial function $f \in \mathcal{R}$ such that $f(x) \neq f(y)$. For each $f \in \mathcal{R}$, let G_f be a graph with vertex set $\mathrm{dom}(f)$ and edge set $\{xy : f(x) \neq f(y)\}$. Clearly each G_f is a complete t-partite graph, and the set $\{G_f : f \in \mathcal{R}\}$ is an edge cover of K_n.

Now, for each $f \in \mathcal{R}$, let U_f be one of the t classes of G_f, chosen independently and uniformly at random, and let U be the intersection of all the U_f's. For each edge xy of K_n, there is an $f \in \mathcal{R}$ such that x and y are in different classes of G_f, so x and y cannot both be in U_f. Therefore U is an independent set and so has at most one element. Now, for each $x \in P$, the probability that $x \in U$ is $t^{-\mu_{\mathcal{R}}(x)}$, so $\mathbb{E}[|U|] = \sum_{x \in P} t^{-\mu_{\mathcal{R}}(x)} \leq 1$. Now let $\mu = \frac{1}{n} \sum_{x \in P} \mu_{\mathcal{R}}(x)$ (i.e., the average multiplicity of $\mathcal{R}$). By convexity, $nt^{-\mu} \leq \sum_{x \in P} t^{-\mu_{\mathcal{R}}(x)} \leq 1$, and hence $\mu \geq \log_t n$.

Note that the case $t = 2$ was proved by Hansel [27]; see also Bollobás and Scott [8]. □

This bound is clearly sharp, as $\mathbf{t}^n$ has cardinality t^n and local t–dimension at most (and hence exactly) n. Hiraguchi [29] proved that a poset of dimension $n \geq 3$ has cardinality at least $2n - 1$, which implies inequality 5.3.

The next proposition shows that chains also have the smallest local t–dimension possible given their cardinality. This contrasts with t–dimension; it's a simple exercise to prove that $\dim_t(\mathbf{n}) = \lceil \frac{n-1}{t-1} \rceil$.

Proposition 74. *For all $n \in \mathbb{N}$, $\mathrm{ldim}_t(\mathbf{n}) = \lceil \log_t n \rceil$.*

Proof. Obviously, $\mathrm{ldim}_t(\mathbf{1}) = 0$. Now let $\mathcal{R}$ be a local t–realiser of $\mathbf{n}$. We construct a local t–realiser of $\mathbf{tn}$ splitting $\mathbf{tn}$ into t equal segments and taking a copy of $\mathcal{R}$ covering each segment, as well as a total function from $\mathbf{tn}$ to $\mathbf{t}$ that sends the i^{th} segment to i, for each $i \in \mathbf{t}$. This shows that $\mathrm{ldim}_t(\mathbf{tn}) \leq \mathrm{ldim}_t(\mathbf{n}) + 1$, and hence by induction $\mathrm{ldim}_t(\mathbf{n}) \leq \lceil \log_t n \rceil$ for all n. The matching lower bound follows from Theorem 73. □

Corollary 75. *For every poset P with cardinality n and every integer $t \geq 2$,*

$$\mathrm{ldim}_t(P) \leq \lceil \log_t n \rceil \,\mathrm{ldim}(P).$$

For every poset P and every pair of integers $t \geq s \geq 2$,

$$\mathrm{ldim}_t(P) \leq \lceil \log_t s \rceil \,\mathrm{ldim}_s(P).$$

□

For antichains, a similar argument shows that $\mathrm{ldim}_t(A_n) \leq 2\lceil \log_t n \rceil$. In the case $t = 2$, we can do better. It follows from Sperner's theorem that

$$\mathrm{ldim}_2(A_n) \leq \dim_2(A_n) = \min\left\{ m : \binom{m}{\lfloor m/2 \rfloor} \geq n \right\}. \tag{5.4}$$

The corresponding upper bound follows from a theorem of Bollobás and Scott [8].

Proposition 76. *For all $n \in \mathbb{N}$,*

$$\operatorname{ldim}_2(A_n) \geq \min\left\{m : \binom{m+1}{\lfloor (m+1)/2 \rfloor} \geq n+1\right\}.$$

Proof. A local 2–realiser of A_n is a set $\mathcal{R}$ of partial functions from $[n]$ to $\{0,1\}$ such that, for every ordered pair $(x, y) \in [n]^2$ with $x \neq y$, there is an $f \in \mathcal{R}$ such that $f(x) = 0$ and $f(y) = 1$. Such a set is also known as a *strongly separating system* on $[n]$. Bollobás and Scott [8] proved that, for every strongly separating system $\mathcal{R}$ on $[n]$, the sum of the cardinalities of the domains of the functions in $\mathcal{R}$ is at least kn, where k is the smallest integer such that

$$\binom{k+1}{\lfloor (k+1)/2 \rfloor} \geq n+1. \tag{5.5}$$

It follows that there exists an element of A_n whose multiplicity in $\mathcal{R}$ is at least k. □

Using Stirling's inequality to estimate the upper and lower bounds, as we did in Chapter 3, it follows that $\operatorname{ldim}_2(A_n) = \log n + \frac{1}{2}\log\log n + O(1)$. Using more precise estimates, one can show that the $O(1)$ term is at most 2.

5.3 Local 2–dimension and complete bipartite edge-coverings of graphs

Let P be a two-level poset with minimal elements A and maximal elements B, with A and B disjoint. The bipartite imcomparability graph of P is the graph with vertex set $A \cup B$ and edge set $\{ab : a \in A, b \in B, a \not< b\}$.

In [32], Kim, Martin, Masařík, Shull, Smith, Uzzell, and Wang showed that the local dimension of a two-level poset P is essentially the same (up to an additive constant) as the local difference graph covering number of the bipartite incomparability graph of P. In this section, we will show that local 2–dimension has a similar connection with the local complete bipartite covering number.

Note that a random bipartite graph with classes of cardinality n has local complete bipartite covering number $\Omega(n/\log n)$ with high probability, so in the following theorem, $\log|A|$ is typically much smaller than $\operatorname{lbc}(G)$.

Theorem 77. *Let P be a two-level poset P with minimal elements A and maximal elements B, and let G be the bipartite incomparability graph of P. Assume without loss of generality that $|A| \geq |B|$. Then*

$$\operatorname{lbc}(G) \leq \operatorname{ldim}_2(P) \leq \operatorname{lbc}(G) + \log|A| + \tfrac{1}{2}\log\log|A| + 3.$$

Proof. First we show that $\operatorname{lbc}(G) \leq \operatorname{ldim}_2(P)$. To this end, let $\mathcal{R}$ be a local 2–realiser of P. For each partial function $f \in \mathcal{R}$, let B_f be the complete bipartite graph with classes $f^{-1}(1) \cap A$ and $f^{-1}(0) \cap B$. Let $\mathcal{C} = \{B_f : f \in \mathcal{R}\}$. Now $\mathcal{C}$ is a complete bipartite edge-covering of G, and, for each $v \in P$, $\mu_{\mathcal{C}}(v) = \mu_{\mathcal{R}}(v)$.

Now we show that $\operatorname{ldim}_2(P) \leq \operatorname{lbc}(G) + \log|A| + \frac{1}{2}\log\log|A| + 3$. Let $\mathcal{C}$ be a complete bipartite edge-covering of G. For each $B \in \mathcal{C}$, define a partial function f_B with domain $V(B)$ by $f_B(a) = 1$ if $a \in A$ and $f(b) = 0$ if $b \in B$. Each such partial function is monotone and, for each $a \in A$, $b \in B$ with a and b incomparable, there is a $B \in \mathcal{C}$ such that $f_B(b) < f_B(a)$. Now define a function f with domain P by $f(a) = 0$ if $a \in A$, $f(b) = 1$ if $b \in B$. Finally, let $\mathcal{R}$ and $\mathcal{S}$ be local 2–realisers of the antichains A and B respectively. The set $\mathcal{T} = \{B_f : B \in \mathcal{C}\} \cup \mathcal{R} \cup \mathcal{S} \cup \{f\}$ is a local 2–realiser of P. For each $a \in A$, $\mu_{\mathcal{T}}(a) = \mu_{\mathcal{C}}(a) + \mu_{\mathcal{R}}(a) + 1$ and, for each $b \in B$, $\mu_{\mathcal{T}}(b) = \mu_{\mathcal{C}}(b) + \mu_{\mathcal{S}}(b) + 1$. As we saw in the previous section, $\mathcal{R}$ and $\mathcal{S}$ can be chosen so that each element has multiplicity at most

$$\min\left\{m : \tbinom{m}{\lfloor m/2 \rfloor} \geq |A|\right\} \leq \log|A| + \tfrac{1}{2}\log\log|A| + 2. \tag{5.6}$$

$\square$

Corollary 78. *Let S_n be the standard example of a poset of dimension n, namely the suborder of the n–dimensional Boolean lattice consisting of all subsets of $[n]$ of cardinality 1 and all subsets of cardinality $n-1$. For all $n \geq 2$,*

$$\operatorname{ldim}_2(S_n) \leq \log n + \tfrac{1}{2}\log\log n + 4.$$

Proof. This follows from Theorem 77 and the fact that the bipartite incomparability graph of S_n is a matching. □

The split of a poset P, first defined by Kimble (see [55]), is defined as the two-level poset Q with minimal elements $P' = \{x' : x \in P\}$ and maximal elements $P'' = \{x'' : x \in P\}$, where $x' \leq y'$ if and only if $x \leq y$ in P.

The following lemma is analogous to a lemma proved for local dimension by Barrera-Cruz, Prag, Smith, Taylor, and Trotter in [2].

Lemma 79. *Let P be a poset with n elements and let Q be the split of P. Then*

$$\operatorname{ldim}_2(Q) - \log n - \tfrac{1}{2}\log\log n - 3 \leq \operatorname{ldim}_2(P) \leq 2\operatorname{ldim}_2(Q) - 2.$$

Proof. Let $\mathcal{R}$ be a local 2–realiser of P. For each partial function $f \in \mathcal{R}$, define a partial function f' with domain $\{x' : f(x) = 1\} \cup \{x'' : f(x) = 0\}$, sending each x' and each x'' to $f(x)$. Let $\mathcal{S}$ and $\mathcal{T}$ be local 2–realisers of the antichains P' and P'', and let g be the total function that maps P' to 0 and P'' to 1. The union $\{f' : f \in \mathcal{R}\} \cup \mathcal{S} \cup \mathcal{T} \cup \{g\}$ is a local 2–realiser of Q, and $\mathcal{R}$, $\mathcal{S}$, and $\mathcal{T}$ can be chosen so that each element of Q has multiplicity at most $\operatorname{ldim}_2(P) + \operatorname{ldim}_2(A_n) + 1$. Therefore, $\operatorname{ldim}_2(Q) \leq \operatorname{ldim}_2(P) + \operatorname{ldim}_2(A_n) + 1$.

Now let $\mathcal{R}$ be a local 2–realiser of Q. For each $f \in \mathcal{R}$, we define a partial function f' with domain $\{x \in P : f(x') = 1 \text{ or } f(x'') = 0\}$, mapping x to 1 if $f(x') = 1$ and 0 if $f(x'') = 0$. Because each f is monotone and $x' < x''$ for all $x \in P$, only one of these cases can be true for each $x \in \mathrm{dom}(f')$. It is easy to check that f' is monotone (there are four cases to consider). For each x and y in P with $x \not\geq y$, $x'' \not\geq y'$, so there is an $f \in \mathcal{R}$ such that $f(x'') = 0$ and $f(y') = 1$, and hence $f'(x) < f'(y)$. Therefore $\mathcal{S} = \{f' : f \in \mathcal{R}\}$ is a local 2–realiser of P. For each $x \in P$, there is a $g \in \mathcal{R}$ such that $g(x') = 0$ and $f(x'') = 1$, so $x \notin \mathrm{dom}(g')$. It follows that $\mu_{\mathcal{S}}(x) \leq (\mu_{\mathcal{R}}(x') - 1) + (\mu_{\mathcal{R}}(x'') - 1)$, and hence $\mathrm{ldim}_2(P) \leq 2\,\mathrm{ldim}_2(Q) - 2$. □

Recall the definition of a *difference graph* from Chapter 3, and in particular the difference graphs H_n. For any $n \in \mathbb{N}$, H_n is the bipartite incomparability graph of the split of $\mathbf{n}$. A *difference graph with n steps* is a graph that can be obtained from H_n by a sequence of vertex duplications. As we saw in Chapter 3, every difference graph can be obtained in this way for some n.

In Chapter 3, we proved that a difference graph with n steps has local complete bipartite covering number equal to

$$\min\left\{k : \binom{2k}{k} \geq n+1\right\} = \tfrac{1}{2}\log n + \tfrac{1}{4}\log\log n + O(1). \tag{5.7}$$

This implies a version of Corollary 75 for graphs, namely, for every graph G on n vertices,

$$\mathrm{lbc}\,(G) \leq \mathrm{lbc}\left(H_{\lceil n/2\rceil}\right)\mathrm{ldc}\,(G) \leq \left(\tfrac{1}{2}\log n + \tfrac{1}{4}\log\log n + \tfrac{3}{2}\right)\mathrm{ldc}\,(G)\,. \tag{5.8}$$

The Erdős-Pyber theorem [19] states that, for every graph G with n vertices, $\mathrm{lbc}\,(G) = O\left(\frac{n}{\log n}\right)$. Csirmaz, Ligeti, and Tardos [13] showed that $\mathrm{lbc}\,(G) \leq (1+o(1))\frac{n}{\log n}$. We can use this to bound the local 2–dimension of any poset from above.

Theorem 80. *For every poset P with cardinality n,*

$$\operatorname{ldim}_2(P) \leq (4+o(1))\frac{n}{\log n}.$$

Proof. Let Q be the split of P and let G be the bipartite incomparability graph of Q. By Lemma 79 and Theorem 77, $\operatorname{ldim}_2(P) \leq 2\operatorname{ldim}_2(Q) \leq 2\operatorname{lbc}(G) + O(\log n)$. Since $|G| = 2n$, $\operatorname{lbc}(G) \leq (2+o(1))\frac{n}{\log n}$. $\square$

Kim et al. [32] proved that, as $n \to \infty$, there exist n-element posets with local dimension $\Omega\left(\frac{n}{\log n}\right)$. Of course the same is true for local t–dimension for every t. In Chapter 4, we improved this lower bound by a constant factor, showing that there exists an n-element poset with local dimension (and hence local t–dimension for every t) at least $\frac{n}{4\log 3n}$ for all $n \geq 2$.

By a theorem of Kierstead [31], for all integers ℓ, k, and n with $1 \leq \ell \leq k \leq n$, $\operatorname{ldim}_2\left(\mathcal{Q}^n_{\ell,k}\right) \leq \dim_2\left(\mathcal{Q}^n_{\ell,k}\right) \leq \dim_2\left(\mathcal{Q}^n_{1,k}\right) \leq \lceil e(k+1)^2 \ln n\rceil$. By Theorem 73, this bound is the best possible up to a constant factor when k is constant. Kierstead's argument can also be used to show that $\operatorname{ldim}_2\left(\mathcal{Q}^n_{1,2}\right) \leq \dim_2\left(\mathcal{Q}^n_{1,2}\right) \leq \lceil 3\log_{27/23} n\rceil$ (see also Theorem 54 in Chapter 3).

5.4 Fractional t–dimension and local t–dimension

Each of the poset parameters we have discussed can be described as the optimal solution to a certain integer program. In this section, we consider the linear programming relaxations of these programs, whose solutions are called the fractional variants of the original parameters.

A *fractional local realiser* of a poset P is a function w that assigns a nonnegative weight to each monotone partial function from P to a chain C in such a way that, for every pair $x \not\geq y$, $\sum\{w(f) : f(x) < f(y)\} \geq 1$. A fractional realiser is a fractional local realiser that assigns positive weight only to total functions, and a fractional local t–realiser is a fractional local realiser where the chain C has t elements. The fractional (local) (t)-dimension of a poset P is the minimum over all fractional (local) (t)-realisers w of $\max\left\{\sum_{x\in\mathrm{dom}(f)} w(f) : x \in P\right\}$. Following Biró, Hamburger, and Pór [4], we denote the fractional variant of a parameter by adding a superscript $\star$ to the corresponding integer parameter. Fractional dimension was introduced and studied by Brightwell and Scheinerman [11] and fractional local dimension by Smith and Trotter [49], but, as far as we know, fractional t–dimension and fractional local t–dimension have never been studied.

Like the corresponding integer parameters, these fractional parameters are easily shown to be subadditive and monotonic. Also, all of the inequalities that hold for the integer parameters hold for the fractional variants as well.

It is trivial to show that $\mathrm{ldim}^\star_{\mathrm{t}}(A_n) \leq \dim^\star_{\mathrm{t}}(A_n) \leq \frac{2t}{t-1}$ for all n and all t – just take w to be the constant function $\frac{2t}{t-1} \cdot t^{-n}$ – so fractional (local) t–dimension cannot be bounded below by a function of cardinality.

We can determine the fractional t–dimension of a chain exactly.

Theorem 81. *For all integers $n \geq t \geq 2$, $\dim^\star_{\mathrm{t}}(\mathbf{n}) = \frac{n-1}{t-1}$.*

Proof. Let w be a fractional t–realiser of $\mathbf{n}$. For each x and y in $\mathbf{n}$ such that y covers x, w must assign total weight at least 1 to the set of monotone functions f such that $f(x) < f(y)$. Conversely, each such f separates at most $t-1$ covering relations. Since $\mathbf{n}$ has $n-1$ covering relations, we have

$$\begin{aligned} \sum_{\substack{f:\mathbf{n}\to\mathbf{t} \\ f \text{ monotone}}} (t-1)w(f) \quad &\geq \quad \sum_{\substack{f:\mathbf{n}\to\mathbf{t} \\ f \text{ monotone}}} \sum_{\substack{x,y\in\mathbf{n} \\ y \text{ covers } x \\ f(x)<f(y)}} w(f) \quad = \\ & \sum_{\substack{x,y\in\mathbf{n} \\ y \text{ covers}}} \sum_{\substack{f:\mathbf{n}\to\mathbf{t} \\ f \text{ monotone} x \\ f(x)<f(y)}} w(f) \geq n-1. \end{aligned} \tag{5.9}$$

It follows that $\dim^\star_{\mathrm{t}}(\mathbf{n}) \geq \frac{n-1}{t-1}$.

To show that $\dim_t^\star(\mathbf{n}) \leq \frac{n-1}{t-1}$ when $n > t$ (the case $n = t$ is trivial), we will define a set F of montone functions from $\mathbf{n}$ to $\mathbf{t}$ such that $|F| = n - 1$ and, for every pair $x, y \in \mathbf{n}$ such that y covers x, there are exactly $t - 1$ functions $f \in F$ such that $f(x) < f(y)$. Then the function w that assigns weight $t - 1$ to each element of F and weight 0 to each monotone function not in F is a fractional t–realiser of $\mathbf{n}$ with total weight $\frac{n-1}{t-1}$.

Let F be the set of all monotone functions f from $\mathbf{n}$ to $\mathbf{t}$ with the following properties:

1. f is surjective;
2. for each $x \in \mathbf{t}$ that is not the top or bottom element, $|f^{-1}\{x\}| \leq 2$;
3. for all $x < y < z \in \mathbf{t}$, if $|f^{-1}(x)| \geq 2$ and $|f^{-1}(z)| \geq 2$, then $|f^{-1}(y)| \geq 2$.

For example, when $n = 9$ and $t = 4$, F consists of the following functions (where the first block is mapped to 1, the second to 2, and so on):

$$
\begin{array}{llll}
123456 & 7 & 8 & 9 \\
12345 & 67 & 8 & 9 \\
1234 & 56 & 78 & 9 \\
123 & 45 & 67 & 89 \\
12 & 34 & 56 & 789 \\
1 & 23 & 45 & 6789 \\
1 & 2 & 34 & 56789 \\
1 & 2 & 3 & 456789,
\end{array}
$$

and, when $n = 7$ and $t = 5$, F consists of these functions:

$$
\begin{gathered}
123\ 4\ 5\ 6\ 7 \\
12\ 34\ 5\ 6\ 7 \\
1\ 23\ 45\ 6\ 7 \\
1\ 2\ 34\ 56\ 7 \\
1\ 2\ 3\ 45\ 67 \\
1\ 2\ 3\ 4\ 567.
\end{gathered}
$$

Now we will show that F has the desired properties. First, denote the bottom and top elements of $\mathbf{t}$ by α and ω respectively, and define $A_f = |f^{-1}\{\alpha\}|$ and $\Omega_f = |f^{-1}\{\omega\}|$. If $n \geq 2t-1$, then there are $n-2t+1$ different functions $f \in F$ such that $A_f \geq 2$ and $\Omega_f \geq 2$, $t-1$ functions $f \in F$ such that $\Omega_f = 1$, and $t-1$ functions $f \in F$ such that $A_f = 1$, so $|F| = n-1$. Otherwise, $n \leq 2t-2$. In this case, the number of functions $f \in F$ such that $A_f = \Omega_f = 1$ is $2t-n-1$, the number of $f \in F$ such that $A_f \geq 2$ is $n-t$, and the number of $f \in F$ such that $\Omega_f \geq 2$ is $n-t$, so $|F| = n-1$.

Now label $\mathbf{n} = \{x_1, x_2, \ldots, x_n\}$ and $\mathbf{t} = \{y_1, y_2, \ldots, y_t\}$ in order. For each $i \in [n-1]$, we must show that there are $t-1$ functions $f \in F$ that separate x_i and x_{i+1}. This is the same as showing that there are $n-t$ functions $f \in F$ such that $f(x_i) = f(x_{i+1})$. For each $i \in [n-1]$ and each $j \in [t]$, let $F_{i,j}$ be the number of functions $f \in F$ such that $f(x_i) = f(x_{i+1}) = y_j$. First, observe that

$$F_{i,1} = \begin{cases} n-t+1-i & \text{if } i \leq n-t, \\ 0 & \text{otherwise,} \end{cases} \tag{5.10}$$

and that

$$F_{i,t} = \begin{cases} i-t+1 & \text{if } i \geq t, \\ 0 & \text{otherwise.} \end{cases} \tag{5.11}$$

For $2 \leq j \leq t-1$, it's easy to see that, if there is an $f \in F$ such that $f(x_i) = f(x_{i+1}) = y_j$, then it is unique. Therefore $F_{i,j}$ is either 1 or 0, and

$$F_{i,j} = \begin{cases} 1 & \text{if } j \leq i \leq n-t+j-1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.12}$$

Finally, for each $i \in [n-1]$, the number of functions $f \in F$ such that $f(x_i) = f(x_{i+1})$ is equal to $\sum_{j=1}^{t} F_{i,j}$. To compute this sum, we need to consider four cases. If $t \leq i \leq n-t$, then

$$\begin{aligned} \sum_{j=1}^{t} F_{i,j} = n-2t+2+|\{j : 2 \leq j \leq t-1, F_{i,j} = 1\}| = \\ n-2t+2+(t-2) = n-t. \end{aligned} \tag{5.13}$$

If $i \leq n-t$ and $i \leq t-1$, then

$$\begin{aligned} \sum_{j=1}^{t} F_{i,j} = n-t+1-i+|\{j : 2 \leq j \leq t-1, F_{i,j} = 1\}| = \\ n-t+1-i+(i-1) = n-t. \end{aligned} \tag{5.14}$$

If $i \geq n-t+1$ and $i \geq t$, then

$$\begin{aligned} \sum_{j=1}^{t} F_{i,j} = n-t+1-i+|\{j : 2 \leq j \leq t-1, F_{i,j} = 1\}| = \\ i-t+1+(n-i-1) = n-t. \end{aligned} \tag{5.15}$$

If $n-t+1 \le i \le t-1$, then

$$\sum_{j=1}^{t} F_{i,j} = |\{j : 2 \le j \le t-1, F_{i,j} = 1\}| = i - (i-n+t+1) + 1 = n-t. \quad (5.16)$$

In all four cases, $|\{f \in F : f(x_i) = f(x_{i+1})\}| = n-t$, so there are $t-1$ functions in F that separate x_i from x_{i+1}. □

We now define a concept that will be useful in proving lower bounds on fractional local t–dimension. Given a poset P and integer $t \ge 2$, a *fractional local t–antirealiser* of P is an ordered pair of functions (I, D), where $I : \{(x,y) \in P^2 : x \not\ge y\} \to [0,1]$ and $D : P \to [0,1]$, such that $\sum_{x \in P} D(x) = 1$ and, for each monotone partial function $f : P \to \mathbf{t}$,

$$\sum_{f(x)<f(y)} I(x,y) \le \sum_{x \in \operatorname{dom} f} D(x). \quad (5.17)$$

A fractional local t–antirealiser can be thought of an an obstacle to constructing a fractional local t–realiser with small local weight at every point. It can be shown using the strong linear programming duality theorem that $\operatorname{ldim}^\star_{\mathrm{t}}(P)$ is equal to the maximum of $\sum_{x \not\ge y} I(x,y)$ over all fractional local t–antirealisers (I, D) of P.

Other types of fractional antirealisers can be defined in a similar way; for example, we can define a fractional t–antirealiser of P as a function $I : \{(x,y) \in P^2 : x \not\ge y\} \to [0,1]$ such that $\sum_{f(x)<f(y)} I(x,y) \le 1$ for all monotone total functions $f : P \to \mathbf{t}$. The fractional t–dimension of P is then equal to the maximum of $\sum_{x \not\ge y} I(x,y)$ over all t–antirealisers I of P. In fact, we have already used fractional t–antirealisers implicitly in the proof of the lower bound in Theorem 81.

Proposition 82. *For all $t \ge 2$, $\dim^\star_{\mathrm{t}}(A_n) = \frac{2t}{t-1} - o(1)$ as $n \to \infty$, and the same is true of $\operatorname{ldim}^\star_{\mathrm{t}}(A_n)$.*

Proof. As mentioned earlier, the upper bound $\dim^\star_{\mathrm{t}}(A_n) \le \frac{2t}{t-1}$ is trivial. In fact, we need only assign positive weight to nonconstant functions, so

$$\dim_t^\star(A_n) \leq \frac{2t}{t-1} \cdot t^{-n}(t^n - t) = \frac{2t}{t-1}\left(1 - t^{1-n}\right). \tag{5.18}$$

Let $D(x) = \frac{1}{n}$ for all $x \in A_n$ and let $I(x,y) = \frac{2t}{(t-1)n^2}$ for all $x \neq y$. Suppose f is a partial function from A_n to $\mathbf{t}$ whose domain has k elements. Then f separates at most $\frac{t-1}{2t}k^2$ ordered pairs (i.e., the number of edges in a t-partite Turán graph on k vertices), so

$$\sum_{f(x)<f(y)} I(x,y) \leq \frac{t-1}{2t}k^2 \cdot \frac{2t}{(t-1)n^2} = \frac{k^2}{n^2} \leq \frac{k}{n}. \tag{5.19}$$

Therefore (I, D) is a fractional local t–antirealiser of A_n, so

$$\operatorname{ldim}_t^\star(A_n) \geq \frac{2t}{t-1}\left(1 - \frac{1}{n}\right). \tag{5.20}$$

$\square$

5.4.1 Fractional local t-dimension of chains

Unlike the other dimension variants, determining the fractional local t–dimension of a chain is not trivial, and in general we are unable to determine the exact value of $\operatorname{ldim}_t^\star(\mathbf{n})$. An argument similar to the proof of Proposition 74 shows that, for all integers $t \geq 2$ and $n \in \mathbb{N}$, $\operatorname{ldim}_t^\star(\mathbf{tn}) \leq \operatorname{ldim}_t^\star(\mathbf{n}) + 1$. Therefore an improvement over the trivial bound $\operatorname{ldim}_t^\star(\mathbf{n}) \leq \lceil \log_t n \rceil$ for any chain automatically yields an improvement (by an additive constant) for all chains.

The smallest n such that $\operatorname{ldim}_2^\star(\mathbf{n}) < \operatorname{ldim}_2(\mathbf{n})$ is 5. Indeed, the fractional local 2–antirealiser shown in Figure 5.1 shows that $\operatorname{ldim}_2^\star(\mathbf{3})$ is at least 2, and by the trivial bound $\operatorname{ldim}_2^\star(\mathbf{3}) \leq \operatorname{ldim}_2(\mathbf{3}) = 2$, it is exactly 2. By monotonicity, the same is true for $\operatorname{ldim}_2^\star(\mathbf{4})$. The fractional local 2–antirealiser in Figure 5.2 shows that $\operatorname{ldim}_2^\star(\mathbf{5}) \geq \frac{5}{2}$.

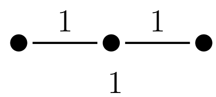

Figure 5.1: A fractional local 2–antirealiser of **3**.

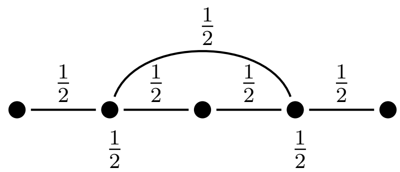

Figure 5.2: A fractional local 2–antirealiser of **5**.

The following fractional local 2–realiser covers each point with total weight at most $\frac{5}{2}$, showing that $\operatorname{ldim}_2^\star(\mathbf{5}) = \frac{5}{2}$. We denote by $w(ab\ldots c\ xy\ldots z)$ the weight of the partial monotone function that sends $a, b, \ldots c$ to 1 and $x, y, \ldots z$ to 2.

$$\begin{aligned} w(123\ 45) &= \tfrac{1}{2} \\ w(12\ 345) &= \tfrac{1}{2} \\ w(12\ 3) &= \tfrac{1}{2} \\ w(3\ 45) &= \tfrac{1}{2} \\ w(1\ 2) &= 1 \\ w(4\ 5) &= 1. \end{aligned}$$

It follows that $\operatorname{ldim}_2^\star(\mathbf{n}) \leq \lceil \log \frac{n}{5} \rceil + \frac{5}{2}$ for all $n \in \mathbb{N}$.

To find lower bounds on the fractional local t–dimension of chains, we reformulate the problem as follows. Suppose $n \in \mathbb{N}$, and consider the complete graph K_n with vertex set $[n]$. Given a natural number $t \geq 2$, an ordered t-partite graph is a complete t-partite subgraph of K_n whose parts can be ordered so that every element of the first part is less than every element of the second part, every element of the second is less than every element of the third, and so on. Then

$\mathrm{ldim}^\star_t(\mathbf{n})$ is equal to the maximum value of $\sum\limits_{e \in [n]^{(2)}} I(e)$ over all pairs of functions $D : [n] \to [0,1]$ and $I : [n]^{(2)} \to [0,1]$ such that $\sum\limits_{v \in [n]} D(v) = 1$ and, for every ordered t-partite graph G, $\sum\limits_{e \in E(G)} I(e) \leq \sum\limits_{v \in V(G)} D(v)$.

Given an edge in $xy \in E(K_n)$, the *length* of xy, denoted $\mathrm{length}(xy)$, is $|x-y|$. For each $\ell \in [n-1]$, K_n contains $n-\ell$ edges of length ℓ. An ordered t-partite graph contains at most $(t-1)\ell$ edges of length ℓ.

We will prove a lower bound on the fractional local t-dimension of chains using the following observation by Hunter Spink [52]. Let $f : [n-1] \to \mathbb{R}$ be a monotone decreasing function and let B be an ordered t-partite graph with k vertices. We claim that $\sum\limits_{e \in E(B)} f(\mathrm{length}(e))$ is maximised when B is compressed (i.e., the vertex set of B is a contiguous subset of $[n]$) and B is a Turán graph. To prove the first claim, take the largest contiguous set of vertices in B containing the leftmost vertex, and move all the vertices in this set one step to the right. Observe that this does not increase the length of any edge in B. Repeat this process until B is compressed. For the second claim, assume B is compressed. Label the parts of B $B_1, B_2, \ldots, B_t$ in order and suppose that $|B_i| > |B_{i+1}|$. Let v be the last vertex of B_i. If we move v to B_{i+1}, we lose an edge of length ℓ for each $\ell \in \big[|B_{i+1}|\big]$ and gain an edge of length k for each $i \in \big[|B_i\| - 1\big]$. Since $\big[|B_i| - 1\big] \subseteq \big[|B_{i+1}|\big]$, $\sum\limits_{e \in E(B)} f(\mathrm{length}(e))$ does not increase, and, by repeating this process, we can transform B into a Turán graph.

Before proving the theorem, we need to introduce some notation. Given an integer $a \geq 0$, the a^{th} harmonic number, denoted H_a, is equal to $\sum\limits_{k=1}^{a} \frac{1}{k}$. For all $a \in \mathbb{N}$, $H_a \geq \ln a + \gamma$, where γ is Euler's constant. If $1 \leq a \leq b$, then

$$H_b - H_a = \sum_{k=a+1}^{b} \frac{1}{k} = \int_a^b \frac{dx}{\lceil x \rceil} \leq \int_a^b \frac{dx}{x} = \ln b - \ln a. \tag{5.21}$$

Theorem 83. *For every integer $t \geq 2$,*

$$\operatorname{ldim}^{\star}_{\mathrm{t}}(\mathbf{n}) \geq \log_{\sqrt{e}\cdot t} n - O_t(1)$$

as $n \to \infty$.

Proof. Let $D(v) = \frac{1}{n}$ for each $v \in [n]$. For each edge $e \in E(K_n)$, let $I(e) = \frac{2}{(2\ln t+1)n} \cdot \frac{1}{\mathrm{length}(e)}$.

Suppose B is an ordered t-partite subgraph of K_n with k vertices. We want to show that $\sum\limits_{e \in E(B)} I(e) \leq \sum\limits_{v \in V(B)} D(v)$. By the above observation and the fact that $\sum\limits_{v \in V(B)} D(v)$ depends only on k, we may assume that B is a compressed Turán graph. Write $k = tq + r$, where q and r are integers and $0 \leq r \leq t-1$. For each natural number ℓ, the number of edges of length ℓ in $E(B)$ is equal to the number of edges of length ℓ in a compressed K_k, minus the number of edges of length ℓ in $t-r$ compressed copies of K_q and r compressed copies of K_{q+1}. For $1 \leq \ell \leq q-1$, the number of edges of length ℓ is therefore

$$tq + r - \ell - (t-r)(q-\ell) - r(q+1-\ell) = (t-1)\ell. \tag{5.22}$$

The number of edges of length q is

$$tq + r - q - r = (t-1)q. \tag{5.23}$$

For $q+1 \leq \ell \leq k$, the number of edges of length ℓ is

$$tq + r - \ell. \tag{5.24}$$

Now let $c = \frac{2}{(2\ln t+1)n}$. It follows from equations 5.22, 5.23, and 5.24 that

$$\begin{aligned}\sum_{e\in E(B)} I(e) = c\sum_{\ell=1}^{q}\frac{(t-1)\ell}{\ell} + c\sum_{\ell=q+1}^{k}\frac{tq+r-\ell}{\ell} =\\ c\big((t-1)q + k(H_k - H_q) - (tq+r-q)\big) =\\ c\big(k(H_k-H_q)-r\big).\end{aligned} \tag{5.25}$$

If $q \geq 1$, then, by inequality 5.21,

$$\begin{aligned}c\big(k(H_k-H_q)-r\big) \leq c\left(k\ln\frac{tq+r}{q} - r\right) \leq\\ c\left(k\ln t + k\frac{r}{tq} - r\right) = c\left(k\ln t + \frac{r^2}{tq}\right),\end{aligned} \tag{5.26}$$

and, since $k > 2r$,

$$\begin{aligned}c\left(k\ln t + \frac{r^2}{tq}\right) < c\left(k\ln t + \frac{r}{q}\right) \leq c\,(k\ln t + r) <\\ c\left(k\ln t + \frac{k}{2}\right) \leq ck\left(\ln t + \frac{1}{2}\right) = \frac{k}{n}.\end{aligned} \tag{5.27}$$

If $q = 0$, then $k = r \leq t-1$ and $H_q = 0$, so

$$c\big(k(H_k-H_q)-r\big) = c\big(k(H_k-1)\big) \leq ck\ln k < ck\ln t \leq \frac{k}{n}. \tag{5.28}$$

In both cases, $\sum\limits_{e\in E(B)} I(e) \leq \frac{k}{n} = \sum\limits_{v\in V(B)} D(v)$. Therefore,

$$\begin{aligned}\mathrm{ldim}^{\star}_{t}(\mathbf{n}) \geq \sum_{e\in[n]^{(2)}} I(e) = c\sum_{\ell=1}^{n}\frac{n-\ell}{\ell} =\\ cn(H_n-1) \geq \frac{2}{2\ln t+1}(\ln n + \gamma - 1) =\\ \log_{\sqrt{e}\cdot t} n - \frac{2-2\gamma}{2\ln t+1}.\end{aligned} \tag{5.29}$$

□

5.4.2 Suborders of the hypercube and posets of bounded degree

In this subsection, we consider the fractional t-dimension of two-layer suborders of the hypercube.

Brightwell and Scheinerman [11] proved that, for all $n \in \mathbb{N}$ and $2 \le k \le n-1$, $\dim^\star\left(\mathcal{Q}^n_{1,k}\right) = k+1$. Smith and Trotter [49] determined the exact value of $\lim_{n\to\infty} \operatorname{ldim}^\star\left(\mathcal{Q}^n_{1,k}\right)$ for all k, and showed that it is equal to $\frac{k}{\ln k - \ln\ln k - o(1)}$ as $k \to \infty$.

The following theorem shows that Brightwell and Sheinerman's result is within a constant factor of the correct value for fractional t–dimension.

Theorem 84. *For every integer $k \ge 1$, as $n \to \infty$,*

$$\dim^\star_2\left(\mathcal{Q}^n_{1,k}\right) \to \left(1 - \tfrac{1}{k+1}\right)^{-k} \cdot (k+1) \le e(k+1).$$

Proof. Every function f from $[n]$ to $\mathbf{2}$ can be extended to a monotone function $f' : \mathcal{Q}^n_{1,k} \to \mathbf{2}$, where $f'(A) = \max\{f(a) : a \in A\}$. Assume n is a multiple of $k+1$ and write $\ell = k+1$, $m = \frac{n}{\ell}$. Define a function w that assigns weight $\binom{\ell(m-1)}{m-1}^{-1}$ to every function f', where $f : [n] \to \mathbf{2}$ and $|f^{-1}\{1\}| = \frac{k}{k+1}n$, and weight 0 to every other monotone function. For every pair (A,x) where $A \in [n]^{(k)}$ and $x \in [n] \setminus A$, there are $\binom{\ell(m-1)}{m-1}$ functions $f : [n] \to \mathbf{2}$ such that $f'(A) < f'\{x\}$, so the total weight of all such functions is 1. It's easy to check that all other non-relations are covered with total weight at least 1. Now the total number of monotone functions with positive weight is $\binom{\ell m}{m}$, so the total weight of all these functions is

$$\begin{aligned}\binom{\ell m}{m}\binom{\ell(m-1)}{m-1}^{-1} &= \frac{(\ell m)!}{m!((\ell-1)m)!} \cdot \frac{(m-1)!((\ell-1)(m-1))!}{(\ell(m-1))!} = \\ \ell \cdot &\frac{(\ell m - 1)\cdot(\ell m - 2)\cdot \dots \cdot (\ell(m-1)+1)}{((\ell-1)m)\cdot((\ell-1)m-1)\cdot\dots\cdot((\ell-1)(m-1)+1)} \le \\ &\ell\left(\frac{\ell m}{(\ell-1)(m-1)}\right)^{\ell-1},\end{aligned} \tag{5.30}$$

which goes to $\ell \cdot \left(\frac{\ell}{\ell-1}\right)^{\ell-1} = \ell \cdot \left(1 - \frac{1}{\ell}\right)^{1-\ell} \leq e\ell$ as $m \to \infty$. Because $\dim_2^\star\left(\mathcal{Q}_{1,k}^n\right) = \dim_2^\star\left(\mathcal{Q}_{1,\ell-1}^{\ell m}\right)$ is monotone increasing in m, we have $\dim_2^\star\left(\mathcal{Q}_{1,\ell-1}^n\right) \leq \ell \cdot \left(1 - \frac{1}{\ell}\right)^{1-\ell}$ for all n.

Now, for the lower bound, we will construct a fractional 2-antirealiser of $\mathcal{Q}_{1,k}^n$. As before, let $\ell = k+1$ and assume $n = \ell m$, where m is an integer. For every pair (A, x) with $A \in [n]^{(k)}$ and $x \in [n] \setminus A$, let $I(A, x) = \frac{k!}{m(km)^k}$. Now suppose f is a montone function from $\mathcal{Q}_{1,k}^n$ to $\mathbf{2}$, and define a function $g : [n] \to \mathbf{2}$, where $g(x) = f\{x\}$. Now, if given $A \in [n]^{(k)}$ and $x \in [n] \setminus A$, if $f(A) < f\{x\}$, then $g'(A) < f\{x\}$. We may therefore assume that $f = g'$ without reducing the number of separated pairs. Now let p be the number of elements $x \in [n]$ such that $g(x) = 2$. The number of pairs (A, x) separated by g' is $p \cdot \binom{n-p}{k} \leq \frac{1}{k!}p(n-p)^k$, and the right side of this inequality is maximised when $p = \frac{n}{k+1} = m$. Hence every monotone function separates at most $\frac{1}{k!}m(km)^k$ pairs, so the sum of $I(A, x)$ over all such pairs is at most 1. Therefore I is a fractional 2-antirealiser of $\mathcal{Q}_{1,k}^n$, so

$$\begin{aligned}\dim_2^\star\left(\mathcal{Q}_{1,k}^n\right) \geq \ell\binom{\ell m}{\ell} \cdot \frac{(\ell-1)!}{m\big((\ell-1)m\big)^{\ell-1}} =\\ \frac{(\ell m)^\ell - O(m^{\ell-1})}{m\big((\ell-1)m\big)^{\ell-1}} = \ell \cdot \left(\frac{\ell}{\ell-1}\right)^{\ell-1} - O\left(\tfrac{1}{m}\right) =\\ \left(1 - \tfrac{1}{k+1}\right)^{-k} \cdot (k+1) - O\left(\tfrac{1}{n}\right).\end{aligned} \tag{5.31}$$

□

Recall that the *outdegree* of an element x of a poset P is the number of elements of P that are strictly greater than x. Using Theorem 84, we can bound the fractional 2-dimension of any poset by a function of its maximum outdegree. We first need the following lemma.

Lemma 85. *Let P be a poset and let Q be the split of P. Then $\dim_2^\star(P) \leq \dim_2^\star(Q)$.*

Proof. Let w be a fractional 2-realiser of Q. For each monotone $f : Q \to \mathbf{2}$, define a function $f' : P \to \mathbf{2}$, where $f'(x) = \max\{f(y') : y \leq x\}$. It's clear that f' is monotone. Now, for each monotone $g : P \to \mathbf{2}$, let $w'(g) = w(f)$ if $g = f'$ for some monotone $f : Q \to \mathbf{2}$ and $w'(g) = 0$ otherwise. Suppose $a \not\geq_P b$. Then $a'' \not\geq_Q b'$, so the total w-weight of all montone functions f such that $f(a'') = 0$ and $f(b') = 1$ is at least 1. For each such f, $f'(c') = 0$ for all $c \leq a$, so $f'(a) = 0$, and $f'(b) = 1$. Hence the pair (a, b) is separated with total weight at least 1, so w' is a fractional 2-realiser of P with the same total weight as w. □

Corollary 86. *Let P be a poset with maximum outdegree υ. Then $\dim_2^\star(P) \leq e(\upsilon + 2)$.*

Proof. Let Q be the split of P. By Lemma 85, $\dim_2^\star(P) \leq \dim_2^\star(Q)$. Since Q has maximum outdegree $\upsilon + 1$, its dual can be embedded into $\mathcal{Q}^n_{1,\upsilon+1}$ for some large n. Therefore, by Theorem 84, $\dim_2^\star(Q) \leq e(\upsilon + 2)$. □

5.5 Open problems

By Theorem 73 and Kierstead's theorem, we know that, for fixed $\ell < k$, $\operatorname{ldim}_t\left(\mathcal{Q}^n_{\ell,k}\right) = \Theta_{t,\ell,k}(\log n)$ as $n \to \infty$. However, the constant factors on the upper and lower bounds are very far apart, and we would like to know if they can be improved.

Question 11. *Given $1 \leq \ell < k \leq n$ and $t \geq 2$, what is $\operatorname{ldim}_t\left(\mathcal{Q}^n_{\ell,k}\right)$? In particular, what is $\operatorname{ldim}_2\left(\mathcal{Q}^n_{1,2}\right)$?*

The local dimension of $\mathcal{Q}^n$ is still unknown. The best known lower bound is $\Omega\left(\frac{n}{\log n}\right)$, but the only known upper bound is n. Maybe studying the local t–dimension of $\mathcal{Q}^n$ will help solve this problem.

Question 12. *What is $\operatorname{ldim}_t(\mathcal{Q}^n)$ for $t \geq 3$? Is it ever strictly less than n? In general, what is $\operatorname{ldim}_t(\mathbf{s}^n)$ when $t > s$?*

The maximum local t–dimension of an n-element poset is $\Theta\left(\frac{n}{\log n}\right)$, with upper and lower bounds that do not depend on t. This leads to the next question.

Question 13. *What is the maximum local t–dimension of an n-element poset? Does it depend on t?*

Of course, all of the natural questions asked of the other parameters (e.g., the maximum and minimum value for n-element posets, the value for the Boolean lattice and for its two-layer suborders, etc.) can be asked of fractional t–dimension and fractional local t–dimension as well.

It follows from Theorem 81 that, for every integer $t \geq 2$ and every $n \in \mathbb{N}$, $\dim_t^\star(\mathbf{n}) = \lceil \dim_t^\star(\mathbf{n}) \rceil$. This motivates the following problem.

Problem 14. *Characterise the posets P for which* $\dim_t(P) = \lceil \dim_t^\star(P) \rceil$.

Proposition 74 and Theorem 83 together imply that $\mathrm{ldim}_t^\star(\mathbf{n}) = \Theta_t(\log n)$ as $n \to \infty$, for every fixed $t \geq 2$. However, we do not have a formula for the exact fractional local t-dimension of a chain.

Problem 15. *What is the exact value of* $\mathrm{ldim}_t^\star(\mathbf{n})$, *for all integers $n \geq t \geq 2$?*

By an argument similar to the proof of Inequality 4.4 in Proposition 58 of Chapter 4, for any $t \geq 2$ and all $m, n \in \mathbb{N}$, $\mathrm{ldim}_t^\star(\mathbf{mn}) \leq \mathrm{ldim}_t^\star(\mathbf{m}) + \mathrm{ldim}_t^\star(\mathbf{n})$. It follows that, if $\mathrm{ldim}_t^\star(\mathbf{m}) < \log_t m$ for any m, then we can improve the trivial upper bound $\mathrm{ldim}_t^\star(\mathbf{n}) \leq \lceil \log_t n \rceil$ by a constant factor for all n. However, we do not know of any examples of such m for any t.

One immediate corollary of Theorem 84 is that the functions $\mathrm{FD}_t(k) = \lim_{n\to\infty} \dim_t^\star\left(\mathcal{Q}_{1,k}^n\right)$ and $\mathrm{FLD}_t(k) = \lim_{n\to\infty} \mathrm{ldim}_t^\star\left(\mathcal{Q}_{1,k}^n\right)$ are well-defined for every integer $t \geq 2$. Theorem 84 establishes the exact value of $\mathrm{FD}_2(k)$, and shows that it is equal to $(e - o(1))(k+1)$ as $k \to \infty$. Brightwell and Scheinerman's results in [11] give a lower bound for $\mathrm{FD}_t(k)$, and Smith and Trotter's results in [49] give a lower bound for $\mathrm{FLD}_t(k)$.

Question 16. *What is the exact value of* $\mathrm{FD}_t(k)$ *and* $\mathrm{FLD}_t(k)$*, for all integers* $t \geq 2$ *and* $k \in \mathbb{N}$*?*

Let $\mathrm{MFD}_t(\Delta)$ be the supremum of $\dim_{\mathrm{t}}^{\star}(P)$ over all posets whose comparability graphs have maximum degree Δ. Similarly, let $\mathrm{MFLD}_t(\Delta)$ be the supremum of $\mathrm{ldim}_{\mathrm{t}}^{\star}(P)$ over all posets whose comparability graphs have maximum degree Δ. It follows from Corollary 86 that these functions are well-defined.

Question 17. *What is the exact value of* $\mathrm{MFD}_t(\Delta)$ *and* $\mathrm{MFLD}_t(\Delta)$*, for all integers* $t \geq 2$ *and* $\Delta \in \mathbb{N}$*?*

We hope to solve these problems in the near future.

BIBLIOGRAPHY

[1] Rodrigo Angelo. A Cameron and Erdős conjecture on counting primitive sets. *arXiv preprint arXiv:1711.08107*, 2017.

[2] Fidel Barrera-Cruz, Thomas Prag, Heather C. Smith, Libby Taylor, and William T. Trotter. Comparing Dushnik-Miller dimension, Boolean dimension and local dimension. *Order*, pages 1–27, 2019.

[3] M. Barrus, M. Ferrara, J. Vandenbussche, and P. Wenger. Colored saturation parameters for rainbow subgraphs. *J. Graph Theory*, 86(4):375–386, 2017.

[4] Csaba Biró, Peter Hamburger, and Attila Pór. The proof of the removable pair conjecture for fractional dimension. *arXiv preprint arXiv:1312.7332*, 2013.

[5] B. Bollobás. Determination of extremal graphs by using weights. *Wiss. Z. Techn. Hochsch. Ilmenau*, 13:419–421, 1967.

[6] B. Bollobás. *Combinatorics: set systems, hypergraphs, families of vectors, and combinatorial probability.* Cambridge University Press, 1986.

[7] Béla Bollobás. *Modern graph theory*, volume 184. Springer Science & Business Media, 2013.

[8] Béla Bollobás and Alex Scott. On separating systems. *European Journal of Combinatorics*, 28, 2007.

[9] Bartłomiej Bosek, Jarosław Grytczuk, and William T. Trotter. Local dimension is unbounded for planar posets. *arXiv preprint arXiv:1712.06099*, 2017.

[10] G. R. Brightwell, H. A. Kierstead, Alexandr V. Kostochka, and W. T. Trotter. The dimension of suborders of the Boolean lattice. *Order*, 11(2):127–134, 1994.

[11] Graham R. Brightwell and Edward R. Scheinerman. Fractional dimension of partial orders. *Order*, 9(2):139–158, 1992.

[12] Peter J. Cameron and P. Erdős. On the number of sets of integers with various properties. *Number Theory (RA Mollin, ed.)*, pages 61–79, 1990.

[13] László Csirmaz, Péter Ligeti, and Gábor Tardos. Erdős–Pyber theorem for hypergraphs and secret sharing. *Graphs and Combinatorics*, 31(5):1335–1346, 2015.

[14] Gábor Damásdi, Stefan Felsner, António Girão, Balász Keszegh, David Lewis, Dániel T. Nagy, and Torsten Ueckerdt. A note on covering Young diagrams with applications to local dimension of posets. *arXiv preprint arXiv:2001.06367*, 2020.

[15] Gábor Damásdi, Balázs Keszegh, and Dániel T. Nagy. Covering difference graphs. *Unpublished note*, 2019.

[16] Ben Dushnik. Concerning a certain set of arrangements. *Proceedings of the American Mathematical Society*, 1, 12 1950.

[17] P. Erdős, A. Hajnal, and J. Moon. A problem in graph theory. *Am. Math. Mon.*, 71(10):1107–1110, 1964.

[18] P. Erdős and G. Szekeres. Über die anzahl der abelschen gruppen gegebener ordnung und über ein verwandtes zahlentheoretisches problem. *Acta Univ. Szeged*, 7(2):95–102, 1934.

[19] P Erdős and László Pyber. Covering a graph by complete bipartite graphs. *Discrete mathematics*, 170(1-3):249–251, 1997.

[20] J. Faudree, R. Faudree, and J. Schmitt. A survey of minimum saturated graphs. *Electron. J. Comb.*, 1000:19–29, 2011.

[21] Stefan Felsner and Torsten Ueckerdt. A note on covering Young diagrams with applications to local dimension of posets. *arXiv preprint arXiv:1902.08223*, 2019.

[22] M. Ferrara, D. Johnston, S. Loeb, F. Pfender, A. Schulte, H. C. Smith, E. Sullivan, M. Tait, and C. Tompkins. On edge-colored saturation problems, 2017.

[23] Zoltán Füredi. The order dimension of two levels of the Boolean lattice. *Order*, 11(1):15–28, 1994.

[24] Zoltán Füredi and Jeff Kahn. On the dimensions of ordered sets of bounded degree. *Order*, 3(1):15–20, 1986.

[25] Carl Friedrich Gauss. Untersuchungen über höhere arithmetik. deutsch herausgegeben von h. maser, 1965.

[26] António Girão, David Lewis, and Kamil Popielarz. Rainbow saturation of graphs. *Journal of Graph Theory*, 2017.

[27] G. Hansel. Nombre minimal de contacts de fermeture nécessaires pour réaliser une fonction booléenne symétrique de n variables. *C. R. Acad. Sci. Paris*, 258, 1964.

[28] D. Hanson and B. Toft. Edge-colored saturated graphs. *J. Graph theory*, 11(2):191–196, 1987.

[29] Toshio Hiraguchi. On the dimension of orders. *The science reports of the Kanazawa University*, 4(4):1–20, 1955.

[30] L. Kászonyi and Z. Tuza. Saturated graphs with minimal number of edges. *J. Graph theory*, 10(2):203–210, 1986.

[31] Hal A Kierstead. On the order dimension of 1-sets versus k-sets. *Journal of Combinatorial Theory, Series A*, 73(2):219–228, 1996.

[32] Jinha Kim, Ryan R. Martin, Tomáš Masařík, Warren Shull, Heather C. Smith, Andrew Uzzell, and Zhiyu Wang. On difference graphs and the local dimension of posets. *European Journal of Combinatorics*, 86:103074, 2020.

[33] T.P. Kirkman. On a problem in combinatorics. *Cambridge Dublin Math. J.*, 2:191–204, 1847.

[34] D Kleitman and B Rothschild. The number of finite topologies. *Proceedings of the American mathematical Society*, 25(2):276–282, 1970.

[35] J. Knopfmacher and W. Zhang. *Number theory arising from finite fields: analytic and probabilistic theory.* CRC Press, 2001.

[36] John Knopfmacher. *Abstract analytic number theory.* North-Holland, 1975.

[37] D. Korándi. Rainbow saturation and graph capacities. *SIAM J. Discrete Math.*, 32(2):1261–1264, 2018.

[38] Alexandr V. Kostochka. The dimension of neighboring levels of the Boolean lattice. *Order*, 14(3):267–268, 1997.

[39] Patryk Kozieł and Małgorzata Sulkowska. Uniform random posets. *Information Sciences*, 515:294–301, 2020.

[40] Edmund Landau. Neuer beweis des primzahlsatzes und beweis des primidealsatzes. *Mathematische Annalen*, 56(4):645–670, 1903.

[41] David Lewis and Victor Souza. The order dimension of divisibility. *arXiv preprint arXiv:2001.08549*, 2020.

[42] Hong Liu, Péter Pál Pach, and Richárd Palincza. The number of maximum primitive sets of integers. *arXiv preprint arXiv:1805.06341*, 2018.

[43] Ben Dushnik; E. W. Miller. Partially ordered sets. *American Journal of Mathematics*, 63, 07 1941.

[44] Vítězslav Novák. On the well dimension of ordered sets. *Czechoslovak Mathematical Journal*, 19(1):1–16, 1969.

[45] Øystein Ore. *Theory of graphs*, volume 38. American Mathematical Soc., 1965.

[46] Herbert Robbins. A remark on Stirling's formula. *American Mathematical Monthly*, 62, 04 1955.

[47] Alex Scott and David R. Wood. Better bounds for poset dimension and boxicity. *arXiv preprint arXiv:1804.03271*, 2018.

[48] Claude Elwood Shannon. A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423, 1948.

[49] Heather C. Smith and William T. Trotter. Fractional local dimension. *arXiv preprint arXiv:1906.05839*, 2019.

[50] Joel Spencer. Minimal scrambling sets of simple orders. *Acta Mathematica Hungarica*, 22, 1971.

[51] Joel Spencer. Asymptotic lower bounds for Ramsey functions. *Discrete Mathematics*, 20:69–76, 1977.

[52] Hunter Spink. personal communication, 2020.

[53] Peter L. Hammer; Uri N. Peled; Xiaorong Sun. Difference graphs. *Discrete Applied Mathematics*, 28, 1990.

[54] William T. Trotter. Inequalities in dimension theory for posets. *Proceedings of the American Mathematical Society*, 47, 02 1975.

[55] William T. Trotter. *Combinatorics and partially ordered sets: Dimension theory*, volume 6. JHU Press, 1992.

[56] Zoltán Füredi; Péter Hajnal; Vojtech Rödl; Willian T. Trotter. Interval orders and shift graphs. *Sets, Graphs, and Numbers*, 60, 1991.

[57] P. Turán. Eine extremalaufgabe aus der graphentheorie. *Mat. Fiz. Lapok*, 48(436-452):61, 1941.

[58] Torsten Ueckerdt. Order and geometry workshop, 2016.

[59] Zhiyu Wang. personal communication, 2019.

[60] A. Zykov. On some properties of linear complexes (О некоторых свойствах линейных комплексов). *Mat. Sb.*, 24(2):163–188, 1949.