Extremal problems in the cube and the grid and other combinatorial results

Eero-Pekka Johannes Räty

Trinity College
and
Department of Pure Mathematics and Mathematical Statistics
University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

April 2021

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared here and specified in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification.

Chapter 6 is based on joint work with Imre Leader.

Eero-Pekka Johannes Räty

Extremal problems in the cube and the grid and other combinatorial results

Eero-Pekka Johannes Räty

Abstract

This dissertation contains results from various areas of combinatorics.

In Chapters 2, 3 and 4 we consider questions in the area of isoperimetric inequalities. In Chapter 2, we find the exact classification of all subsets $A \subseteq \{0,1\}^n$ for which both A and A^c minimise the size of the neighbourhood, which answers a question of Aubrun and Szarek. Harper's inequality implies that the initial segments of the simplicial order satisfy these conditions, but we prove that in general there are non-trivial examples of such sets as well.

In Chapter 3, we consider the zero-deletion shadow, which is closely related to the general coordinate deletion shadow introduced by Danh and Daykin. We prove that there is a certain order on $[k]^n = \{0, \dots, k-1\}^n$, the n-dimensional grid of side-length k, whose initial segments minimise the size of the zero-deletion shadow.

In Chapter 4, we consider the following generalisation of the Kruskal-Katona theorem on $[k]^n$. For a set $A \subseteq [k]^n$, define the d-shadow of A to be the set of all points x obtained from any $y \in A$ by replacing one non-zero coordinate of y by 0. We find an order on $[k]^n$ whose initial segments minimise the size of the d-shadow.

In Chapter 5, we consider a certain combinatorial game called Toucher-Isolator game that is played on the edges of a given graph G. The value of the game on G measures how many vertices of G one of the players can achieve by using the edges claimed by her. We find the exact value of the game when G is a path or a cycle of a given length, and we prove that among the trees on n vertices, the path on n vertices has the least value of the game. These results improve previous bounds obtained by Dowden, Kang, Mikalački and Stojaković.

In Chapter 6, we consider a problem in Ramsey Theory related to the Hales-Jewett theorem. We prove that for any 2-colouring of $[3]^n$ there exists a monochromatic combinatorial line whose active coordinate set is an interval, provided that n is large. This disproves a conjecture of Conlon and Kamčev.

In Chapter 7, we give a construction of a graph G that is P_6-induced-saturated, where P_6 is the path on 6 vertices. This answers a question of Axenovich and Csikós.

Contents

1 Introduction **1**

2 Uniqueness in Harper's vertex-isoperimetric theorem **5**

2.1 Introduction . . . 5

2.2 Construction of an example . . . 8

2.3 Classifying all extremal sets . . . 9

2.4 The weak version . . . 21

2.5 A uniqueness result for certain sizes . . . 26

3 Coordinate deletion of zeroes **30**

3.1 Introduction . . . 30

3.2 Deletion on $\{0, 1\}$-sequences . . . 33

3.3 The main theorem . . . 34

3.4 An extremal result for the generalised shadow . . . 40

4 A grid generalisation of the Kruskal-Katona theorem **43**

4.1 Introduction . . . 43

4.2 Proof of Theorem 34 . . . 46

4.2.1 The compression operators . . . 46

4.2.2 The special case $n = 2$. . . . 48

4.2.3 General observations . . . 49

4.2.4 Rest of the proof . . . 51

4.3 Minimal d-shadow for a given rank . . . 62

5 The Toucher-Isolator game **63**

5.1 Introduction . . . 63

5.2 Toucher-Isolator game on a cycle and a path . . . 66

5.2.1 The lower bound . . . 67

5.2.2 The upper bound . . . 70

5.3 Toucher-Isolator game on trees . . . 78

5.3.1 First phase of the game . . . 81

5.3.2 Delayed version of the game . . . 83

6 Intervals in the Hales-Jewett theorem **100**
6.1 Introduction 100
6.2 Proof of Theorem 46 101

7 Induced saturation of P_6 **103**
7.1 Introduction 103
7.2 The construction 103

8 Bibliography **106**

Chapter 1

Introduction

In this dissertation, we consider six questions in the area of extremal combinatorics.

Chapters 2, 3 and 4 are concerned with questions in the area of discrete isoperimetric inequalities. Perhaps the best-known discrete isoperimetric inequality is Harper's theorem on the n-dimensional hypercube $Q_n = \{0,1\}^n$. We say that vertices x, y in Q_n are *neighbours* if they differ in exactly one coordinate, and more generally define the *distance* of x and y to be $d(x,y) = |\{i : x_i \neq y_i\}|$. The *neighbourhood* of a set $A \subseteq Q_n$ consists of the set A itself together with all other elements that are a neighbour of some element of A. The classical result of Harper [23] states that among the sets of a given size, the initial segment of the simplicial order has the smallest neighbourhood.

Let A be an initial segment of the simplicial order. Since its complement A^c is also isomorphic to an initial segment of the simplicial order, Harper's theorem implies that both A and A^c have neighbourhoods of minimal size. Aubrun and Szarek [1] asked if the initial segments are the only sets (up to isomorphism) for which both A and A^c have neighbourhoods of minimal size. In Chapter 2, we give a negative answer to their question. We go on to give an exact classification of all subsets A for which both A and A^c have neighbourhoods of minimal size.

Define the *exact Hamming ball* of radius r centered at x to be $B(x,r) = \{y \in Q_n : d(x,y) \leq r\}$, and define A to be a *Hamming ball* if there exist x and r so that $B(x,r) \subseteq A \subset B(x,r+1)$. It turns out that all the extremal sets A are contained between two exact Hamming balls with the same center and radius differing by 2 - that is, there exist x and r for which $B(x,r) \subseteq A \subset B(x,r+2)$. Rather surprisingly, it turns out that the only Hamming balls which are extremal are the initial segments of the simplicial order. This chapter is based on [41].

Another well-known discrete isoperimetric inequality is the Kruskal-Katona theorem [25, 28]. Let $\{0,1\}^n_r$ denote the set of all $\{0,1\}$-sequences of length n containing exactly r coordinates that equal 1. The *lower shadow* of $A \subseteq \{0,1\}^n_r$ is the set of points in $\{0,1\}^n_{r-1}$ that can be obtained by flipping exactly one 1-entry to 0 from a point in A. Similarly, the *upper shadow* of $A \subseteq \{0,1\}^n_r$ is the set of points in $\{0,1\}^n_{r+1}$ obtained by flipping exactly one 0-entry to 1 from a point in A. The Kruskal-Katona theorem states that the size of the lower shadow is minimised by the initial segments of the colexicographic order, or equivalently, the size of the upper shadow is minimised by the initial segments of the lexicographic order.

Instead of changing the value of a coordinate, it is also natural to define an operator which acts by deleting a coordinate. For $A \subseteq [k]^n = \{0,\ldots,k-1\}^n$, define the *coordinate deletion* *shadow* of A to be the set of points obtained by deleting exactly one coordinate from a point in A. Danh and Daykin [16] proved that the initial segments of the simplicial order minimise the size of the coordinate deletion shadow when $k = 2$. However, Leck [30] proved that no such order exists when $k \geq 3$. Bollobás and Leader [9] proved that the exact subcubes $[m]^n$ have minimal coordinate deletion shadow in $[k]^n$ for all m. Furthermore, they conjectured that for all m and r, the subset of $[m]^n$ containing the sequences with at most r coordinates that equal $m-1$ has minimal coordinate deletion shadow. This conjecture remains open.

Define the *rank* of a sequence to be the sum of its coordinates. The usual lower shadow decreases the rank by 1 and preserves the dimension n, while the coordinate deletion shadow decreases the dimension by 1 but there is no control on how it changes the rank. There is, however, an operator called *zero-deletion shadow* which preserves the rank but reduces the dimension by one, and hence comes 'between' the lower shadow and the coordinate deletion shadow. Define the zero-deletion shadow of $A \subseteq [k]^n$ to be the set of points obtained by removing one coordinate that equals 0 from any point in A.

Note that the zero-deletion shadow of any subset of $\{1, \dots, k-1\}^n \subseteq [k]^n$ is empty, and in order to minimise the size of the zero-deletion shadow in general, it seems natural to choose points with as few zeroes as possible. In particular, it is natural to guess that for each $0 \leq i \leq n$ the sets containing all points with at most i zeroes have zero-deletion shadow of minimal size. In Chapter 3 we prove that the initial segments of a certain order minimise the size of the zero-deletion shadow, and indeed points with fewer zeroes are preferred by this order. As a consequence, it follows that the sets containing all points with at most i zeroes indeed have zero-deletion shadow of minimal size. This chapter is based on [38].

Various generalisations of the Kruskal-Katona theorem itself has been studied in general grids $[k]^n$. One such generalisation of the upper shadow was considered by Clements [13]. For a set $A \subseteq [k]^n$, define the d^+*-shadow* of A to be the set of all points x obtained from any $y \in A$ by replacing one coordinate of y that equals 0 by any element in $\{1, \dots, k-1\}$. Clements found an order whose initial segments minimise the size of the d^+-shadow. In Chapter 4, we consider the following closely related operator. For a set $A \subseteq [k]^n$, define the d-shadow of A to be the set of all points x obtained from any $y \in A$ by replacing one coordinate of y that is in $\{1, \dots, k-1\}$ by 0.

Again, it seems natural to prefer points containing as many zeroes as possible, and this indeed turns out to be the case. We prove that there is a certain order whose initial segments minimise the size of the d-shadow, and it follows that the sets containing points with at least i zeroes have d-shadow of minimal size. Furthermore, we prove that the restrictions of the initial segments of our order minimise the size of the d-shadow on levels containing points with a given number of zeroes. We have recently learnt that the main result of this chapter may be deduced from a result of Frankl, Füredi and Kalai [19], and of London [32]. Given that, this chapter should be viewed as giving a new proof of their result. This chapter is based on [36].

In Chapter 5, we consider the following combinatorial game introduced by Dowden, Kang, Mikalački and Stojaković. The game is played on the edges of a given graph G by two players, Toucher and Isolator. They claim edges on alternating turns, with the first move given to Toucher. Toucher is aiming to maximise the number of vertices incident with some edge she has claimed, and Isolator is aiming to minimise this number. Equivalently, Isolator is trying to maximise the number of vertices so that she has claimed all of their incident edges.

We say that a vertex is *isolated* if it is not incident with any of the edges claimed by Toucher. Define the value of the game $u(G)$ to be the number of isolated vertices at the end of the game when both players play under optimal strategies. Dowden, Kang, Mikalački and Stojaković had specific interest in the case when G is a tree, and they gave bounds for the value of the game in terms of the degree sequence of G. As a consequence, they proved that when G is a tree, the asymptotic proportion of isolated vertices is between $1/8$ and $1/2$. Note that the upper bound is attained when G is a star on n vertices. For certain trees they noted that the general bounds can be improved. As an example, they proved that for P_n, the path on n vertices, the asymptotic proportion of isolated vertices is between $3/16$ and $1/4$. Furthermore, they conjectured that the correct asymptotic proportion of isolated vertices on the path should be $1/5$.

In Chapter 5, we improve both the general bound and the bound on the path. We prove that $1/5$ is the correct asymptotic proportion of isolated vertices for a path, and in fact, we find the exact value of the game for a path. We also find the exact value of the game for a cycle. Our other main result is to improve the general lower bound for trees, and we prove that the value of the game for any tree T with n vertices is at least as large as the value of the game for P_n. Therefore, the asymptotic proportion of isolated vertices in a tree is at least $1/5$. These results are based on [37, 40].

In Chapter 6, we consider a problem related to the Hales-Jewett theorem [22]. A set $L \subseteq [k]^n$ is called a *combinatorial line* if there exist a non-empty set S and integers $a_i \in [k]$ so that

$$L = \{(x_1, \ldots, x_n) : x_i = a_i \text{ for all } i \notin S \text{ and } x_i = x_j \text{ for all } i, j \in S\}.$$

The set S is called the *active coordinate set* of L. The Hales-Jewett theorem states that whenever $[k]^n$ is r-coloured, there is a monochromatic combinatorial line provided that n is large enough. The least such n is denoted by $HJ(k, r)$.

Following Shelah's proof [42] of Hales-Jewett theorem, it can be shown that provided n is large, one can always find a monochromatic combinatorial line whose active coordinate set is a union of at most $HJ(k-1, r)$ intervals. Since $HJ(2, r) = r$, it follows that one can find a monochromatic combinatorial line in $[3]^n$ whose active coordinate set is a union of at most r intervals. Conlon and Kamčev proved that this bound is tight when r is odd, and they conjectured the same to be true when r is even. In this chapter we prove that whenever $[3]^n$ is 2-coloured and n is sufficiently large, there exists a monochromatic combinatorial line whose active coordinate set is actually an interval. This disproves the conjecture of Conlon and Kamčev. This chapter is joint work with Leader based on [29].

In Chapter 7, we consider a graph saturation question. A graph G is said to be *H-induced-saturated* if G does not contain H as an induced subgraph, but removing any edge from G or adding any edge from G^c to G creates an induced copy of H. It is easy to see that an empty graph is P_2-induced-saturated and a clique is P_3-induced saturated, where again P_n denotes the path on n vertices. Martin and Smith studied a similar problem, and their results imply that there is no graph that is P_4-induced-saturated. Axenovich and Csikós gave examples of families of trees for which an induced saturated graph exists. However, they noted that these results did not cover P_n for any $n \geq 5$, and they asked whether P_n is induced saturated for any $n \geq 5$. In Chapter 7 we give a construction which proves that P_6 is induced saturated. This chapter is based on [39].

Throughout the thesis, we use standard graph theoretic and combinatorial notation. We write $[k]$ for $\{0, \ldots, k-1\}$ and $[k]^n$ for $\{0, \ldots, k-1\}^n$, and we often use the shorthand $X = \{1, \ldots, n\}$. As usual, for any set A we write $A^{(r)}$ for $\{B \subseteq A : |B| = r\}$, and also $A^{(\leq r)}$ for $\{B \subseteq A : |B| \leq r\}$ and $A^{(\geq r)}$ for $\{B \subseteq A : |B| \leq r\}$.

Chapter 2

Uniqueness in Harper's vertex-isoperimetric theorem

2.1 Introduction

The n-dimensional hypercube Q_n has vertex-set the power set $\mathcal{P}(\{1,\ldots,n\})$ with metric $d(x,y) = |x\Delta y|$. We can also view Q_n as $\{0,1\}^n$, the set of $\{0,1\}$-sequences of length n, with the metric $d(x,y) = |\{i : x_i \neq y_i\}|$. For a subset A of the hypercube Q_n, define the *neighbourhood* of A to be the set $N(A) = \{x \in Q_n : d(x,A) \leq 1\}$ where $d(x,A) = \min_{y\in A} d(x,y)$. Also more generally for each $t > 0$ define $N^t(A) = \{x \in Q_n : d(x,A) \leq t\}$, and note that we have $N^t(A) = N\left(N^{t-1}(A)\right)$.

In order to state Harper's vertex-isoperimetric theorem we need a few definitions. For any n and $0 \leq r \leq n$ define the *lexicographic order* on $\{x : x \subseteq \{1,\ldots,n\}, |x| = r\}$ to be given by $x <_{lex} y$ if $\min(x\Delta y) \in x$, and define the *colexicographic order* to be given by $x <_{colex} y$ if $\max(x\Delta y) \in y$. Define the *simplicial order* on Q_n to be given by $x <_{sim} y$ if

$$|x| < |y| \text{ or } (|x| = |y| \text{ and } x <_{lex} y).$$

Theorem 1. *(Harper, [23]). Let A be a subset of Q_n and let B be the initial segment of the simplicial order of size $|A|$. Then we have $|N(A)| \geq |N(B)|$.* □

For a general introduction to the vertex-isoperimetric theorem, see e.g. Bollobás [6, Chapter 16].

It turns out that the sets for which Harper's theorem holds with equality are not in general unique. As a trivial example, any subset of Q_2 of size 2 has minimal vertex boundary and not all such sets are isomorphic. There are more interesting and less trivial examples as well.

It is easy to verify that if A is an initial segment of the simplicial order then so is $N(A)$. Hence Harper's theorem implies that among the subsets $A \subseteq Q_n$ of a given size, the size of $N^t(A)$ is minimised when A is chosen to be the initial segment of the simplicial order. We say that $N^t(A)$ is *minimal* if $\left|N^t(B)\right| \geq \left|N^t(A)\right|$ for all $B \subseteq Q_n$ satisfying $|B| = |A|$. Let C be the initial segment of the simplicial order of size $|A|$. It is useful to observe that Harper's theorem implies that $N^t(A)$ is minimal if and only if $\left|N^t(A)\right| = \left|N^t(C)\right|$.

We say that two subsets A and B of Q_n are *isomorphic* if there exists an isometry θ of Q_n satisfying $\theta(A) = B$. In this chapter we consider the following question of Aubrun and Szarek [1, Exercise 5.66]: If $A \subseteq Q_n$ for which both $N^t(A)$ and $N^t(A^c)$ are minimal for all $t > 0$, does it follow that A is isomorphic to an initial segment of the simplicial order? For convenience, we say that A is *extremal* if $N^t(A)$ and $N^t(A^c)$ are minimal for all $t > 0$. Let C be the initial segment of the simplicial order of size $|A|$. Since C^c is isomorphic to the initial segment of the simplicial order of size $|C^c|$, it follows from Harper's inequality that A is extremal if and only if we have $\left|N^t(A)\right| = \left|N^t(C)\right|$ and $\left|N^t(A^c)\right| = \left|N^t(C^c)\right|$ for all $t > 0$. In particular, the initial segments of the simplicial order are extremal.

Define the *exact Hamming ball* of radius r centred at x to be $B(x,r) = \{y \in Q_n : d(x,y) \leq r\}$, and define A to be a *Hamming ball* if there exist x and r for which we have $B(x,r) \subseteq A \subset B(x,r+1)$. Note that $B(\emptyset, r)$ is the initial segment of the simplicial order of size $\sum_{i=0}^{r} \binom{n}{i}$, and every initial segment of the simplicial order is a Hamming ball. Later in the chapter we sometimes consider exact Hamming balls of radius r centred at x in $\mathcal{P}(\{1,\ldots,n\} \setminus \{i\})$ rather than in $\mathcal{P}(\{1,\ldots,n\})$. In order to highlight this difference, we write $B_i(x,r)$ for the exact Hamming ball of radius r centred at x with respect to ground set $\{1,\ldots,n\} \setminus \{i\}$.

Note that requiring only $N^t(A)$ to be minimal for all $t > 0$ is not a strong enough condition to guarantee that A should be isomorphic to an initial segment of the simplicial order. Indeed, as a trivial example one could take $A = B(x,r) \setminus \{x\}$ for $r \geq 1$. Then $N^t(A) = B(x, r+t)$ for all $t > 0$, and hence $N^t(A)$ is always minimal, yet A is not isomorphic to an initial segment of the simplicial order.

It turns out that the answer to the question of Aubrun and Szarek is negative, and we present a counterexample in Section 2.2. It turns out that all the extremal sets are contained between two exact Hamming balls with the same centre and radius differing by 2 - that is, if A is extremal, there exist x and r with $B(x,r) \subseteq A \subset B(x,r+2)$. Rather surprisingly, it turns out that the only Hamming balls which are extremal *are* the initial segments of the simplicial order.

The second aim of this chapter is to classify all the extremal sets A up to isomorphism. In order to state the result, we need some notation. We write $X = \{1,\ldots,n\}$, $X^{(r)} = \{x \subseteq X : |x| = r\}$, $X^{(\geq r)} = \{x \subseteq X : |x| \geq r\}$, $X^{(\leq r)} = \{x \subseteq X : |x| \leq r\}$, $X_i = \{1,\ldots,n\} \setminus \{i\}$ and $X_{i,j} = \{1,\ldots,n\} \setminus \{i,j\}$. Throughout this chapter, we denote the elements of Q_n by lower case letters, the subsets of Q_n by upper case letters and the set systems on $X^{(r)}$ by calligraphy letters.

Define the maps $\pi_i : X^{(r+1)} \to X_i^{(r)} \cup X_i^{(r+1)}$ by setting $\pi_i(x) = x \setminus \{i\}$ for all $x \in X^{(r+1)}$. For a set system $\mathcal{B} \subseteq X^{(r+1)}$ define $\pi_i(\mathcal{B}) = \{\pi_i(x) : x \in \mathcal{B}\}$. Note that π_i is a bijection from $X^{(r+1)}$ to $X_i^{(r)} \cup X_i^{(r+1)}$, and hence it follows that $|\pi_i(\mathcal{B})| = |\mathcal{B}|$ for all i.

It is known that the exact Hamming balls are the only sets of their respective sizes for which the inequality in Harper's theorem holds with equality. That is, if $A \subseteq Q_n$ is a set of size $\left|X^{(\leq r)}\right|$ for which the size of $N(A)$ is minimal among the subsets of Q_n of the same size, then $A = B(x,r)$ for some $x \in Q_n$. Hence if $A \subseteq Q_n$ is an extremal set of size $\left|X^{(\leq r)}\right|$ for some r, it certainly follows that A has to be isomorphic to the initial segment of the simplicial order.

Note that if A is extremal, then so is A^c, as the definition of extremality is symmetric under taking complements. Let $G_r = X^{(\leq r)} \cup \left\{b \in X^{(r+1)} : 1 \in b\right\}$. It is easy to check that we have

$|G_r| + |G_{n-r-2}| = 2^n$ for all r. Hence if $A \subseteq Q_n$ satisfies $|A| \neq \left|X^{(\leq r)}\right|$ for all r, then at least one of the inequalities $\left|X^{(\leq r)}\right| < |A| \leq |G_r|$ or $\left|X^{(\leq r)}\right| < |A^c| \leq |G_r|$ is satisfied for some r. Hence it is sufficient to only classify the extremal sets $A \subseteq Q_n$ for which there exists r such that $\left|X^{(\leq r)}\right| < |A| \leq |G_r|$.

For convenience, we write $f_r = f_{n,r} = \left|X^{(\leq r)}\right| = \sum_{j=0}^{r} \binom{n}{j}$ and $g_r = g_{n,r} = |G_r| = \sum_{j=0}^{r} \binom{n}{j} + \binom{n-1}{r}$. In both cases, the dependence on n will not be highlighted if the value of n is clear from the context.

Let s be an integer of the form $s = f_r + k$ for some $0 \leq k \leq \binom{n-1}{r}$. Note that for a fixed n, the value of s uniquely determines the values of r and k. Furthermore, observe that we have $f_n + \binom{n-1}{r} = g_r$. Hence the set of integers that can be written as $f_r + k$ for some $0 \leq k \leq \binom{n-1}{r}$ is exactly the set of those integers s' for which there exists r satisfying $f_r \leq s' \leq g_r$.

Given an integer s of the form $s = f_r + k$, let $\mathcal{A}$ be the initial segment of the lexicographic order on $X^{(r+1)}$ of size k. For $A \subseteq \mathcal{P}(X_i)$ write $\{i\} + A$ for the family $\{\{i\} \cup a : a \in A\}$. For each i define

$$A_i = X^{(\leq r)} \cup (\{i\} + \pi_i(\mathcal{A})).$$

Note that we have $\pi_i(\mathcal{A}) \subseteq X_i^{(r)} \cup X_i^{(r+1)}$, and hence $\{i\} + \pi_i(\mathcal{A})$ is a well-defined subset of $X^{(r+1)} \cup X^{(r+2)}$ of the same size as $\mathcal{A}$. In particular, the sets $X^{(\leq r)}$ and $\{i\} + \pi_i(\mathcal{A})$ are disjoint, and hence each set A_i contains exactly s elements. We also have $\mathcal{A} \subseteq \{1\} + X_1^{(r)}$ since $k \leq \binom{n-1}{r}$. Hence it follows that $\{1\} + \pi_1(\mathcal{A}) = \mathcal{A}$, and thus A_1 is the initial segment of the simplicial order. Note that some of the sets A_i might be isomorphic to each other.

Now we are ready to state the classification of extremal sets.

Theorem 2. *(Classification of extremal sets). Let $A \subseteq Q_n$ be a subset of size s, where $s = f_r + k$ for some r and $k \leq \binom{n-1}{r}$. Let $A_1, \ldots, A_n$ be the sets defined as above for these choices of r and k. Then A is extremal if and only if A is isomorphic to some A_i.*

It is natural to ask what happens if we weaken the notion of extremality. A natural way to do this is to seek subsets $A \subseteq Q_n$ for which both $N(A)$ and $N(A^c)$ have minimal sizes among the subsets of Q_n of the same sizes. We say that $A \subseteq Q_n$ is *weakly extremal* if for all $B \subseteq Q_n$ with $|B| = |A|$ we have $|N(B)| \geq |N(A)|$ and $|N(B^c)| \geq |N(A^c)|$. Let C be the initial segment of the simplicial order of size $|A|$. Again, by Harper's theorem, weak extremality of A is equivalent to having $|N(A)| = |N(C)|$ and $|N(A^c)| = |N(C^c)|$. Rather surprisingly, we prove in Section 2.4 that the notions of weak extremality and extremality coincide.

Theorem 3. *Let $A \subseteq Q_n$ be a subset for which every $B \subseteq Q_n$ with $|B| = |A|$ satisfies $|N(B)| \geq |N(A)|$ and $|N(B^c)| \geq |N(A^c)|$. Then for all $t > 0$ and $B \subseteq Q_n$ with $|B| = |A|$ we have $\left|N^t(B)\right| \geq \left|N^t(A)\right|$ and $\left|N^t(B^c)\right| \geq \left|N^t(A^c)\right|$.*

Hence it immediately follows that Theorem 2 holds when extremality is replaced with weak extremality. In fact, Theorem 3 follows from the following slightly stronger Theorem which we also prove in Section 2.4.

Theorem 4. *Let $A \subseteq Q_n$ be a subset for which every $B \subseteq Q_n$ with $|B| = |A|$ satisfies $|N(B)| \geq |N(A)|$. Then for all $t > 0$ and $B \subseteq Q_n$ with $|B| = |A|$ we have $\left|N^t(B)\right| \geq \left|N^t(A)\right|$.*

The plan of this chapter is as follows. In Section 2.2 we construct an extremal set which is not isomorphic to an initial segment of the simplicial order. In Section 2.3 we prove Theorem 2. In Section 2.4 we consider weakly extremal sets and prove Theorem 3.

Recall that exact Hamming balls are the unique sets of their respective sizes for which equality holds in Harper's inequality. In Section 2.5 we prove another near-uniqueness result: we show that there exists only one set B_r of size g_r, apart from the initial segment, for which equality holds in Harper's inequality. In fact, the set B_r is also an extremal set, and we describe it already in Section 2.2.

2.2 Construction of an example

In this section we find for every r an extremal set $B_r \subseteq Q_n$ satisfying $|B_r| = g_r$ for which B_r is not isomorphic to an initial segment of the simplicial order. Let C_r be the initial segment of the simplicial order of size g_r on Q_n. Then C_r can be written as $C_r = B(\emptyset, r) \cup B(\{1\}, r)$. Define $B_r = B(\emptyset, r) \cup B(\{1,2\}, r)$. That is, B_r is the union of two exact Hamming balls of same radius and centres within distance 2 apart from each other.

Note that if $B(s,r) \subseteq B_r \subseteq B(s, r+1)$ holds for some $s \in Q_n$, then the first inclusion implies that we have $s = \emptyset$ or $s = \{1,2\}$. However, the second inclusion is violated in both cases. Hence B_r is not a Hamming ball, and hence it cannot be isomorphic to an initial segment of the simplicial order.

It is easy to verify that we have $N^t(C_r) = C_{r+t}$ and $N^t(B_r) = B_{r+t}$ for all $t > 0$. Hence in order to prove that $N^t(B_r)$ are minimal for all $t > 0$, it suffices to prove that we have $|B_m| = |C_m|$ for all m. Observe that B_r can be written as $B_r = X^{(\leq r)} \cup \left(\{1,2\} + \left(X_{1,2}^{(r-1)} \cup X_{1,2}^{(r)}\right)\right)$. Hence it follows that

$$|B_r| = f_r + \binom{n-2}{r-1} + \binom{n-2}{r} = f_r + \binom{n-1}{r} = g_r,$$

and thus we have $|B_r| = |C_r|$ for all r.

Since

$$C_r^c = B(\{1,\ldots,n\}, n-r-1) \cap B(\{2,\ldots,n\}, n-r-1),$$

it is easy to check that we have

$$C_r^c = B(\{1,\ldots,n\}, n-r-2) \cup B(\{2,\ldots,n\}, n-r-2).$$

Hence C_r^c is isomorphic to C_{n-r-2} under the isometry $\theta : Q_n \to Q_n$ given by $\theta(a) = a^c$ for all $a \in Q_n$. In particular, it follows that $g_r + g_{n-r-2} = 2^n$.

Similarly we have

$$B_r^c = B(\{1,\ldots,n\}, n-r-1) \cap B(\{3,\ldots,n\}, n-r-1)$$

and our aim is to show that this implies that

$$B_r^c = B(\{1,3,\ldots,n\}, n-r-2) \cup B(\{2,\ldots,n\}, n-r-2). \tag{2.1}$$

Indeed, note that for any $x \in B(\{1,3,\ldots,n\}, n-r-2)$ we have

$$d(x,\{1,\ldots,n\}) \leq d(x,\{1,3\ldots,n\}) + d(\{1,3,\ldots,n\},\{1,\ldots,n\}) \leq (n-r-2)+1 = n-r-1$$

by the triangle inequality. One can similarly deduce that we have $d(x,\{3,\ldots,n\}) \leq n-r-1$. Hence it follows that

$$B(\{1,3,\ldots,n\}, n-r-2) \subseteq B(\{1,\ldots,n\}, n-r-1) \cap B(\{3,\ldots,n\}, n-r-1),$$

and similarly we have

$$B(\{2,\ldots,n\}, n-r-2) \subseteq B(\{1,\ldots,n\}, n-r-1) \cap B(\{3,\ldots,n\}, n-r-1).$$

These two observations imply that the $\supseteq$-part of (2.1) holds.

Note that $B(\{1,3,\ldots,n\}, n-r-2) \cup B(\{2,\ldots,n\}, n-r-2)$ is isomorphic to B_{n-r-2} under the isometry ϕ given by $\phi(a) = a\Delta\{2,\ldots,n\}$ for all $a \in Q_n$. Hence we have

$$|B(\{1,3\ldots,n\}, n-r-2) \cup B(\{2,\ldots,n\}, n-r-2)| = |B_{n-r-2}| = g_{n-r-2} = 2^n - |B_r| = |B_r^c|.$$

This, together with the fact that the inclusion holds in the $\supseteq$-direction, implies that (2.1) holds. In particular, B_r^c is isomorphic to B_{n-r-2}.

Hence $N^t(B_r^c)$ is isomorphic to $B_{n-r+t-2}$ and $N^t(C_r^c)$ is isomorphic to $C_{n-r+t-2}$ for all $t > 0$. Since $|B_{n-r+t-2}| = |C_{n-r+t-2}|$, it follows that $\left|N^t(B_r^c)\right| = \left|N^t(C_r^c)\right|$ for all $t > 0$, and hence $N^t(B_r^c)$ are minimal for all $t > 0$. Therefore B_r is an extremal set.

2.3 Classifying all extremal sets

Recall that $f_r = \sum_{i=0}^{r}\binom{n}{i}$ is the size of the exact Hamming ball of radius r, and $g_r = \sum_{i=0}^{r}\binom{n}{i} + \binom{n-1}{r}$ is the size of the initial segment $X^{(\leq r)} \cup \left(\{1\} + X_1^{(r)}\right)$. It is convenient to exclude the sets of size f_r from the classification, and this is possible due to the following result.

Proposition 5. *Let $A \subseteq Q_n$ be a set satisfying $|A| = f_r$ and suppose that for any $B \subseteq Q_n$ satisfying $|B| = f_r$ we have $|N(B)| \geq |N(A)|$. Then $A = B(x,r)$ for some $x \in Q_n$.* □

Since this is a well-known fact, the proof is omitted. This could be deduced by induction on n and applying Lemma 6 of Katona from [24]. A similar technique will be used in Section 2.5 in the proof of Claim 1 in Theorem 21.

Since the classification of extremal sets $A \subseteq Q_n$ satisfying $|A| = f_r$ for some r is covered by Proposition 5, it is enough to consider only those sets $A \subseteq Q_n$ satisfying $f_r < |A| < f_{r+1}$ for some r. Furthermore, since $g_r + g_{n-2-r} = 2^n$ and $f_r + f_{n-1-r} = 2^n$, by considering A^c if necessary, it is enough to classify only those extremal sets $A \subseteq Q_n$ satisfying $f_r < |A| \leq g_r$ for some r. Hence, from now on, we will assume that $A \subseteq Q_n$ is an extremal set satisfying $f_r < |A| \leq g_r$ for some r.

Lemma 6. *Let $A \subseteq Q_n$ be an extremal set satisfying $f_r < |A| \leq g_r$ for some r. Then there exist x, y, $z \in Q_n$ with $y \neq z$ satisfying $d(x,y) \leq 1$, $d(x,z) \leq 1$ and for which $B(x,r) \subseteq A \subseteq$*

$B(y, r+1) \cap B(z, r+1)$.

Since $d(x, y) \leq 1$, the condition $A \subseteq B(y, r+1)$ implies that we also have $A \subseteq B(x, r+2)$. Hence it follows that there exists an element $x \in Q_n$ for which we have $B(x, r) \subseteq A \subseteq B(x, r+2)$. This implies that the interesting behaviour in the set A occurs only on two layers of the cube, namely on those which are distance $r+1$ and $r+2$ apart from x.

Proof. Let $A \subseteq Q_n$ satisfying the condition $f_r < |A| \leq g_r$. Let C_r be the initial segment of size g_r, and recall from Section 2.2 that we have $N^t(C_r) = C_{r+t}$. Since $|A| \leq g_r$ and A is extremal, it follows that $\left|N^{n-r-2}(A)\right| \leq \left|N^{n-r-2}(C_r)\right| = g_{n-2} = 2^n - 2$. Hence there exist distinct elements u, $v \in N^{n-r-2}(A)^c$. Let $y = u^c$ and $z = v^c$. By the choice of u and v, it follows that $B(u, n-r-2) \cup B(v, n-r-2) \subseteq A^c$. Taking complements implies that we have $A \subseteq B(y, r+1) \cap B(z, r+1)$.

Since $|A| > f_r$, it follows that $|A^c| < f_{n-r-1}$. Combining this with the extremality of A, it follows that $|N^r(A^c)| \leq |N^r(B(\emptyset, n-r-1))| = f_{n-1} = 2^n - 1$. Hence there exists $x \in (N^r(A^c))^c$ and therefore we have $B(x, r) \subseteq A$. Combining this with the previous observations implies that $B(x, r) \subseteq A \subseteq B(y, r+1) \cap B(z, r+1)$. Since $B(x, r) \subseteq B(y, r+1)$, we must have $d(x, y) \leq 1$, and similarly it follows that $d(x, z) \leq 1$. □

The proof of Lemma 6 gives some insight on why it is convenient to assume that the size of A satisfies the condition $f_r < |A| \leq g_r$ rather than only $f_r < |A| < f_{r+1}$. Indeed, the condition $f_r < |A| < f_{r+1}$ would not be strong enough to guarantee the existence of both y and z.

Given this result, we can split the rest of the classification into two parts: considering those A which are Hamming balls, i.e. for which there exist $x \in Q_n$ and r satisfying $B(x, r) \subseteq A \subseteq B(x, r+1)$, and considering those A for which no such x and r exist. It turns out that all the extremal sets apart from the initial segment appear in the second case. This is proved in Proposition 11, but before that we need a few preliminary results. Many of these preliminary results are used later as well.

For $A \subseteq Q_n$, define the *i-sections* $A_{i,+}$ and $A_{i,-}$ of A by setting $A_{i,+} = \{a \setminus \{i\} : a \in A, i \in a\}$ and $A_{i,-} = \{a : a \in A, i \notin a\}$. Note that $A_{i,+}$ and $A_{i,-}$ are subsets of $\mathcal{P}(X_i)$ which is naturally isomorphic to Q_{n-1}. If the direction i is clear from the context, they will be denoted as A_+ and A_-.

We often want to relate the neighbourhood of A to the neighbourhoods of the i-sections $A_{i,+}$ and $A_{i,-}$. However, for $A_{i,+}$ and $A_{i,-}$ the neighbourhood is always taken inside $\mathcal{P}(X_i)$, i.e. with respect to the ground set X_i rather than X. Since for i-sections $A_{i,+}$ and $A_{i,-}$ we always consider the neighbourhood with respect to X_i, and otherwise we always consider neighbourhood with respect to X, we will use the same notation $N(A)$ and $N(A_{i,\pm})$ in both cases to avoid excessive use of subscripts. In the second case the neighbourhood $N(A_{i,\pm})$ that is considered is a neighbourhood of an i-section, and hence it should always be understood as the neighbourhood with respect to the ground set X_i.

Lemma 7. *For all $r \geq 0$ and for any distinct elements x, $y \in Q_n$ we have $|B(x, r) \cup B(y, r)| \geq g_r$, with equality if and only if $d(x, y) \leq 2$.*

Proof. We may assume that $y \nsubseteq x$. Set $A = B(x, r) \cup B(y, r)$. Recall that $B_i(x, r)$ denotes the exact Hamming ball of radius r centred at x with respect to the ground set X_i. For any

$i \in y \setminus x$ we have $A_{i,-} = B_i(x,r) \cup B_i(y \setminus \{i\}, r-1)$ and $A_{i,+} = B_i(x, r-1) \cup B_i(y \setminus \{i\}, r)$. In particular, we have $B_i(x,r) \subseteq A_{i,-}$ and $B_i(y \setminus \{i\}, r) \subseteq A_{i,+}$, and thus it follows that $|A| = |A_{i,+}| + |A_{i,-}| \geq 2f_{n-1,r}$.

Let C be the initial segment of the simplicial order of size $g_{n,r}$, and recall that we have $C = B(\emptyset, r) \cup B(\{1\}, r)$. Note that $C_{i,-} = C_{i,+} = B_1(\emptyset, r)$, and hence it follows that $g_{n,r} = 2f_{n-1,r}$. In particular, we obtain that $|A| \geq g_{n,r}$.

Note that the equality holds if and only if $A_{i,-} = B_i(x,r)$ and $A_{i,+} = B_i(y \setminus \{i\}, r)$. Hence we must have $B_i(y \setminus \{i\}, r-1) \subseteq B_i(x,r)$ and $B_i(x, r-1) \subseteq B_i(y \setminus \{i\}, r)$, which are satisfied if and only if $d(y \setminus \{i\}, x) \leq 1$. Since $i \in y \setminus x$, this inequality is satisfied if and only if $d(x,y) \leq 2$. □

Lemma 8. *Let G be the set of isometries ϕ of Q_n satisfying $\phi(\emptyset) = \emptyset$. For each $\sigma \in S_n$, define ϕ_σ by setting $\phi_\sigma(a) = \{\sigma(i) : i \in a\}$ for $a \in Q_n$. Then we have $G = \{\phi_\sigma : \sigma \in S_n\}$.*

Proof. It is clear that each ϕ_σ is an isometry satisfying $\phi_\sigma(\emptyset) = \emptyset$. Let $g \in G$. Since $g(\emptyset) = \emptyset$, it follows that for all $i \in X$ we have $d(g(\{i\}), \emptyset) = 1$, and thus $g(\{i\})$ is a set containing exactly one element. Denote this unique element by g_i, and define σ by setting $\sigma(i) = g_i$. Since g is a bijection, it follows that $\sigma : X \to X$ is an injection, and hence we have $\sigma \in S_n$.

Our aim is to prove that we have $g(a) = \phi_\sigma(a)$ for all $a \in Q_n$. We prove this by induction on the number of elements in a. Note that the claim is true for any $a \in Q_n$ with $|a| \in \{0,1\}$ by the construction of σ and the fact that $g(\emptyset) = \emptyset$.

Suppose that the claim is true for all $b \in Q_n$ with $|b| \leq m-1$ where $m \geq 2$, and let $a \in Q_n$ with $|a| = m$. For each $i \in a$ we have $g(a \setminus \{i\}) = \phi_\sigma(a \setminus \{i\})$ by induction. Note that $|\phi_\sigma(a \setminus \{i\})| = |a| - 1 = m - 1$ and $|g(a)| = d(g(a), \emptyset) = d(a, \emptyset) = m$. We also have $d(g(a), g(a \setminus \{i\})) = d(a, a \setminus \{i\}) = 1$. In particular, $g(a)$ is a set containing m elements, $g(a \setminus \{i\})$ is a set containing $m-1$ elements and the symmetric difference of $g(a)$ and $g(a \setminus \{i\})$ contains one element. Hence it follows that $g(a) = g(a \setminus \{i\}) \cup \{j\} = \phi_\sigma(a \setminus \{i\}) \cup \{j\}$ for some $j \notin g(a \setminus \{i\})$.

Note that $d(g(a), g(\{i\})) = d(a, \{i\}) = m - 1$. Combining this with the previous observations $|g(a)| = m$ and $|g(\{i\})| = |\{\sigma(i)\}| = 1$, it follows that $\sigma(i) \in g(a)$. Since $\sigma(i) \notin \phi_\sigma(a \setminus \{i\})$, we must have $j = \sigma(i)$. Hence it follows that $g(a) = \phi_\sigma(a \setminus \{i\}) \cup \{\sigma(i)\} = \phi_\sigma(a)$, as required. □

Note that ϕ_σ induces a bijection on $X^{(r)}$ for every $\sigma \in S_n$. We say that $\mathcal{A} \subseteq X^{(r)}$ and $\mathcal{B} \subseteq X^{(r)}$ are *isomorphic as subsets of* $X^{(r)}$ if there exists ϕ_σ for which we have $\phi_\sigma(\mathcal{A}) = \mathcal{B}$. In order to avoid confusion between this notion and the notion of being isomorphic on Q_n, we will explicitly write down 'isomorphic as subsets of $X^{(r)}$' rather than 'isomorphic'. The aim of the following lemma is to relate these two notions.

Lemma 9. *Let A and C be subsets of Q_n of the form $A = B(\emptyset, r) \cup \mathcal{A}$ and $C = B(\emptyset, r) \cup \mathcal{C}$ for some $\mathcal{A}$, $\mathcal{C} \subseteq X^{(r+1)}$. If A and C are isomorphic, then $\mathcal{A}$ and $\mathcal{C}$ are isomorphic as subsets of $X^{(r+1)}$.*

Proof. Note that it suffices to show that if A and C are isomorphic, then there exists $\phi_\sigma \in G$ with $\phi_\sigma(A) = C$. Let g be an isometry mapping A to C. Hence g also maps A^c to C^c. If $g(\emptyset) = \emptyset$, we are done by Lemma 8, and hence we may assume that $g(\emptyset) \neq \emptyset$.

Since g is an isometry, it follows that $B(g(\emptyset), r) = g(B(\emptyset, r))$, and hence we have $B(g(\emptyset), r) \subseteq C$. Thus $B(\emptyset, r) \cup B(g(\emptyset), r) \subseteq C$, and by Lemma 7 we have $|C| \geq g_r$. Furthermore, the equality holds if and only if $C = B(\emptyset, r) \cup B(g(\emptyset), r)$ and $d(g(\emptyset), \emptyset) \leq 2$.

Since $B(\{1, \ldots, n\}, n-r-2) \subseteq A^c$, it follows similarly that $B(g(\{1, \ldots, n\}), n-r-2) \subseteq C^c$. Since g is an isometry with $g(\emptyset) \neq \emptyset$, it follows that $g(\{1, \ldots, n\}) \neq \{1, \ldots, n\}$ as $\emptyset$ and $\{1, \ldots, n\}$ are antipodal points. Hence we have

$$B(\{1, \ldots, n\}, n-r-2) \cup B(g(\{1, \ldots, n\}), n-r-2) \subseteq C^c$$

and thus Lemma 7 implies that $|C^c| \geq g_{n-r-2}$.

Since $g_r + g_{n-r-2} = 2^n$, the equality must hold in both cases. Thus $C = B(\emptyset, r) \cup B(g(\emptyset), r)$, and since $C \subseteq B(\emptyset, r+1)$ we must have $d(g(\emptyset), \emptyset) = 1$. Hence $A = B\left(g^{-1}(\emptyset), r\right) \cup B(\emptyset, r)$, and we also have $d\left(g^{-1}(\emptyset), \emptyset\right) = 1$ since g is an isometry.

Let i, $j \in X$ be chosen so that $g(\emptyset) = \{i\}$ and $g^{-1}(\emptyset) = j$. Since $A = B(\{j\}, r) \cup B(\emptyset, r)$ and $C = B(\{i\}, r) \cup B(\emptyset, r)$, by choosing $\sigma = (ij) \in S_n$ it follows that $\phi_\sigma(A) = C$, as required. □

Lemma 10. *Let t be a fixed positive integer, $A \subseteq Q_n$ and let j be a direction for which we have $|A_{j,-}| \geq |A_{j,+}|$. Let $C_{j,-}$ and $C_{j,+}$ be the initial segments of the simplicial order on $\mathcal{P}(X_j)$ with $|C_{j,-}| = |A_{j,-}|$ and $|C_{j,+}| = |A_{j,+}|$. If $N^t(A)$ is minimal and $C_{j,-} \subseteq N(C_{j,+})$, then $N^t(A_{j,-})$ and $N^t(A_{j,+})$ are minimal and we have $N^{t-1}(A_{j,+}) \subseteq N^t(A_{j,-})$ and $N^{t-1}(A_{j,-}) \subseteq N^t(A_{j,+})$.*

The idea of comparing the neighbourhoods of A_+ and A_- with the neighbourhoods of C_+ and C_- was used similarly by Bollobás and Leader [8].

Proof. For simplicity we denote the j-sections by A_+, A_-, C_+ and C_-. Define $C \subseteq Q_n$ by setting $C = C_- \cup (C_+ + \{j\})$. It is easy to observe that we have $\left(N^t(A)\right)_+ = N^{t-1}(A_-) \cup N^t(A_+)$ and $\left(N^t(A)\right)_- = N^{t-1}(A_+) \cup N^t(A_-)$. Hence it follows that

$$\left|N^t(A)\right| = \left|N^{t-1}(A_-) \cup N^t(A_+)\right| + \left|N^{t-1}(A_+) \cup N^t(A_-)\right|. \tag{2.2}$$

Since C_+ and C_- are initial segments of the simplicial order on $\mathcal{P}(X_j)$ with $|C_+| \leq |C_-|$, it follows that $C_+ \subseteq C_-$ as initial segments are nested. Hence we also have $N^{t-1}(C_+) \subseteq N^{t-1}(C_-) \subseteq N^t(C_-)$, as taking neighbourhoods preserves the inclusion of sets, and a set is always contained in its neighbourhood. Similarly the assumption $C_- \subseteq N(C_+)$ implies that we have $N^{t-1}(C_-) \subseteq N^t(C_+)$. Applying (2.2) for C implies that

$$\left|N^t(C)\right| = \left|N^t(C_+)\right| + \left|N^t(C_-)\right|. \tag{2.3}$$

Since C_+ and C_- are initial segments, Harper's inequality implies that we have $\left|N^t(A_+)\right| \geq \left|N^t(C_+)\right|$ and $\left|N^t(A_-)\right| \geq \left|N^t(C_-)\right|$. In particular, it follows that

$$\left|N^{t-1}(A_-) \cup N^t(A_+)\right| \geq \left|N^t(A_+)\right| \geq \left|N^t(C_+)\right| \tag{2.4}$$

and

$$\left|N^{t-1}(A_+) \cup N^t(A_-)\right| \geq \left|N^t(A_-)\right| \geq \left|N^t(C_-)\right|. \tag{2.5}$$

Hence combining (2.2), (2.3), (2.4) and (2.5) it follows that

$$\left|N^t(A)\right| \geq \left|N^t(C)\right|. \tag{2.6}$$

However, since $|A| = |C|$ and $N^t(A)$ is minimal, the equality must hold in (2.6), and in particular it must hold in (2.4) and (2.5). Thus (2.4) implies that we have $\left|N^t(A_+)\right| = \left|N^t(C_+)\right|$, and since C_+ is the initial segment of the simplicial order, it follows that $N^t(A_+)$ is minimal. Since equality holds in (2.4), we also obtain that $N^{t-1}(A_-) \cup N^t(A_+) = N^t(A_+)$. Hence it follows that $N^{t-1}(A_-) \subseteq N^t(A_+)$. Similarly the fact that equality holds in (2.5) implies that $N^t(A_-)$ is minimal, and that $N^{t-1}(A_+) \subseteq N^t(A_-)$. □

Proposition 11. *Suppose that $A \subseteq Q_n$ is an extremal set for which there exist $t \in Q_n$ and r so that $B(t,r) \subseteq A \subseteq B(t,r+1)$. Then A is isomorphic to an initial segment of the simplicial order.*

Proof. The proof is by induction on n. When $n \leq 2$ it is easy to verify that the claim is true. Suppose that the claim holds for $n-1$, and let $A \subseteq Q_n$ be an extremal set for which there exist $t \in Q_n$ and r so that $B(t,r) \subseteq A \subseteq B(t,r+1)$. By the symmetry of Q_n, we may assume that $t = \emptyset$.

If A satisfies $|A| = f_r$, then Proposition 5 implies that A is an exact Hamming ball of radius r. In particular, A is isomorphic to the initial segment of the simplicial order, and the claim follows in this case. Otherwise, by taking complements if necessary, we may assume that A satisfies $f_r < |A| \leq g_r$. Indeed, this follows from the earlier observation that extremality is preserved under taking complements.

Since $f_r < |A| \leq g_r$, Lemma 6 implies that there exist x, y and z with $y \neq z$ satisfying $B(x,r) \subseteq A \subseteq B(y,r+1) \cap B(z,r+1)$, $d(x,y) \leq 1$ and $d(x,z) \leq 1$. We split the rest of the proof into two cases based on whether we have $|A| = g_r$ or $f_r < |A| < g_r$.

Case 1. The size of A satisfies $|A| = g_r$.

Since $A \subseteq B(y,r+1) \cap B(z,r+1)$, it follows that $B(y^c, n-r-2) \cup B(z^c, n-r-2) \subseteq A^c$. Since $y^c \neq z^c$, Lemma 7 implies that we have $|A^c| \geq g_{n-r-2}$. Since $|A^c| = 2^n - g_r = g_{n-r-2}$, the equality must hold. In particular, we must have $A^c = B(y^c, n-r-2) \cup B(z^c, n-r-2)$, and hence it follows that $A = B(y,r+1) \cap B(z,r+1)$.

If $d(y,z) = 1$, by applying a suitable isometry of Q_n if necessary we may assume that $y = \emptyset$ and $z = \{1\}$. Hence A equals $B(\emptyset, r+1) \cap B(\{1\}, r+1) = X^{(\leq r)} \cup \left(\{1\} + X_1^{(r)}\right)$, which is the initial segment of the simplicial order of size g_r. If $d(y,z) = 2$, note that the only elements s satisfying the condition $A \subseteq B(s,r+1)$ are $s = y$ and $s = z$. However, since $d(y,z) = 2$, it follows that neither of $B(y,r)$ nor $B(z,r)$ is a subset of A. This contradicts the existence of t satisfying $B(t,r) \subseteq A \subseteq B(t,r+1)$. Hence A has to be isomorphic to an initial segment of the simplicial order.

Case 2. The size of A satisfies $f_r < |A| < g_r$.

Recall that $t = \emptyset$ is the element satisfying the condition $B(t,r) \subseteq A \subseteq B(t,r+1)$, and that x, y and z are elements with $y \neq z$ satisfying $B(x,r) \subseteq A \subseteq B(y,r+1) \cap B(z,r+1)$, $d(x,y) \leq 1$

and $d(x,z) \leq 1$. If $x \neq \emptyset$, Lemma 7 implies that we have $|A| \geq |B(\emptyset, r) \cup B(x,r)| \geq g_r$, which contradicts the assumption $|A| < g_r$. Hence we must have $x = \emptyset$.

Since $y \neq z$, it follows that at least one of y and z does not equal $\emptyset$, and hence we may assume that $y \neq \emptyset$. Hence the condition $d(y,x) \leq 1$ implies that $y = \{i\}$ for some i, and by applying a suitable isometry we may assume that $i = 1$. Since $A \subseteq B(y, r+1)$ and $B(\emptyset, r) \subseteq A \subseteq B(\emptyset, r+1)$, it follows that $B(\emptyset, r) \subseteq A \subseteq B(\emptyset, r+1) \cap B(\{1\}, r+1)$. Hence there exists $\mathcal{B} \subseteq X_1^{(r)}$ for which we have $A = X^{(\leq r)} \cup (\{1\} + \mathcal{B})$.

Consider the j-sections in the direction $j = 1$, and for simplicity denote them by A_+ and A_- for the rest of the proof. Since $A = X^{(\leq r)} \cup (\{1\} + \mathcal{B})$, it follows that $A_+ = X_1^{(\leq r-1)} \cup \mathcal{B}$ and $A_- = X_1^{(\leq r)}$. Let C_+ and C_- be the initial segments of the simplicial order of the same sizes as A_+ and A_- on $\mathcal{P}(X_1)$.

Since we have $f_{n-1,r-1} \leq |C_+| < |C_-| = f_{n-1,r}$, it follows that $C_- \subseteq N(C_+)$. Since A is an extremal set, it follows that $N^t(A)$ are minimal for all $t > 0$. Hence Lemma 10 applied for all $t > 0$ implies that $N^t(A_+)$ and $N^t(A_-)$ are minimal for all $t > 0$.

Let D_- and D_+ be the initial segments of the simplicial order of the same sizes as A_-^c and A_+^c on $\mathcal{P}(X_1)$. Since we have $f_{n-1,n-r-2} = |A_-^c| < |A_+^c| \leq f_{n-1,n-r-1}$, it follows that $D_+^c \subseteq N(D_-^c)$. Hence Lemma 10 implies that $N^t(A_+^c)$ and $N^t(A_-^c)$ are minimal for all $t > 0$, and thus both A_+ and A_- are extremal, although note that the extremality of A_- is also evident from the fact that $A_- = X_1^{(\leq r)}$.

The extremality of A_+ implies that $X_1^{(\leq r-1)} \cup \mathcal{B}$ is extremal as a subset of $\mathcal{P}(X_1)$. Since $B_1(\emptyset, r-1) \subseteq X_1^{(\leq r-1)} \cup \mathcal{B} \subseteq B_1(\emptyset, r)$, the inductive hypothesis implies that $X_1^{(\leq r-1)} \cup \mathcal{B}$ is isomorphic to an initial segment of the simplicial order on $\mathcal{P}(X_1)$. Recall that the initial segment of the simplicial order of size $|A_+|$ on $\mathcal{P}(X_1)$ is of the form $X_1^{(\leq r-1)} \cup \mathcal{C}$, where $\mathcal{C} \subseteq X_1^{(r)}$ is the initial segment of the lexicographic order on $X_1^{(r)}$ of size $|\mathcal{B}|$. Hence Lemma 9 implies that one can choose the isomorphism to be of the form ϕ_σ for some $\sigma \in S_n$. Hence we have $\phi_\sigma(\mathcal{B}) = \mathcal{C}$, which implies that $\mathcal{B}$ is isomorphic to an initial segment of the lexicographic order on $X_1^{(r)}$ as subsets of $X_1^{(r)}$. Thus $\{1\} + \mathcal{B}$ is isomorphic to an initial segment of the lexicographic order on $X^{(r+1)}$ as subsets of $X^{(r+1)}$, and hence $A = X^{(\leq r)} \cup (\{1\} + \mathcal{B})$ is isomorphic to an initial segment of the simplicial order. □

Define the *lower shadow* of a set system $\mathcal{A} \subseteq X^{(r)}$ by setting $\partial^- \mathcal{A} = \{b \setminus \{i\} : b \in \mathcal{A}, i \in b\}$, and the iterated lower shadow by setting $\partial^{-t} \mathcal{A} = \partial^- \left(\partial^{-(t-1)} \mathcal{A}\right)$. Similarly define the *upper shadow* of $\mathcal{A} \subseteq X^{(r)}$ by setting $\partial^+ \mathcal{A} = \{b \cup \{i\} : i \in X \setminus b, b \in \mathcal{A}\}$, and the iterated upper shadow by setting $\partial^{+t} \mathcal{A} = \partial^+ \left(\partial^{+(t-1)} \mathcal{A}\right)$. Note that the upper shadow depends on the ground set, which will be X unless otherwise highlighted in the notation. For $\mathcal{A} \subseteq X^{(r)}$ define $\overline{\mathcal{A}} = \{a^c : a \in \mathcal{A}\}$. Note that we have $|\overline{\mathcal{A}}| = |\mathcal{A}|$ and $\overline{\mathcal{A}} \subseteq X^{(n-r)}$. It turns out that the upper and lower shadows can be related to each others via $\partial^+ \mathcal{A} = \overline{\partial^- \overline{\mathcal{A}}}$.

It is natural to ask how one should choose a set system $\mathcal{A} \subseteq X^{(r)}$ of a given size in order to minimise the size of the lower shadow. Note that answering this question would answer the same question concerning the upper shadow by using the fact that $\partial^+ \mathcal{A} = \overline{\partial^- \overline{\mathcal{A}}}$. These questions are answered by the Kruskal-Katona theorem.

Theorem 12. *(Kruskal-Katona theorem [25, 28])*

1. *Let* $\mathcal{A} \subseteq X^{(r)}$*, and let* $\mathcal{B} \subseteq X^{(r)}$ *be the initial segment of the colexicographic order of size* $|\mathcal{A}|$*. Then we have* $|\partial^{-}\mathcal{A}| \geq |\partial^{-}\mathcal{B}|$*.*

2. *Let* $\mathcal{A} \subseteq X^{(r)}$*, and let* $\mathcal{C} \subseteq X^{(r)}$ *be the initial segment of the lexicographic order of size* $|\mathcal{A}|$*. Then we have* $|\partial^{+}\mathcal{A}| \geq |\partial^{+}\mathcal{C}|$*.* □

Note that the lower shadow of an initial segment of the colexicographic order on $X^{(r)}$ is an initial segment of the colexicographic order on $X^{(r-1)}$, and similarly the upper shadow of an initial segment of the lexicographic order on $X^{(r)}$ is an initial segment of the lexicographic order on $X^{(r+1)}$. Hence we can strengthen the conclusion of Theorem 12 by replacing $|\partial^{-}\mathcal{A}| \geq |\partial^{-}\mathcal{B}|$ with $\left|\partial^{-t}\mathcal{A}\right| \geq \left|\partial^{-t}\mathcal{B}\right|$ for all $t > 0$, and by replacing $|\partial^{+}\mathcal{A}| \geq |\partial^{+}\mathcal{C}|$ with $\left|\partial^{+t}\mathcal{A}\right| \geq \left|\partial^{+t}\mathcal{C}\right|$ for all $t > 0$.

Let $\mathcal{A} \subseteq X^{(r)}$, let $\mathcal{B}$ be the initial segment of the colexicographic order on $X^{(r)}$ of size $|\mathcal{A}|$ and let $\mathcal{C}$ be the initial segment of the lexicographic order on $X^{(r)}$ of size $|\mathcal{A}|$. We say that $\partial^{-t}\mathcal{A}$ are *minimal for all* $t > 0$ if we have $\left|\partial^{-t}\mathcal{A}\right| = \left|\partial^{-t}\mathcal{B}\right|$ for all $t > 0$, and we say that $\partial^{+t}\mathcal{A}$ are *minimal for all* $t > 0$ if we have $\left|\partial^{+t}\mathcal{A}\right| = \left|\partial^{+t}\mathcal{C}\right|$ for all $t > 0$. Since the Kruskal-Katona theorem implies that we always have $\left|\partial^{-t}\mathcal{A}\right| \geq \left|\partial^{-t}\mathcal{B}\right|$ and $\left|\partial^{+t}\mathcal{A}\right| \geq \left|\partial^{+t}\mathcal{C}\right|$ for all $t > 0$, in order to verify minimality it suffices to prove that we have $\left|\partial^{-t}\mathcal{B}\right| \geq \left|\partial^{-t}\mathcal{A}\right|$ and $\left|\partial^{+t}\mathcal{C}\right| \geq \left|\partial^{+t}\mathcal{A}\right|$ for all $t > 0$.

Let $\mathcal{C}$ be an initial segment of the lexicographic order on $X^{(r)}$, and let $\mathcal{B} = X^{(r)} \backslash \mathcal{C}$. Then $\mathcal{B}$ is isomorphic to an initial segment of the colexicographic order as subsets of $X^{(r)}$. Hence it follows that $\partial^{+t}\mathcal{C}$ are minimal for all $t > 0$ and $\partial^{-t}\mathcal{B}$ are minimal for all $t > 0$. We now prove that if $\mathcal{A} \subseteq X^{(r)}$ for which $\partial^{+t}\mathcal{A}$ are minimal for all $t > 0$ and $\partial^{-t}\left(X^{(r)} \setminus \mathcal{A}\right)$ are minimal for all $t > 0$, then $\mathcal{A}$ and the initial segment of the lexicographic order on $X^{(r)}$ of size $|\mathcal{A}|$ are isomorphic as subsets of $X^{(r)}$.

Corollary 13. *Let* $\mathcal{A} \subseteq X^{(r)}$ *and set* $\mathcal{D} = X^{(r)} \setminus \mathcal{A}$*. Suppose that* $\partial^{+t}\mathcal{A}$ *and* $\partial^{-t}\mathcal{D}$ *are minimal for all* $t > 0$*. Then* $\mathcal{A}$ *and the initial segment of the lexicographic order of size* $|\mathcal{A}|$ *are isomorphic as subsets of* $X^{(r)}$*.*

Proof. Let $A = X^{(\leq r-1)} \cup \mathcal{A}$, and note that $A^c = X^{(\geq r+1)} \cup \mathcal{D}$. Let $\mathcal{C} \subseteq X^{(r)}$ be the initial segment of the lexicographic order of size $|\mathcal{A}|$. Then $C = X^{(\leq r-1)} \cup \mathcal{C}$ is the initial segment of the simplicial order of size $|A|$. Define $\mathcal{B} = X^{(r)} \setminus \mathcal{C}$, and note that we have $C^c = X^{(\geq r+1)} \cup \mathcal{B}$.

Since $\partial^{+t}\mathcal{A}$ are minimal for all $t > 0$ and $\mathcal{C}$ is the initial segment of the lexicographic order on $X^{(r)}$ of size $|\mathcal{A}|$, it follows that $\left|\partial^{+t}\mathcal{A}\right| = \left|\partial^{+t}\mathcal{C}\right|$ for all $t > 0$. Since $N^t(A) = X^{(\leq r+t-1)} \cup \partial^{+t}\mathcal{A}$ and $N^t(C) = X^{(\leq r+t-1)} \cup \partial^{+t}\mathcal{C}$, it follows that

$$\left|N^t(A)\right| = \left|X^{(\leq r+t-1)}\right| + \left|\partial^{+t}\mathcal{A}\right| = \left|X^{(\leq r+t-1)}\right| + \left|\partial^{+t}\mathcal{C}\right| = \left|N^t(C)\right| \tag{2.7}$$

for all $t > 0$.

Since $\partial^{-t}\mathcal{D}$ are minimal for all $t > 0$ and $\mathcal{B}$ is isomorphic to the initial segment of the colexicographic order of size $|\mathcal{D}|$ as subsets of $X^{(r)}$, it follows that $\left|\partial^{-t}\mathcal{D}\right| = \left|\partial^{-t}\mathcal{B}\right|$ for all $t > 0$. Note that similarly we also have $N^t(A^c) = X^{(\geq r-t+1)} \cup \partial^{-t}\mathcal{D}$ and $N^t(C^c) = X^{(\geq r-t+1)} \cup \partial^{-t}\mathcal{B}$.

Hence it follows that

$$\left|N^t\left(A^c\right)\right| = \left|X^{(\geq r-t+1)}\right| + \left|\partial^{-t}\mathcal{D}\right| = \left|X^{(\geq r-t+1)}\right| + \left|\partial^{-t}\mathcal{B}\right| = \left|N^t\left(C^c\right)\right| \tag{2.8}$$

for all $t > 0$. Hence A is extremal.

Since $B(\emptyset, r-1) \subseteq A \subseteq B(\emptyset, r)$, Proposition 11 implies that A is isomorphic to the initial segment of the simplicial order, and hence A is isomorphic to C. Thus Lemma 9 implies that there exists $\sigma \in S_n$ for which we have $\phi_\sigma(A) = C$. Since ϕ_σ maps the elements of $X^{(r)}$ to the elements of $X^{(r)}$, it follows that $\phi_\sigma(\mathcal{A}) = \mathcal{C}$. Hence $\mathcal{A}$ and $\mathcal{C}$ are isomorphic as subsets of $X^{(r)}$. $\square$

For convenience, we recall the definition of the sets A_i and restate Theorem 2. Define the maps $\pi_i : X^{(r+1)} \to X_i^{(r+1)} \cup X_i^{(r)}$ by setting $\pi_i(x) = x \setminus \{i\}$ for all $x \in X^{(r+1)}$, and for a set system $\mathcal{B} \subseteq X^{(r+1)}$ define $\pi_i(\mathcal{B}) = \{\pi_i(x) : x \in \mathcal{B}\}$. Let s be an integer of the form $s = f_r + k$ for some $0 \leq k \leq \binom{n-1}{r}$. Let $\mathcal{A}$ be the initial segment of the lexicographic order on $X^{(r+1)}$ of size k. Finally, for each i set

$$A_i = X^{(\leq r)} \cup (\{i\} + \pi_i(\mathcal{A})),$$

and note that we have $|A_i| = s$ for all i.

Theorem 2. *Let $A \subseteq Q_n$ be a subset of size s, where $s = f_r + k$ for some r and $k \leq \binom{n-1}{r}$. Let $A_1, \ldots, A_n$ be the sets defined as above for these choices of r and k. Then A is extremal if and only if A is isomorphic to some A_i.*

Proof. Let $0 \leq r \leq n$, $0 \leq k \leq \binom{n-1}{r}$ and set $s = f_r + k$. Let $A_1, \ldots, A_n$ be the subsets of Q_n defined as above for these choices of n, r, k and s. We start by proving that every extremal set $A \subseteq Q_n$ of size s is isomorphic to A_i for some i, and then we prove that each A_i is extremal. The second part is much easier, but since it follows very quickly as a consequence of the first part, we start with the more difficult part.

If $k = 0$, it follows that $|A| = f_r$. Hence Proposition 5 implies that A is an exact Hamming ball of radius r, and so is each of the sets A_i. Hence the result follows in this case. Hence we may assume that $k > 0$.

Since $f_r < |A| \leq g_r$, Lemma 6 implies that there exist x, y, $z \in Q_n$ with $y \neq z$ satisfying $d(x,y) \leq 1$, $d(x,z) \leq 1$ and $B(x,r) \subseteq A \subseteq B(y, r+1) \cap B(z, r+1)$. We may also assume that $x = \emptyset$. If $y = \emptyset$ or $z = \emptyset$, then we have $B(\emptyset, r) \subseteq A \subseteq B(\emptyset, r+1)$ and hence Proposition 11 implies that A is isomorphic to the initial segment of the simplicial order, and thus A is isomorphic to A_1. Hence we may assume that $y \neq \emptyset$ and $z \neq \emptyset$, and thus we have $y = \{i\}$ and $z = \{j\}$ for distinct elements i, $j \in X$.

The condition $B(x,r) \subseteq A \subseteq B(y, r+1) \cap B(z, r+1)$ guaranteed by Lemma 6 can be rewritten as

$$X^{(\leq r)} \subseteq A \subseteq X^{(\leq r)} \cup \left(\{i,j\} + \left(X_{i,j}^{(r-1)} \cup X_{i,j}^{(r)}\right)\right).$$

Hence it follows that $A = X^{(\leq r)} \cup (\{i,j\} + \mathcal{A}_1) \cup (\{i,j\} + \mathcal{A}_2)$ where $\mathcal{A}_t \subseteq X_{i,j}^{(r-2+t)}$ for both $t \in \{1,2\}$. Define a set system $\mathcal{F} \subseteq X^{(r+1)}$ by setting $\mathcal{F} = (\{i,j\} + \mathcal{A}_1) \cup (\{i\} + \mathcal{A}_2)$, and $C \subseteq Q_n$ by setting $C = X^{(\leq r)} \cup \mathcal{F}$. Note that $\mathcal{F}$ can be written as $\mathcal{F} = \{i\} + ((\{j\} + \mathcal{A}_1) \cup \mathcal{A}_2)$.

We start by verifying that it suffices to prove that C is extremal. If C is extremal, C is isomorphic to the initial segment of the simplicial order of size $|A|$ by Proposition 11, and Lemma 9 implies that this isomorphism can be chosen to be of the form ϕ_σ for some $\sigma \in S_n$. Hence $\sigma(\mathcal{F})$ is the initial segment of the lexicographic order on $X^{(r+1)}$ of size $|\mathcal{F}|$.

We will show that such σ always exists with $\sigma(i) = 1$, and hence suppose that we have $\sigma(i) \neq 1$. Since $\mathcal{F} \subseteq \{i\} + X_i^{(r)}$, we also have $\sigma(\mathcal{F}) \subseteq \sigma(i) + X_{\sigma(i)}^{(r)}$, and since $\sigma(\mathcal{F})$ is an initial segment of the lexicographic order on $X^{(r+1)}$ with $|\sigma(\mathcal{F})| \leq \binom{n-1}{r}$, it follows that $\sigma(\mathcal{F}) \subseteq \{1\} + X_1^{(r)}$. Hence every element of $\sigma(\mathcal{F})$ contains $\{1, \sigma(i)\}$ as a subset. Thus σ' given by $\sigma' = (1\,\sigma(i))\,\sigma$ also maps $\mathcal{F}$ to the initial segment of the lexicographic order, and hence we may assume that $\sigma(i) = 1$.

It is easy to see that we have $\phi_\sigma(A) = A_{\sigma(j)}$ whenever σ maps $\mathcal{F}$ to the initial segment of the lexicographic order and satisfies $\sigma(i)$ =1. Hence in order to prove that A is isomorphic to one of $A_1, \ldots, A_n$, it suffices to show that C is extremal.

Let $\partial^+_{i,j}$ denote the upper shadow operator with respect to the ground set $X_{i,j}$. Note that we have

$$\partial^{+t}\mathcal{F} = \left(\{i,j\} + \left(\partial^{+t}_{i,j}\mathcal{A}_1 \cup \partial^{+(t-1)}_{i,j}\mathcal{A}_2\right)\right) \cup \left(\{i\} + \partial^{+t}_{i,j}\mathcal{A}_2\right),$$

and hence it follows that

$$\left|N^t(C)\right| = \left|X^{(\leq r+t)}\right| + \left|\partial^{+t}\mathcal{F}\right| = \left|X^{(\leq r+t)}\right| + \left|\partial^{+t}_{i,j}\mathcal{A}_1 \cup \partial^{+(t-1)}_{i,j}\mathcal{A}_2\right| + \left|\partial^{+t}_{i,j}\mathcal{A}_2\right|. \tag{2.9}$$

On the other hand, we have

$$N^t(A) = X^{(\leq r+t)} \cup \left(\{i,j\} + \left(\partial^{+t}_{i,j}\mathcal{A}_1 \cup \partial^{+(t-1)}_{i,j}\mathcal{A}_2\right)\right) \cup \left(\{i,j\} + \partial^{+t}_{i,j}\mathcal{A}_2\right),$$

and hence it follows that

$$\left|N^t(A)\right| = \left|X^{(\leq r+t)}\right| + \left|\partial^{+t}_{i,j}\mathcal{A}_1 \cup \partial^{+(t-1)}_{i,j}\mathcal{A}_2\right| + \left|\partial^{+t}_{i,j}\mathcal{A}_2\right|. \tag{2.10}$$

Combining (2.9) with (2.10) we obtain that $\left|N^t(A)\right| = \left|N^t(C)\right|$ for all $t > 0$.

Let $\mathcal{B}_1 = X^{(r-1)}_{i,j} \setminus \mathcal{A}_1$ and $\mathcal{B}_2 = X^{(r)}_{i,j} \setminus \mathcal{A}_2$. Let $S_r = \left\{x \in X^{(r)} : \{i,j\} \not\subseteq x\right\}$, i.e. $S_r = X^{(r)} \setminus \left(\{i,j\} + X^{(r-2)}_{i,j}\right)$. It is easy to verify that we have

$$A^c = X^{(\geq r+3)} \cup (\{i,j\} + (\mathcal{B}_1 \cup \mathcal{B}_2)) \cup S_{r+1} \cup S_{r+2}. \tag{2.11}$$

It is easy to see that $\partial^{-t}S_m = S_{m-t}$ for any $m > t$. Also note that we have

$$\{i,j\} + \partial^{-t}\mathcal{B}_m \subseteq \partial^{-t}(\{i,j\} + \mathcal{B}_m) \subseteq \left(\{i,j\} + \partial^{-t}\mathcal{B}_m\right) \cup S_{r+m-t}$$

for both $m \in \{1,2\}$. Hence (2.11) implies that we have

$$N^t(A^c) = X^{(\geq r-t+3)} \cup \left(\{i,j\} + \left(\partial^{-t}\mathcal{B}_2 \cup \partial^{-(t-1)}\mathcal{B}_1\right)\right) \cup \left(\{i,j\} + \partial^{-t}\mathcal{B}_1\right) \cup S_{r-t+1} \cup S_{r-t+2}. \tag{2.12}$$

Note that $X^{(\geq r-t+3)}$, $\{i,j\} + \left(\partial^{-t}\mathcal{B}_2 \cup \partial^{-(t-1)}\mathcal{B}_1\right)$, S_{r-t+2} , $\{i,j\} + \partial^{-t}\mathcal{B}_1$ and S_{r-t+1} are pairwise disjoint subsets of Q_n. Indeed, the first one has only sets containing at least $r-t+3$ elements, whereas the next two contain only sets of size $r-t+2$ and the last two contain only sets of size $r-t+1$. Also, S_{r-t+2} and S_{r-t+1} do not have any sets contained in $\{i,j\}+\mathcal{P}(X_{i,j})$, whereas the second and fourth set systems are contained in $\{i,j\}+\mathcal{P}(X_{i,j})$.

Combining these observations together with (2.12), we obtain that

$$\left|N^t(A^c)\right| = \left|X^{(\geq r-t+3)}\right| + \left|\partial^{-t}\mathcal{B}_2 \cup \partial^{-(t-1)}\mathcal{B}_1\right| + \left|\partial^{-t}\mathcal{B}_1\right| + \left|S_{r-t+1}\right| + \left|S_{r-t+2}\right|. \tag{2.13}$$

Note that we have $\left|X_{i,j}^{(r-1)}\right| + \left|X_{i,j}^{(r)}\right| = \binom{n-2}{r-1} + \binom{n-2}{r} = \binom{n-1}{r}$ for all r. Since $S_r = X^{(r)} \setminus \left(\{i,j\} + X_{i,j}^{(r-2)}\right)$, it follows that

$$\begin{aligned}\left|S_{r-t+1}\right| + \left|S_{r-t+2}\right| &= \left|X^{(r-t+1)}\right| + \left|X^{(r-t+2)}\right| - \left(\left|X_{i,j}^{(r-t-1)}\right| + \left|X_{i,j}^{(r-t)}\right|\right) \\ &= \left|X^{(r-t+2)}\right| + \binom{n}{r-t+1} - \binom{n-1}{r-t} = \left|X^{(r-t+2)}\right| + \binom{n-1}{r-t+1}.\end{aligned} \tag{2.14}$$

Combining (2.13) and (2.14) together with the fact that $\left|X_i^{(r-t+1)}\right| = \binom{n-1}{r-t+1}$, it follows that

$$\left|N^t(A^c)\right| = \left|X^{(\geq r-t+2)}\right| + \left|X_i^{(r-t+1)}\right| + \left|\partial^{-t}\mathcal{B}_2 \cup \partial^{-(t-1)}\mathcal{B}_1\right| + \left|\partial^{-t}\mathcal{B}_1\right|. \tag{2.15}$$

Let $\mathcal{G} = X^{(r+1)} \setminus \mathcal{F}$. Since $\mathcal{F} \subseteq X^{(r+1)}$ and $C = X^{(\leq r)} \cup \mathcal{F}$, it follows that $C^c = X^{(\geq r+2)} \cup \mathcal{G}$. Hence we have

$$\left|N^t(C^c)\right| = \left|X^{(\geq r-t+2)}\right| + \left|\partial^{-t}\mathcal{G}\right| \tag{2.16}$$

for all $t > 0$. We now find an expression for $\left|\partial^{-t}\mathcal{G}\right|$ in terms of $\mathcal{B}_1$ and $\mathcal{B}_2$.

Recall that $\mathcal{F}$ can be written as $\mathcal{F} = \{i\} + ((\{j\} + \mathcal{A}_1) \cup \mathcal{A}_2)$. Hence it follows that $\mathcal{G} = X_i^{(r+1)} \cup (\{i,j\} + \mathcal{B}_1) \cup (\{i\} + \mathcal{B}_2)$, and it is easy to verify that we have

$$\partial^{-t}\mathcal{G} = X_i^{(r-t+1)} \cup \left(\{i,j\} + \partial^{-t}\mathcal{B}_1\right) \cup \left(\{i\} + \left(\partial^{-(t-1)}\mathcal{B}_1 \cup \partial^{-t}\mathcal{B}_2\right)\right). \tag{2.17}$$

Note that the sets $X_i^{(r-t+1)}$, $\{i,j\} + \partial^{-t}\mathcal{B}_1$ and $\{i\} + \left(\partial^{-(t-1)}\mathcal{B}_1 \cup \partial^{-t}\mathcal{B}_2\right)$ are pairwise disjoint. Indeed, this follows by noting that the sets in $X_i^{(r-t+1)}$ do not contain i, the sets in $\{i,j\}+\partial^{-t}\mathcal{B}_1$ contain both i and j, and the sets in $\{i\} + \left(\partial^{-(t-1)}\mathcal{B}_1 \cup \partial^{-t}\mathcal{B}_2\right)$ contain i but not j. Hence (2.16) and (2.17) imply that

$$\left|N^t(C^c)\right| = \left|X^{(\geq r-t+2)}\right| + \left|X_i^{(r-t+1)}\right| + \left|\partial^{-t}\mathcal{B}_1\right| + \left|\partial^{-(t-1)}\mathcal{B}_1 \cup \partial^{-t}\mathcal{B}_2\right|. \tag{2.18}$$

In particular, (2.15) and (2.18) imply that we have $\left|N^t(C^c)\right| = \left|N^t(A^c)\right|$ for all $t > 0$. Hence C is extremal, and thus by our earlier observations it follows that A is isomorphic to A_i for some i.

Conversely, suppose that $A = A_i$ for some i. Define the set systems $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{B}_1$, $\mathcal{B}_2$, $\mathcal{F}$ and $\mathcal{G}$ as before. By the construction of the set A_i, it follows that $\mathcal{F}$ is isomorphic to an initial segment of the lexicographic order as subsets of $X^{(r+1)}$. As before, define $C = X^{(\leq r)} \cup \mathcal{F}$. Then C is isomorphic to an initial segment of the simplicial order, so C is extremal. Then (2.9) and (2.10) imply that $\left|N^t(A)\right| = \left|N^t(C)\right|$ for all $t > 0$, and similarly (2.15) and (2.18) imply that

$\left|N^t(A^c)\right| = \left|N^t(C^c)\right|$ for all $t > 0$. Since A and C are subsets of the same size and C is extremal, it follows that A is an extremal set as well, which completes the proof. $\square$

Note that if $s = f_r$ for some r, then Proposition 5 implies that if A is an extremal set of size s, then A is isomorphic to the initial segment of the simplicial order. We now show that for all other sizes s there exists a non-trivial extremal set of size s.

Corollary 14. *For all n and for all $s \notin \{f_0, \dots, f_n\}$ there exists an extremal set $A \subseteq Q_n$ of size s which is not isomorphic to the initial segment of the simplicial order.*

Proof. If $s = g_r$ for some r, we may take $A = B_r$ where B_r is the set defined in Section 2.2. Hence by taking complements if necessary, we may assume that $f_r < s < g_r$ for some r. Let $k = s - f_r$, and let $\mathcal{D}$ be the initial segment of the lexicographic order of size k on $X^{(r+1)}$. Let $\mathcal{D}_{i,+}$ and $\mathcal{D}_{i,-}$ be the i-sections of $\mathcal{D}$. Since $\mathcal{D} \neq \emptyset$, there exists i for which $\mathcal{D}_{i,-} \neq \emptyset$.

Consider the set A_i for this particular choice of i, and note that A_i is given by $A_i = X^{(\leq r)} \cup (\{i\} + \mathcal{D}_{i,+}) \cup (\{i\} + \mathcal{D}_{i,-})$. Since $\mathcal{D}_{i,-} \neq \emptyset$, it follows that $A_i \cap X^{(r+2)} \neq \emptyset$. Since $s < g_r$, it follows that $B(\emptyset, r)$ is the unique exact Hamming ball of radius r contained in A_i. Indeed, if $B(x,r) \subseteq A$ for some $x \neq \emptyset$, then Lemma 7 implies that $|A| \geq |B(\emptyset, r) \cup B(x,r)| \geq g_r$, which contradicts the fact that $|A| < g_r$. Thus at least one of the conditions $B(x,r) \subseteq A_i$ and $A_i \subseteq B(x, r+1)$ is violated for any $x \in Q_n$. The first condition is violated for any $x \neq \emptyset$, and the second one is violated for $x = \emptyset$ since $A_i \cap X^{(r+2)} \neq \emptyset$. In particular, it follows that A_i is not a Hamming ball, and thus A_i cannot be isomorphic to an initial segment of the simplicial order. $\square$

It is natural to ask when the sets A_i and A_j are isomorphic as subsets of Q_n. Let $\mathcal{D}$ be the initial segment chosen so that the sets A_i are obtained from $\mathcal{D}$. If $\sigma = (ij) \in S_n$ satisfies the condition $\phi_\sigma(\mathcal{D}) = \mathcal{D}$, then A_i and A_j are certainly isomorphic. The aim of the following lemma is to prove that this is the only case when such an isomorphism occurs.

Lemma 15. *Let $0 \leq r \leq n$, $0 \leq k \leq \binom{n-1}{r}$, $s = f_r + k$, and let A_i be the sets defined before Theorem 2 for these choices of n and s. Let $\mathcal{D}$ be the initial segment of the lexicographic order of size k on $X^{(r+1)}$. Then A_i and A_j are isomorphic if and only if $\phi_\sigma(\mathcal{D}) = \mathcal{D}$ for $\sigma = (ij)$.*

We say that $\mathcal{D} \subseteq X^{(r)}$ is *left compressed* if for all $i < j$ and $a \in \mathcal{D}$ we have

$$(j \in a \text{ and } i \notin a) \Rightarrow (a \setminus \{j\}) \cup \{i\} \in \mathcal{D}.$$

As an example, if $\mathcal{D}$ is an initial segment of the lexicographic order or the colexicographic order, then $\mathcal{D}$ is left compressed.

Proof. If $s = f_r$, then $\mathcal{D} = \emptyset$ and we have $A_i = B(\emptyset, r)$ for all i. Hence for every i and j the sets A_i and A_j are isomorphic. We also have $\phi_\sigma(\mathcal{D}) = \mathcal{D}$ for all $\sigma = (ij)$, and therefore the claim follows.

If $s = g_r$, we have $\mathcal{D} = \{1\} + X_1^{(r)}$. Hence we have $\phi_\sigma(\mathcal{D}) = \mathcal{D}$ for $\sigma = (ij)$ if and only if $1 \notin \{i, j\}$. On the other hand, note that A_1 is isomorphic to the initial segment of the simplicial order, and for $j \geq 2$ each A_j is isomorphic to the set B_r defined in Section 2.2. Hence the claim is true when $s = f_r$ or $s = g_r$, and from now, on we may assume that $f_r < s < g_r$.

If $\phi_\sigma(\mathcal{D}) = \mathcal{D}$ for $\sigma = (ij)$, it follows that $\phi_\sigma(\mathcal{D}_{i,-}) = \mathcal{D}_{j,-}$ and $\phi_\sigma(\mathcal{D}_{i,+}) = \mathcal{D}_{j,+}$. Hence it follows that $\phi_\sigma(A_i) = A_j$, and thus A_i and A_j are isomorphic.

Suppose that A_i and A_j are isomorphic and we have $i < j$. Since $s < g_r$, Lemma 7 implies that $B(\emptyset, r)$ is the unique exact Hamming ball of radius r contained in each set A_t. Hence if g is an isometry mapping A_i to A_j, then it follows that $g(\emptyset) = \emptyset$. Thus Lemma 8 implies that we have $g = \phi_\sigma$ for some $\sigma \in S_n$. Hence it follows that $A_i \cap X^{(m)}$ is mapped to $A_j \cap X^{(m)}$ for all m, and in particular it follows that $|\mathcal{D}_{i,-}| = |\mathcal{D}_{j,-}|$ and $|\mathcal{D}_{i,+}| = |\mathcal{D}_{j,+}|$. In order to prove that $\phi_{(ij)}(\mathcal{D}) = \mathcal{D}$, we need to prove the following two claims

1. For all $a \in \mathcal{D}$ with $i \notin a$ and $j \in a$ we have $(a \setminus \{j\}) \cup \{i\} \in \mathcal{D}$.
2. For all $a \in \mathcal{D}$ with $i \in a$ and $j \notin a$ we have $(a \setminus \{i\}) \cup \{j\} \in \mathcal{D}$.

Since $i < j$ and $\mathcal{D}$ is an initial segment of the lexicographic order, the first claim follows immediately from the fact that $\mathcal{D}$ is left-compressed.

Define $\mathcal{T}_i = \{a \in \mathcal{D} : i \in a, j \notin a\}$, $\mathcal{T}_j = \{a \in \mathcal{D} : j \in a, i \notin a\}$ and $\mathcal{T}_{i,j} = \{a \in \mathcal{D} : i \in a, j \in a\}$. Note that we have $\mathcal{D}_{i,+} = \mathcal{T}_i \cup \mathcal{T}_{i,j}$ and $\mathcal{D}_{j,+} = \mathcal{T}_j \cup \mathcal{T}_{i,j}$. Since $\mathcal{T}_i$, $\mathcal{T}_j$ and $\mathcal{T}_{i,j}$ are disjoint sets, the condition $|\mathcal{D}_{i,+}| = |\mathcal{D}_{j,+}|$ is equivalent to the condition $|\mathcal{T}_i| = |\mathcal{T}_j|$.

Since $\mathcal{D}$ is left compressed, it follows that the map $a \to (a \setminus \{j\}) \cup \{i\}$ is a well-defined injection from $\mathcal{T}_j$ to $\mathcal{T}_i$. Since $|\mathcal{T}_i| = |\mathcal{T}_j|$, this map is a bijection, and its inverse is given by $b \to (b \setminus \{i\}) \cup \{j\}$. Hence the second claim must be true as well, and thus we have $\phi_{(ij)}(\mathcal{D}) = \mathcal{D}$. $\square$

It is natural to ask whether for all t there exist n and s for which there are t pairwise non-isomorphic extremal sets $B_1, \ldots, B_t$ of size s on Q_n. One can use Lemma 15 to conclude that this turns out to be true. Indeed, for a given t it suffices to find n, r and an initial segment $\mathcal{A}$ of the lexicographic order on $X^{(r)}$ with $|\mathcal{A}| \leq \binom{n-1}{r-1}$ for which the sizes $|\mathcal{A}_{j_i,+}|$ are distinct for some $j_1, \ldots, j_t \in X$.

Recall that if $\mathcal{A}$ is isomorphic to an initial segment of the colexicographic order, then $X^{(r)} \setminus \mathcal{A}$ is isomorphic to an initial segment of the lexicographic order. Also note that $|\mathcal{A}_{j,+}| + \left|\left(X^{(r)} \setminus \mathcal{A}\right)_{j,+}\right| = \binom{n-1}{r-1}$ for all $j \in X$. Hence it suffices to find an initial segment $\mathcal{A}$ of the colexicographic order on $X^{(r)}$ with $|\mathcal{A}| \geq \binom{n}{r} - \binom{n-1}{r-1} = \binom{n-1}{r}$ for which the sizes of $\mathcal{A}_{j_i,+}$ are distinct for some $j_1, \ldots, j_t \in X$.

For all t, we inductively construct a proper subset $\mathcal{A}_t$ of $\{1, \ldots, 2t\}^{(t)}$ satisfying $\left|(\mathcal{A}_t)_{2i,+}\right| < \left|(\mathcal{A}_t)_{2j,+}\right|$ for all $1 \leq j < i \leq t$ and with $|\mathcal{A}_t| \geq \binom{2t-1}{t}$. As the base case $t = 1$ we can certainly take $\mathcal{A}_1 = \{\{1\}\}$.

Now suppose that $t \geq 2$ and that the claim holds for $t-1$, and consider the set system $\mathcal{A}_{t+1} = \{1, \ldots, 2t+1\}^{(t+1)} \cup (\mathcal{A}_t + \{2t+2\})$. First of all, $\mathcal{A}_{t+1}$ is certainly an initial segment of the colexicographic order on $\{1, \ldots, 2t+2\}^{(t+1)}$ as $\mathcal{A}_t$ is an initial segment of the colexicographic order on $\{1, \ldots, 2t\}^{(t)}$, and we also have $|\mathcal{A}_{t+1}| \geq \binom{2t+1}{t+1}$.

Note that for all $1 \leq i \leq t$ we have $(\mathcal{A}_{t+1})_{2i,+} = \binom{2t}{t} + \left|(\mathcal{A}_t)_{2i,+}\right| \geq \binom{2t}{t}$. Hence the inductive hypothesis implies that we have $\left|(\mathcal{A}_{t+1})_{2i,+}\right| < \left|(\mathcal{A}_{t+1})_{2j,+}\right|$ for all $1 \leq j < i \leq t$. On the other hand, note that $\left|(\mathcal{A}_{t+1})_{2t+2,+}\right| = |\mathcal{A}_t| < \binom{2t}{t}$, as $\mathcal{A}_t$ is a proper subset of $\{1, \ldots, 2t\}^{(t)}$. Hence we have $\left|(\mathcal{A}_{t+1})_{2t+2,+}\right| < \left|(\mathcal{A}_{t+1})_{2i,+}\right|$ for all $1 \leq i \leq t+1$. Hence for $n = 2t$ there exists a suitable s with $f_{t-1} \leq s \leq f_{t-1} + \binom{2t-1}{t-1}$ for which the sets $A_2, A_4, \ldots, A_{2t}$ are pairwise non-isomorphic.

2.4 The weak version

Suppose we weaken the notion of extremality so that we only require $N(A)$ and $N(A^c)$ to have minimal size among the sets $A \subseteq Q_n$ of a given size. The aim of this section is to prove that this weaker condition actually implies that we have $\left|N^t(B)\right| \geq \left|N^t(A)\right|$ and $\left|N^t(B^c)\right| \geq \left|N^t(A^c)\right|$ for all $B \subseteq Q_n$ of size $|A|$ and for all $t > 0$, i.e. that such a set A is extremal. From now on, we say that A is *weakly extremal* if for all $B \subseteq Q_n$ of size $|A|$ we have $|N(B)| \geq |N(A)|$ and $|N(B^c)| \geq |N(A^c)|$.

Füredi and Griggs [20] proved that if $\mathcal{D} \subseteq X^{(r)}$ is a set system for which the size of $\partial^- \mathcal{D}$ is minimal among the subsets of $X^{(r)}$ of a given size, then $\partial^{-t}\mathcal{D}$ are minimal for all $t > 0$. Our aim is to prove that a similar conclusion holds for subsets of Q_n as well. That is, if $A \subseteq Q_n$ is a subset for which the size of $N(A)$ is minimal among the subsets of Q_n of size $|A|$, then the size of $N^t(A)$ is also minimal for all $t > 0$ among the subsets of Q_n of size $|A|$. This result immediately implies that the notions of weak extremality and extremality coincide. Hence the main task is to prove the following theorem.

Theorem 4. *Let $A \subseteq Q_n$ be a subset for which every $B \subseteq Q_n$ with $|B| = |A|$ satisfies $|N(B)| \geq |N(A)|$. Then for all $t > 0$ and $B \subseteq Q_n$ with $|B| = |A|$ we have $\left|N^t(B)\right| \geq \left|N^t(A)\right|$.*

Let $A \subseteq Q_n$ be a set satisfying the conditions of Theorem 4. A natural way to prove Theorem 4 is to use induction on the dimension n. Let $A_{i,+}$ and $A_{i,-}$ be the i-sections of A for some direction i, and let $C_{i,+}$ and $C_{i,-}$ be the initial segments of the simplicial order on $\mathcal{P}(X_i)$ of sizes $|A_{i,+}|$ and $|A_{i,-}|$ respectively. Without loss of generality assume that we have $|A_{i,+}| \leq |A_{i,-}|$. If the sets $C_{i,+}$ and $C_{i,-}$ satisfy the conditions of Lemma 10 when $t = 1$, i.e. we have $C_{i,-} \subseteq N(C_{i,+})$, then applying Lemma 10 with $t = 1$ implies that both $N(A_{i,+})$ and $N(A_{i,-})$ are minimal, and hence the claim follows from a relatively straightforward application of (2.2). However, we cannot apply this argument if the conditions of Lemma 10 are not satisfied, i.e. if $N(C_{i,+})$ is a proper subset of $C_{i,-}$. It is easy to verify that this may happen even when A is an initial segment of the simplicial order. For example, if $A = B(\emptyset, r) \cup \{\{1, \ldots, r+1\}\}$, then $N(A_{i,+})$ is a proper subset of $A_{i,-}$ for every $i > r+1$.

Note that in the previous example there were some directions i for which the sets $C_{i,-}$ and $C_{i,+}$ satisfy the conditions of Lemma 10, for example any $i \leq r+1$ would work. Conveniently, it turns out that for any set $A \subseteq Q_n$ satisfying the conditions of Theorem 4 there is always a direction i for which both $C_{i,+}$ and $C_{i,-}$ satisfy the conditions of Lemma 10.

We start with some preliminary results which are Lemmas 16, 17 and 18. In a sense, Lemmas 17 and 18 are just necessary tools for subsequent results, and the proofs are mostly calculational, while Lemma 16 is a direct consequence of a result proved in Section 3.3. These statements are in flavour similar to Harper's theorem and the Kruskal-Katona theorem, but it seems that there is no straightforward way of deducing them directly from Harper's theorem or the Kruskal-Katona theorem.

Armed with those, we move on to Lemma 19 in which we prove that there exists a direction i for which $C_{i,+}$ and $C_{i,-}$ satisfy the conditions of Lemma 10. Given Lemma 19, the proof of Theorem 4 follows almost immediately, and we deduce that weak extremality implies extremality in Theorem 3.

Lemma 16. *Let $\mathcal{A}$, $\mathcal{B} \subseteq X^{(r)}$ be initial segments of the colexicographic order with $|\mathcal{A}| + |\mathcal{B}| \leq \binom{n}{r}$, and recall that $\mathcal{A}_+ = \{a \setminus \{1\} : 1 \in a, a \in \mathcal{A}\}$. Let $\mathcal{C} \subseteq X^{(r)}$ be the initial segment of the colexicographic order of size $|\mathcal{A}| + |\mathcal{B}|$ on $X^{(r)}$. Then we have $|\mathcal{A}_+| + |\mathcal{B}_+| \geq |\mathcal{C}_+|$.*

Proof. Let $\mathcal{I}$ and $\mathcal{J}$ be initial segments of the colexicographic order on $X^{(r)}$ with $|\mathcal{I}| \geq |\mathcal{J}|$, and let $\mathcal{K}$ be the initial segment of the colexicographic order on $X^{(r)}$ of size $|\mathcal{I} \setminus \mathcal{J}|$. Since $|\mathcal{K}| + |\mathcal{J}| = |\mathcal{I}|$, Lemma 31 from Section 3.3 states that we have $|\mathcal{K}_+| \geq \left|(\mathcal{I} \setminus \mathcal{J})_+\right|$.

We apply this Lemma for $\mathcal{I} = \mathcal{C}$ and $\mathcal{J} = \mathcal{A}$, and these conditions imply that we have $\mathcal{K} = \mathcal{B}$. Since $\mathcal{A} \subseteq \mathcal{C}$, it follows directly that we have $|\mathcal{C}_+| = \left|(\mathcal{C} \setminus \mathcal{A})_+\right| + |\mathcal{A}_+|$. Since Lemma 31 implies that $\left|(\mathcal{C} \setminus \mathcal{A})_+\right| \leq |\mathcal{B}_+|$, it follows that $|\mathcal{A}_+| + |\mathcal{B}_+| \geq |\mathcal{C}_+|$. □

Lemma 17.

1. *Let $\mathcal{A}$, $\mathcal{B} \subseteq X^{(r)}$ be non-empty initial segments of the lexicographic order satisfying $|\mathcal{A}| + |\mathcal{B}| \leq \binom{n}{r}$ and let $\mathcal{C}$ be the initial segment of the lexicographic order of size $|\mathcal{A}| + |\mathcal{B}|$ on $X^{(r)}$. Then we have $|\partial^+\mathcal{A}| + |\partial^+\mathcal{B}| > |\partial^+\mathcal{C}|$.*

2. *Let $\mathcal{A}$, $\mathcal{B} \subseteq X^{(r)}$ be non-empty initial segments of the colexicographic order satisfying $|\mathcal{A}| + |\mathcal{B}| \leq \binom{n}{r}$ and let $\mathcal{C}$ be the initial segment of the colexicographic order of size $|\mathcal{A}| + |\mathcal{B}|$ on $X^{(r)}$. Then we have $|\partial^-\mathcal{A}| + |\partial^-\mathcal{B}| > |\partial^-\mathcal{C}|$.*

Note that the Kruskal-Katona theorem certainly implies that we have $|\partial^+\mathcal{A}| + |\partial^+\mathcal{B}| \geq |\partial^+\mathcal{C}|$ and $|\partial^-\mathcal{A}| + |\partial^-\mathcal{B}| \geq |\partial^-\mathcal{C}|$. However, there does not seem to be a straightforward way to directly conclude the strict inequalities from the Kruskal-Katona theorem.

Proof. Let $\mathcal{A}$ be an initial segment of the lexicographic order and let $\overline{\mathcal{A}} = \{a^c : a \in \mathcal{A}\}$. Recall that we have $|\mathcal{A}| = |\overline{\mathcal{A}}|$, $\partial^-\overline{\mathcal{A}} = \overline{\partial^+\mathcal{A}}$ and that $\overline{\mathcal{A}}$ is isomorphic to an initial segment of the colexicographic order. Hence these claims are equivalent, so it suffices to prove the second one.

The proof is by induction on r. When $r = 1$, we have $|\partial^-\mathcal{A}| = |\partial^-\mathcal{B}| = |\partial^-\mathcal{C}| = 1$ as each of the set systems $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ is non-empty. Hence we may assume that $r > 1$, and that the claim holds for $r - 1$.

Recall that $\mathcal{A}_+ = \{a \setminus \{1\} : 1 \in a, a \in \mathcal{A}\}$ and $\mathcal{A}_- = \{a : 1 \notin a, a \in \mathcal{A}\}$. Since 1 is the smallest element in X and $\mathcal{A}$ is an initial segment of the colexicographic order and hence left compressed, it follows that $\partial^-\mathcal{A}_- \subseteq \mathcal{A}_+$. Since $\mathcal{A} = (\{1\} + \mathcal{A}_+) \cup \mathcal{A}_-$, it follows that

$$\partial^-\mathcal{A} = \left(\{1\} + \partial^-\mathcal{A}_+\right) \cup \left(\mathcal{A}_+ \cup \partial^-\mathcal{A}_-\right) = \left(\{1\} + \partial^-\mathcal{A}_+\right) \cup \mathcal{A}_+,$$

and hence we have

$$|\mathcal{A}| = |\mathcal{A}_+| + \left|\partial^-\mathcal{A}_+\right|. \tag{2.19}$$

Similarly (2.19) holds for $\mathcal{B}$ and $\mathcal{C}$ as well. Since $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are non-empty initial segments of the colexicographic order, it follows that $\mathcal{A}_+$, $\mathcal{B}_+$ and $\mathcal{C}_+$ are non-empty initial segments of the colexicographic order on $X_1^{(r-1)}$. Thus Lemma 16 implies that we have

$$|\mathcal{A}_+| + |\mathcal{B}_+| \geq |\mathcal{C}_+| . \tag{2.20}$$

Since the lower shadow of a smaller initial segment cannot have larger size than the lower shadow of a larger initial segment, the inductive hypothesis implies that we have

$$\left|\partial^- \mathcal{A}_+\right| + \left|\partial^- \mathcal{B}_+\right| > \left|\partial^- \mathcal{C}_+\right| . \tag{2.21}$$

Combining (2.19), (2.20) and (2.21) we obtain that

$$\begin{aligned}
\left|\partial^- \mathcal{A}\right| + \left|\partial^- \mathcal{B}\right| &= \left|\partial^- \mathcal{A}_+\right| + \left|\partial^- \mathcal{B}_+\right| + |\mathcal{A}_+| + |\mathcal{B}_+| \\
&\geq \left|\partial^- \mathcal{A}_+\right| + \left|\partial^- \mathcal{B}_+\right| + |\mathcal{C}_+| > \left|\partial^- \mathcal{C}_+\right| + |\mathcal{C}_+| = \left|\partial^- \mathcal{C}\right| ,
\end{aligned}$$

which completes the proof of Lemma 17. □

Lemma 18. *Let $A \subseteq Q_n$ and suppose that for all $B \subseteq Q_n$ with $|B| = |A|$ we have $|N(B)| \geq |N(A)|$. Let D_+ be the initial segment of the simplicial order on Q_{n-1} of the largest size for which the initial segment D_- of the simplicial order on Q_{n-1} of size $|A| - |D_+|$ satisfies $N(D_+) \subseteq D_-$. Then for any direction i we have $|A_{i,+}| \geq |D_+|$. Furthermore, we must have $|A_{i,+}| = |D_+|$ whenever $|N(A_{i,+})| \leq |A_{i,-}|$.*

Proof. For simplicity, denote the i-sections by A_+ and A_-. Let D_+ and D_- be defined as in the statement, and let $D = D_- \cup (\{n\} + D_+)$.

Suppose that we have $|A_+| < |D_+|$, and note that then we must also have $|A_-| > |D_-|$. Since D_- is an initial segment of the simplicial order, Harper's theorem implies that $|N(A_-)| \geq |N(D_-)|$. Since $N(D_+) \subseteq D_-$, and hence $D_+ \subseteq N^2(D_+) \subseteq N(D_-)$, (2.2) implies that we have

$$|N(D)| = |N(D_-)| + |D_-| .$$

On the other hand, (2.2) implies that

$$|N(A)| \geq |N(A_-)| + |A_-| > |N(D_-)| + |D_-| = |N(D)|$$

as $|A_-| > |D_-|$ and $|N(A_-)| \geq |N(D_-)|$. This contradicts the fact that $|N(D)| \geq |N(A)|$. Hence we must have $|A_+| \geq |D_+|$.

In order to prove the second part, suppose that we have $|N(A_+)| \leq |A_-|$. Let C_+ and C_- be the initial segments of the simplicial order on $\mathcal{P}(X_i)$ of sizes $|A_+|$ and $|A_-|$. Harper's theorem implies that we have $|N(A_+)| \geq |N(C_+)|$, and since $|A_-| = |C_-|$ it follows that $|N(C_+)| \leq |C_-|$. Since C_+ and C_- are initial segments of the simplicial order on $\mathcal{P}(X_i)$, it follows that $N(C_+) \subseteq C_-$, so C_+ and C_- satisfy the conditions required from D_+ and D_-. Thus by the maximality condition it follows that $|C_+| \leq |D_+|$. By the first part, we know that $|A_+| \geq |D_+|$. Since $|A_+| = |C_+|$, these two inequalities imply that $|A_+| = |D_+|$. □

Lemma 19. *Let $A \subseteq Q_n$ be a set satisfying $|A| \neq f_r$ for all r, and so that for all $B \subseteq Q_n$ with $|B| = |A|$ we have $|N(B)| \geq |N(A)|$. Let D_+ and D_- be defined as in the statement of Lemma*

18 for A. Then there exists a direction i for which we have $\min(|A_{i,+}|, |A_{i,-}|) > |D_+|$.

Proof. Let D_+ and D_- be defined as in the statement of Lemma 18. For $I \subseteq X$, note that the set $A_I \subseteq Q_n$ defined by $A_I = \{x \Delta I : x \in A\}$ is isomorphic to A. By taking $I = \{i : |A_{i,+}| > |A_{i,-}|\}$, it is easy to see that we may assume that the condition $|A_{i,+}| \leq |A_{i,-}|$ holds for all i. Hence it suffices to show that under this condition there exists i satisfying $|A_{i,+}| > |D_+|$.

Suppose that the claim is false. Then, by Lemma 18, we must have $|A_{i,+}| = |D_+|$ for all i. Let r be chosen so that $f_r < |A| < f_{r+1}$, and note that such r exists as $|A| \neq f_s$ for all s. Since $N(B_{n-1}(\emptyset, r-1)) = B_{n-1}(\emptyset, r)$ and $f_{n-1,r-1} + f_{n-1,r} = f_{n,r}$, it follows from the maximality assumption that we have $|D_+| \geq f_{n-1,r-1}$. On the other hand, since $f_{n-1,r} + f_{n-1,r+1} = f_{n,r+1}$, the condition $N(D_+) \subseteq D_-$ will be violated if $|D_+| \geq f_{n-1,r}$. These two conditions imply that we have

$$f_{n-1,r-1} \leq |D_+| < f_{n-1,r}. \tag{2.22}$$

Hence there exists a set system $\mathcal{D} \subseteq X_n^{(r)}$ so that $D_+ = X_n^{(\leq r-1)} \cup \mathcal{D}$. Since $N(D_+) \subseteq D_-$, it follows that $X_n^{(\leq r)} \cup \partial_n^+ \mathcal{D} \subseteq D_-$.

Let C be the initial segment of the simplicial order of size $|A|$ on Q_n. Since $f_r < |A| < f_{r+1}$, it follows that $C = X^{(\leq r)} \cup \mathcal{C}$ for some $\mathcal{C} \subseteq X^{(r+1)}$. Combining the conditions $|A| = |D_+| + |D_-|$, $D_+ = X_n^{(\leq r-1)} \cup \mathcal{D}$ and $X_n^{(\leq r)} \cup \partial_n^+ \mathcal{D} \subseteq D_-$, it follows that $|\mathcal{C}| \geq |\mathcal{D}| + |\partial_n^+ \mathcal{D}|$.

For $B \subseteq Q_n$ define $f(B) = \sum_{x \in B} |x|$. It is easy to see that among the sets $B \subseteq Q_n$ of a given size, $f(B)$ attains its minimum value when B is taken to be the initial segment of the simplicial order. In particular, we have $f(A) \geq f(C)$. It is easy to verify that

$$f(C) = \sum_{j=0}^{r} j \binom{n}{j} + (r+1)|\mathcal{C}| = n \sum_{j=0}^{r-1} \binom{n-1}{j} + (r+1)|\mathcal{C}|. \tag{2.23}$$

Let $\mathbb{I}\{i \in x\}$ denote the indicator function of the event $\{i \in x\}$ for $x \in Q_n$. Then for any $B \subseteq Q_n$ we have

$$f(B) = \sum_{x \in B} |x| = \sum_{x \in B} \sum_{i=1}^{n} \mathbb{I}\{i \in x\} = \sum_{i=1}^{n} \sum_{x \in B} \mathbb{I}\{i \in x\} = \sum_{i=1}^{n} |B_{i,+}|. \tag{2.24}$$

Since $|A_{i,+}| = |D_+|$ for all i, it follows that

$$f(A) = n|D_+| = n \sum_{j=0}^{r-1} \binom{n-1}{j} + n|\mathcal{D}|. \tag{2.25}$$

Since $f(A) \geq f(C)$, (2.23) and (2.25) imply that we have

$$n|\mathcal{D}| \geq |\mathcal{C}|(r+1) \tag{2.26}$$

Since $|\mathcal{C}| \geq |\mathcal{D}| + |\partial_n^+ \mathcal{D}|$, (2.26) implies that

$$(n-r-1)|\mathcal{D}| \geq (r+1)\left|\partial_n^+ \mathcal{D}\right|. \tag{2.27}$$

The Local LYM inequality [31, 35, 44] for upper shadow states that for $\mathcal{D} \subseteq X_n^{(r)}$ we have

$$(r+1)\left|\partial_n^{+}\mathcal{D}\right| \geq ((n-1)-r)\left|\mathcal{D}\right|, \tag{2.28}$$

and the equality holds if and only if $\mathcal{D} = \emptyset$ or $\mathcal{D} = X_n^{(r)}$. Combining (2.27) with (2.28), we obtain that the equality must hold in (2.28).

If $\mathcal{D} = \emptyset$, then (2.26) implies that $\mathcal{C} = \emptyset$. Hence $|A| = f_r$, which contradicts our earlier assumption. If $\mathcal{D} = X_n^{(r)}$, then $D_+ = X_n^{(\leq r)}$ and hence $X_n^{(\leq r+1)} = N(D_+) \subseteq D_-$. Therefore $|A| = |D_+| + |D_-| \geq f_{r+1}$, which contradicts the assumption that $|A| < f_{r+1}$. Thus there exists a direction i satisfying $\min(|A_{i,+}|, |A_{i,-}|) > |D_+|$. □

Now we are ready to prove Theorem 4.

Theorem 4. *Let $A \subseteq Q_n$ be a subset for which every $B \subseteq Q_n$ with $|B| = |A|$ satisfies $|N(B)| \geq |N(A)|$. Then for all $t > 0$ and $B \subseteq Q_n$ with $|B| = |A|$ we have $\left|N^t(B)\right| \geq \left|N^t(A)\right|$.*

Proof. The proof is by induction on n. When $n \leq 2$, it is easy to verify that the result holds. Hence we may assume that $n \geq 3$ and that the result holds for $n-1$.

Let $A \subseteq Q_n$ be a subset so that for all $B \subseteq Q_n$ with $|B| = |A|$ we have $|N(B)| \geq |N(A)|$. By the argument presented in the proof of Lemma 18, we may assume that $|A_{i,+}| \leq |A_{i,-}|$ holds for all directions i. Let D_+ and D_- be defined as in the statement of Lemma 18. Then Lemma 19 implies that there exists a direction i satisfying $|A_{i,+}| > |D_+|$. For notational convenience, denote the i-sections in this direction by A_+ and A_-.

Let C_+ and C_- be the initial segments of the simplicial order on $\mathcal{P}(X_i)$ of the same sizes as A_+ and A_- respectively, and set $C = C_- \cup (\{i\} + C_+)$. Then $|C_+| = |A_+| > |D_+|$, and the maximality assumption on D_+ implies that we have $C_- \subseteq N(C_+)$. Since $|C_+| \leq |C_-|$, it certainly follows that $C_+ \subseteq N(C_-)$. Hence Lemma 10 with $t = 1$ implies that $N(A_+)$ and $N(A_-)$ are minimal, i.e. we can apply the inductive hypothesis on them, and that we have

$$A_{\pm} \subseteq N(A_{\mp}). \tag{2.29}$$

Since $N(A_+)$ and $N(A_-)$ are minimal, the inductive hypothesis combined with Harper's theorem implies that for all $t > 0$ we have

$$\left|N^t(A_{\pm})\right| = \left|N^t(C_{\pm})\right|. \tag{2.30}$$

By taking neighbourhood $t-1$ times from (2.29), it follows that we have $N^{t-1}(A_{\pm}) \subseteq N^t(A_{\mp})$ for all $t > 0$. Since $C_{\pm} \subseteq N(C_{\mp})$, we also have $N^{t-1}(C_{\pm}) \subseteq N^t(C_{\mp})$ for all $t > 0$. Thus (2.2) implies that we have

$$\left|N^t(A)\right| = \left|N^t(A_+)\right| + \left|N^t(A_-)\right| \tag{2.31}$$

and

$$\left|N^t(C)\right| = \left|N^t(C_+)\right| + \left|N^t(C_-)\right| \tag{2.32}$$

for all $t > 0$. Combining (2.30), (2.31) and (2.32), we obtain that $\left|N^t(A)\right| = \left|N^t(C)\right|$ holds for all $t > 0$. Thus Harper's inequality implies that for all $B \subseteq Q_n$ with $|B| = |A|$ we have $\left|N^t(B)\right| \geq \left|N^t(A)\right|$, which completes the proof of Theorem 4. □

Now we can immediately apply Theorem 4 to prove that weak extremality implies extremality.

Theorem 3. *Let $A \subseteq Q_n$ be a weakly extremal subset. Then A is extremal.*

Proof. Since for all $B \subseteq Q_n$ with $|B| = |A|$ we have $|N(B)| \geq |N(A)|$, Theorem 4 implies that for all $B \subseteq Q_n$ with $|B| = |A|$ we have $\left|N^t(B)\right| \geq \left|N^t(A)\right|$. Similarly applying Theorem 4 to A^c implies that for all $B \subseteq Q_n$ with $|B| = |A|$ we have $\left|N^t(B^c)\right| \geq \left|N^t(A^c)\right|$. Hence weak extremality implies extremality. □

Thus we can conclude the classification of weakly extremal sets.

Corollary 20. *Theorem 2 holds when extremality is replaced with weak extremality.* □

2.5 A uniqueness result for certain sizes

Recall that in Section 2.2 we defined an extremal set B_r of size g_r by setting $B_r = B(\emptyset, r) \cup B(\{1,2\}, r)$. The aim of this section is to prove that up to isomorphism the sets B_r introduced in Section 2.2 are the only sets of size g_r, together with the initial segment, for which $N(A)$ is minimal. Recently Keevash and Long [26] studied stability in the vertex isoperimetric inequality. Theorem 21 follows as a consequence of their more general result.

Theorem 21. *Let $r \leq n-1$, and let $A \subseteq Q_n$ satisfying $|A| = g_r$ for which the size of $N(A)$ is minimal among the subsets of Q_n of size g_r. Then either A is isomorphic to the initial segment of the simplicial order or A is isomorphic to B_r.*

Proof. The main idea of the proof is to carefully analyse the codimension-1 compressions. Let $A \subseteq Q_n$ satisfying $|A| = g_r$ for which $N(A)$ is minimal. As in the proof of Lemma 19, by considering $A_I = \{a\Delta I : a \in A\}$ if necessary for a suitably chosen $I \subseteq X$, we may assume that we have $|A_{i,+}| \leq |A_{i,-}|$ for all directions i.

Choose a direction i. Let $C_{i,+}$ and $C_{i,-}$ be the initial segments of the simplicial order with $|C_{i,+}| = |A_{i,+}|$ and $|C_{i,-}| = |A_{i,-}|$, and define $C = C_{i,-} \cup (\{i\} + C_{i,+})$. Let D be the initial segment of the simplicial order of size g_r, i.e. $D = X^{(\leq r)} \cup \left(\{1\} + X_1^{(r)}\right)$. Recall that we have $N(D) = X^{(\leq r+1)} \cup \left(\{1\} + X_1^{(r+1)}\right)$ and $|N(D)| = g_{r+1}$. Our first aim is to find bounds for the sizes of $A_{i,-}$ and $A_{i,+}$. Since their sizes are the same as the sizes of $C_{i,-}$ and $C_{i,+}$, it suffices to prove these bounds for the sizes of $C_{i,-}$ and $C_{i,+}$.

If $|C_{i,-}| > g_{n-1,r}$, we have $|C_{i,+}| < g_{n-1,r-1}$. Hence it follows that $|N(C_{i,+})| \leq g_{n-1,r} < |C_{i,-}|$, and thus we have $N(C_{i,+}) \subseteq C_{i,-}$. In particular, (2.2) implies that

$$|N(C)| = |N(C_{i,-})| + |C_{i,-}| > g_{n-1,r+1} + g_{n-1,r} = g_{n,r+1}.$$

Since $|N(A_{i,-})| \geq |N(C_{i,-})|$ holds by Harper's inequality, (2.2) implies that

$$|N(A)| \geq |N(A_{i,-})| + |A_{i,-}| \geq |N(C_{i,-})| + |C_{i,-}| = |N(C)| > g_{n,r+1}. \tag{2.33}$$

This contradicts the minimality of the size of $N(A)$. Hence we must have $|C_{i,-}| \leq g_{n-1,r}$.

Since $|C_{i,-}| \geq |C_{i,+}|$, it follows that $|C_{i,-}| \geq \frac{1}{2} g_{n,r} = f_{n-1,r}$. These two conditions imply that we have $f_{n-1,r} \leq |C_{i,-}| \leq g_{n-1,r}$. Our next aim is to show that $|C_{i,-}|$, which equals $|A_{i,-}|$, must always be either $f_{n-1,r}$ or $g_{n-1,r}$.

Claim 1 . For each direction i we have $|C_{i,-}| = f_{n-1,r}$. or $|C_{i,-}| = g_{n-1,r}$.

Proof of Claim 1. Since $f_{n-1,r} \leq |C_{i,-}| \leq g_{n-1,r}$ and $|C| = g_{n,r}$, it follows that $g_{n-1,r-1} \leq |C_{i,+}| \leq f_{n-1,r}$. In particular, we have $C_{i,-} \subseteq N(C_{i,+})$, and we also trivially have $C_{i,+} \subseteq N(C_{i,-})$.

Let j be the least element in X_i, i.e. $j = 1$ if $i \neq 1$ and $j = 2$ if $i = 1$. Since $C_{i,-}$ is an initial segment of the simplicial order on $\mathcal{P}(X_i)$ satisfying $f_{n-1,r} \leq |C_{i,-}| \leq g_{n-1,r}$, it follows that $C_{i,-} = X_i^{(\leq r)} \cup (\{j\} + \mathcal{A})$ where $\mathcal{A}$ is an initial segment of the lexicographic order on $X_{i,j}^{(r)}$. Similarly one can deduce that $C_{i,+} = X_i^{(\leq r-1)} \cup \left(\{j\} + X_{i,j}^{(r-1)}\right) \cup \mathcal{B}$, where $\mathcal{B}$ is an initial segment of the lexicographic order on $X_{i,j}^{(r)}$. Note that

$$\begin{aligned} |\mathcal{A}| + |\mathcal{B}| &= |A| - f_{n-1,r-1} - f_{n-1,r} - \left|X_{i,j}^{(r-1)}\right| \\ &= g_{n,r} - f_{n,r} - \binom{n-2}{r-1} = \binom{n-1}{r} - \binom{n-2}{r-1} = \binom{n-2}{r}. \end{aligned} \tag{2.34}$$

It is easy to verify that

$$N(C_{i,-}) = X_i^{(\leq r+1)} \cup \left(\{j\} + \partial_{i,j}^{+} \mathcal{A}\right)$$

and

$$N(C_{i,+}) = X_i^{(\leq r)} \cup \left(\{j\} + X_{i,j}^{(r)}\right) \cup \partial_{i,j}^{+} \mathcal{B},$$

where again $\partial_{i,j}^{+}$ denotes the upper shadow operator with respect to the ground set $X_{i,j}$. In particular, it follows that

$$|N(C_{i,-})| = f_{n-1,r+1} + \left|\partial_{i,j}^{+} \mathcal{A}\right| \tag{2.35}$$

and

$$|N(C_{i,+})| = g_{n-1,r} + \left|\partial_{i,j}^{+} \mathcal{B}\right|. \tag{2.36}$$

Note that $f_{n-1,r+1} + g_{n-1,r} = g_{n,r+1} - \binom{n-2}{r+1}$. Since $C_{i,+} \subseteq N(C_{i,-})$ and $C_{i,-} \subseteq N(C_{i,+})$, combining (2.2) together with (2.35) and (2.36) we obtain that

$$|N(C)| = |N(C_{i,+})| + |N(C_{i,-})| = g_{n,r+1} + \left|\partial_{i,j}^{+} \mathcal{A}\right| + \left|\partial_{i,j}^{+} \mathcal{B}\right| - \binom{n-2}{r+1}. \tag{2.37}$$

Applying the Local LYM inequality to $\mathcal{A}$ as a subset of $X_{i,j}^{(r)}$, it follows that

$$\left|\partial_{i,j}^{+} \mathcal{A}\right| \geq \frac{n-r-2}{r+1} |\mathcal{A}|, \tag{2.38}$$

and the equality holds if and only if $\mathcal{A}_{i,j} = X^{(r)}$ or $\mathcal{A} = \emptyset$. Certainly (2.38) holds for $\mathcal{B}$ as well.

Adding these two inequalities together we obtain that

$$\left|\partial_{i,j}^{+}\mathcal{A}\right| + \left|\partial_{i,j}^{+}\mathcal{B}\right| \geq \frac{n-r-2}{r+1}\left(|\mathcal{A}| + |\mathcal{B}|\right). \tag{2.39}$$

Since $|\mathcal{A}| + |\mathcal{B}| = \binom{n-2}{r}$ and $\frac{n-r-2}{r+1}\binom{n-2}{r} = \binom{n-2}{r+1}$, it follows that

$$\left|\partial_{i,j}^{+}\mathcal{A}\right| + \left|\partial_{i,j}^{+}\mathcal{B}\right| \geq \binom{n-2}{r+1}. \tag{2.40}$$

Combining (2.40) with (2.37), we obtain that

$$|N(C)| \geq g_{n,r+1}. \tag{2.41}$$

As in (2.33), Harper's inequality implies that we have $|N(A)| \geq |N(C)|$. Since $|N(A)| = g_{n,r+1}$, (2.41) implies that the equality holds in both applications of the Local LYM inequality. In particular, we must have $\mathcal{A} = \emptyset$ or $\mathcal{A} = X_{i,j}^{(r)}$. In the first case we have $|C_{i,-}| = f_{n-1,r}$ and in the second case we have $|C_{i,-}| = g_{n-1,r}$, as required. □

If $|A_{i,-}| = f_{n-1,r}$ for some direction i, we could use Proposition 5 to deduce that $A_{i,-}$ and $A_{i,+}$ must be exact Hamming balls. The aim of the next claim is to prove that such a direction i must always exist.

Claim 2. There exists a direction i for which we have $|A_{i,-}| = f_{n-1,r}$.

Proof of Claim 2. Suppose that the claim is false. Then by Claim 1 it follows that $|A_{i,-}| = g_{n-1,r}$ for all i. As in the proof of Lemma 19, for $B \subseteq Q_n$ define $f(B) = \sum_{x \in B}|x|$. Recall that among the sets $B \subseteq Q_n$ of a given size, $f(B)$ attains its minimum value when B is taken to be the initial segment of the simplicial order. Also recall that $f(B) = \sum_{i=1}^{n}|B_{i,+}|$ for any $B \subseteq Q_n$.

Since $|A_{i,-}| = g_{n-1,r}$ for all i, it follows that $|A_{i,+}| = g_{n-1,r-1}$ for all i. Hence $f(A) = ng_{n-1,r-1}$. Let $D = X^{(\leq r)} \cup \left(\{1\} + X_1^{(r)}\right)$ be the initial segment of the simplicial order of size $g_{n,r}$. It is easy to verify that we have $|D_{1,+}| = f_{n-1,r}$ and $|D_{i,+}| = g_{n-1,r-1}$ for all $i \geq 2$. Hence it follows that $f(D) = f_{n-1,r} + (n-1)\, g_{n-1,r-1}$. Since D is the initial segment of the simplicial order of the same size as A, it follows that $f(A) \geq f(D)$. Hence we must have $g_{n-1,r-1} \geq f_{n-1,r}$, which is only true when $r = n-1$. However, this case can only occur when $A = Q_n$, in which case we also have $|A_{i,-}| = 2^{n-1} = f_{n-1,n-1}$. This completes the proof of Claim 2. □

Let i be a direction for which we have $|A_{i,-}| = f_{n-1,r}$, and note that we also have $|A_{i,+}| = f_{n-1,r}$. Hence we have $C_{i,+} \subseteq N(C_{i,-})$ and $C_{i,-} \subseteq N(C_{i,+})$, and thus Lemma 10 with $t = 1$ implies that both $N(A_{i,+})$ and $N(A_{i,-})$ are minimal. Since $|A_{i,+}| = |A_{i,-}| = f_{n-1,r}$, Proposition 5 implies that $A_{i,-}$ and $A_{i,+}$ are exact Hamming balls on $\mathcal{P}(X_i)$ with radius r. Hence $A_{i,+} = B_i(x,r)$ and $A_{i,-} = B_i(y,r)$ for some $x, y \in Q_n$, and by symmetry we may assume that $x = \emptyset$.

Note that (2.2) implies that

$$\begin{aligned} |N(A)| &= |A_{i,+} \cup N(A_{i,-})| + |A_{i,-} \cup N(A_{i,+})| \\ &= |B_i(\emptyset, r) \cup B_i(y, r+1)| + |B_i(\emptyset, r+1) \cup B_i(y, r)|. \end{aligned}$$

Recall that $|N(A)| = g_{n,r+1} = 2f_{n-1,r+1}$ by the minimality of $|N(A)|$. Since $|B_i(y, r+1)| = |B_i(\emptyset, r+1)| = f_{n-1,r+1}$, we must have $B_i(\emptyset, r) \subseteq B_i(y, r+1)$ and $B_i(y, r) \subseteq B_i(\emptyset, r+1)$. In particular, we must have $d(y, \emptyset) \leq 1$.

If $y = \emptyset$, it follows that A is isomorphic to the initial segment of the simplicial order, and the isomorphism is given by any ϕ_σ with $\sigma(i) = 1$. If $y = \{j\}$ for some $j \in X_i$, then A is isomorphic to the set B_r, and the isomorphism is given by any ϕ_σ with $\sigma(i) = 1$ and $\sigma(j) = 2$. $\square$

Chapter 3

Coordinate deletion of zeroes

3.1 Introduction

We start by recalling the Kruskal-Katona theorem concerning the lower shadow of a set system. However, this time we work with $\{0,1\}$-sequences, i.e. we view Q_n as $\{0,1\}^n$ which is the set of $\{0,1\}$-sequences of length n. In this framework, the *lower shadow* of A is defined to be the set of points that can be obtained by flipping exactly one 1-entry to 0 from a point in A.

The *rank* of a point $x \in [k]^n = \{0,\dots,k-1\}^n$ is defined to be $|x| = \sum_{i=1}^n x_i$. Note that the lower shadow operator on $\{0,1\}^n$ decreases the rank of a point by 1. For a given r, it is natural to ask how one should choose a set $A \subseteq \{0,1\}^n$ of a given size containing only points of rank r to minimise the size of the lower shadow. This question was answered by Kruskal and Katona [24, 25].

Recall that the *colexicographic order* on $\{x \in \{0,1\}^n : |x| = r\}$ is defined by setting $x \leq_{colex} y$ if $x = y$ or $\max(X \Delta Y) \in Y$, where $X = \{i : x_i = 1\}$ and $Y = \{i : y_i = 1\}$. The Kruskal-Katona theorem states that for a set $A \subseteq \{0,1\}^n$ of a given size containing only points of rank r, the size of the lower shadow is minimised when A is chosen to be the initial segment of the colexicographic order.

Instead of changing the values of the coordinates, it is also natural to define an operator which acts by deleting the coordinates. For $A \subseteq [k]^n$, define the *coordinate deletion shadow* to be the set of points that can be obtained by deleting one coordinate from a point in A. The coordinate deletion shadow of A is denoted by $\Delta(A)$. For example, the coordinate deletion shadow of $\{000, 001, 002, 121\}$ is $\{00, 01, 02, 12, 11, 21\}$.

Again, it is natural to ask which subsets of $\{0,1\}^n$ minimise the size of the coordinate deletion shadow among the sets of a given size. Recall that the *simplicial order* $<_{sim}$ is defined on $\{0,1\}^n$ by setting

$$x <_{sim} y \text{ if } |x| < |y| \text{ or } (|x| = |y| \text{ and } \min(X \Delta Y) \in X).$$

Danh and Daykin proved that among the subsets of $\{0,1\}^n$ of a given size, the initial segment of the simplicial order minimises the size of the coordinate deletion shadow [16]. They also conjectured a certain order as the best in $[k]^n$ for $k \geq 3$, but Leck [30] showed that this turned out to be false. In fact, he proved that there is no order in general whose initial segments have minimal coordinate deletion shadow.

Bollobás and Leader [9] pointed out that the exact subcubes $[t]^n \subseteq [k]^n$ have minimal coordinate deletion shadow. Indeed, for a set $B \subseteq [k]^n$ we define B_i to be the projection of B onto the hyperplane excluding the i^{th} direction. Then the Loomis-Whitney inequality [33] implies that we have $\left(\prod_{i=1}^{n} B_i\right) \geq |B|^{n-1}$. Since for any i we have $B_i \subseteq \Delta(B)$, it follows that $|\Delta(B)|^n \geq |B|^{n-1}$. In particular, when $|B| = t^n$ we have $|\Delta(B)| \geq t^{n-1}$, which proves that $[t]^n$ has minimal coordinate deletion shadow among the subsets of $[k]^n$ of size t^n.

In addition, Bollobás and Leader made the following conjecture that certain other types of sets also have minimal coordinate deletion shadow.

Conjecture 22. *(Bollobás, Leader [9]). For each $t \leq k$ and $r \leq n$, let $B_{r,t} \subseteq [k]^n$ be the set containing all the points whose coordinates are in $[t] = \{0, \ldots, t-1\}$ and which have at most r coordinates that equal $t-1$. Then the sets $B_{r,t}$ have minimal coordinate deletion shadow among the subsets of $[k]^n$ of the same size.*

Even the case $t = k$ in the conjecture is unknown.

There is, however, a notion that comes 'between' the lower shadow and the coordinate deletion shadow. The usual lower shadow operator decreases the rank by 1 and preserves the dimension n, while the coordinate deletion shadow decreases the dimension by 1 but there is no control on how it changes the rank. Hence it is natural to consider the following operator which preserves the rank, but reduces the dimension by one.

Define the *zero-deletion shadow* of $A \subseteq [k]^n$ to be the set of points in $[k]^{n-1}$ obtained by removing one coordinate that equals 0 from a point in A. We denote the zero-deletion shadow of A by $\delta(A)$. For example, we have $\delta(\{00011, 00101\}) = \{0011, 0101\}$ and $\delta(\{112, 113, 123\}) = \emptyset$. For convenience, we say that A has minimal zero-deletion shadow if for any $B \subseteq [k]^n$ satisfying $|B| = |A|$ we have $|\delta(B)| \geq |\delta(A)|$.

How can we find sets A with minimal zero-deletion shadow? If $|A| \leq (k-1)^n$, the question is trivial as one can take any subset of $\{1, \ldots, k-1\}^n$ of the given size. In general, it is natural to choose A to contain points with as few zeroes as possible. Furthermore, it is natural to guess that for each $0 \leq i \leq n$, the sets containing all points with at most i zeroes have minimal zero-deletion shadow.

Our main result in this chapter is to find an order on $[k]^n$ whose initial segments have minimal zero-deletion shadow. In particular, it follows that the sets containing all points with at most i zeroes have minimal zero-deletion shadow.

In order to state the main result, we need a few definitions. For a point $x \in [k]^n$, let $R(x) = \{i : x_i = 0\}$ and let $w(x) = |R(x)|$. Let $L_r(n) = \{x \in [k]^n : w(x) = r\}$. Note that the zero-deletion operator maps the elements in $L_r(n)$ to elements in $L_{r-1}(n-1)$.

For $x \in [k]^n$, define its *reduced sequence* to be the sequence obtained by removing all coordinates of x that equal 0. Denote the reduced sequence of x by $re(x)$. Note that for any point s and for any $t \in \delta(\{s\})$ we have $re(s) = re(t)$, as removing a coordinate which equals 0 does not change the reduced sequence. Hence we can split $L_r(n)$ into equivalence classes which are characterised by the reduced sequences.

We start by proving that inside an equivalence class one should choose points x so that the set system containing the sets $R(x)$ is an initial segment of the colexicographic order. This is a straightforward consequence of the work of Danh and Daykin in [16].

Since $[k]^n$ splits into equivalence classes based on the reduced sequences, and we know that the initial segments of the colexicographic order minimise the size of the zero-deletion shadow inside each equivalence class, we are left with the question on how to split the points into different equivalence classes. We go on to prove that in order to minimise the zero-deletion shadow of a subset of $[k]^n$, one should prioritise points in $L_r(n)$ over points that are in $L_s(n)$ for $r < s$, and inside $L_r(n)$ one should choose all points from one equivalence class before choosing points from another equivalence class. As a consequence, we obtain an order whose initial segments have minimal zero-deletion shadow.

For $r \in [k]$, define $R_r(x) = \{i : x_i = r\}$ and $w_r(x) = |R_r(x)|$. Note that we have $R = R_0$ and $w = w_0$. For each i, define the order $\leq_c$ on $\{1, \ldots, k-1\}^i$ as follows. For distinct $x, y \in \{1, \ldots, k-1\}^i$ let $r \in \{1, \ldots, k\}$ be the minimal index satisfying $R_r(x) \neq R_r(y)$. We say that $x \leq_c y$ if and only if $\max(R_r(x) \Delta R_r(y)) \in R_r(y)$.

Finally, define the order $\leq$ on $[k]^n$ as follows. For distinct x, $y \in [k]^n$ we set $x \leq y$ if one of the following conditions holds.

1. $w_0(x) < w_0(y)$
2. $w_0(x) = w_0(y)$, $re(x) \neq re(y)$ and $re(x) \leq_c re(y)$
3. $w_0(x) = w_0(y)$, $re(x) = re(y)$ and $R_0(x) \leq_{colex} R_0(y)$

Note that for any distinct x and y exactly one of $x \leq y$ and $y \leq x$ is satisfied. Hence $\leq$ defines an order on $[k]^n$.

Now we are ready to state our main theorem.

Theorem 23. *Let $A \subseteq [k]^n$ and let B be the initial segment of the $\leq$-order of size $|A|$. Then we have $|\delta(A)| \geq |\delta(B)|$.*

In particular, it follows that the sets of the form $L_{\leq r}(n) = \bigcup_{i=0}^{r} L_i(n)$ have minimal zero-deletion shadow. Note that the zero-deletion shadow is affected only by the location of non-zero coordinates and not on their exact values. Hence all the equivalence classes of $L_r(n)$ behave in the same way. Thus for any fixed r, one could replace the $\leq_c$-order with any other order on $\{1, \ldots, k-1\}^r$ in the definition of the $\leq$-order.

The plan of this chapter is as follows. In Section 3.2 we prove that in order to minimise the size of the zero-deletion shadow inside an equivalence class, one should choose points x so that the set system containing the sets $R(x)$ forms an initial segment of the colexicographic order. In Section 3.3 we prove Theorem 23. In Section 3.4 we generalise the zero-deletion shadow to allow deleting a coordinate that is in the set $\{0, \ldots, r\}$ for a chosen r instead of just deleting a coordinate which equals 0. In this case we show that the sets $\{x : \sum_{i=0}^{r} w_i(x) \leq s\}$, which are analogous to the sets $L_{\leq s}(n)$, minimise the size of the shadow for each $0 \leq s \leq n$. In this general case we do not know what happens for sets of other sizes.

As in the previous chapter, we write $X = \{1, \ldots, n\}$ and $X^{(r)} = \{A \subseteq \{1, \ldots, n\} : |A| = r\}$. We write L_r instead of $L_r(n)$ if the value of n is clear. When $k = 1$, we may also write $\{0,1\}^n_r$ instead of $L_r(n)$. This notation will be used to highlight that we are working with $\{0,1\}$-sequences.

3.2 Deletion on $\{0,1\}$-sequences

In this section we always work with subsets of $\{0,1\}^n$ or $\{0,1\}^n_r$. Danh and Daykin proved in [16] the following result for the coordinate deletion shadow on $\{0,1\}^n$.

Theorem 24. *(Danh, Daykin). Let $A \subseteq \{0,1\}^n$ and let B be the initial segment of the simplicial order of size $|A|$. Then we have $|\Delta(A)| \geq |\Delta(B)|$.* □

Recall that there is a natural correspondence between $\{0,1\}^n$ and the power-set $\mathcal{P}(X)$. For our purposes, it is convenient to choose this correspondence to be given by mapping a sequence (x_i) to the set $R_0(x) = \{i : x_i = 0\}$. In this way we can identify a set $A \subseteq \{0,1\}^n_r$ with the set system $\mathcal{A} \subseteq X^{(r)}$ containing the images of the elements of A under this bijection. This correspondence enables us to translate questions related to the zero-deletion shadow to questions related to the properties of the set systems $\mathcal{A} \subseteq X^{(r)}$ instead. We start by proving that the subsets A of $\{0,1\}^n_r$ with minimal zero-deletion shadow are the ones whose corresponding set $\mathcal{A}$ is an initial segment of the colexicographic order.

Lemma 25. *Let $A \subseteq \{0,1\}^n_r$ and let $B \subseteq \{0,1\}^n_r$ be the initial segment of the colexicographic order of size $|A|$. Then we have $|\delta(A)| \geq |\delta(B)|$.*

Proof. Define $C_1 = A \cup L_{>r}(n)$ and $C_2 = B \cup L_{>r}(n)$, where $L_{>r}(n) = \bigcup_{i=r+1}^{n} \{0,1\}^n_i$. Note that C_2 is isomorphic to an initial segment of the simplicial order, where the isomorphism is given by the map which reverses the sequence. Since this map also preserves the size of the coordinate deletion shadow, Theorem 24 implies that we have

$$|\Delta(C_1)| \geq |\Delta(C_2)|. \tag{3.1}$$

It is easy to check that we have $\Delta(C_1) = L_{>r}(n-1) \cup \delta(A)$ and $\Delta(C_2) = L_{>r}(n-1) \cup \delta(B)$. Indeed, $L_{>r}(n-1)$ is certainly contained in both $\Delta(C_1)$ and $\Delta(C_2)$, and the only contribution to the coordinate deletion shadow outside $L_{>r}(n-1)$ arises by removing 0 from a point which contains exactly $n-r$ coordinates that equal 1. Hence we have

$$|\Delta(C_1)| = |L_{>r}(n-1)| + |\delta(A)| \tag{3.2}$$

and

$$|\Delta(C_2)| = |L_{>r}(n-1)| + |\delta(B)|. \tag{3.3}$$

Thus (3.1), (3.2) and (3.3) imply that we have $|\delta(A)| \geq |\delta(B)|$. □

Lemma 25 implies that the initial segments of the colexicographic order have minimal zero-deletion shadow among the subsets of $\{0,1\}^n_r$. Before moving on to the general case from $\{0,1\}$-sequences, we find a way to relate the size of $\delta(A)$ to a certain property of the associated set system $\mathcal{A}$ for $A \subseteq \{0,1\}^n_r$. For convenience, from now on, we say that $A \subseteq \{0,1\}^n_r$ is an initial segment of the colexicographic order if the associated set system $\mathcal{A}$ is an initial segment of the colexicographic order. For $\mathcal{A} \subseteq \mathcal{P}(X)$ define $\mathcal{A}_1 = \{B \in \mathcal{A} : 1 \in B\}$.

Lemma 26. *Let $A \subseteq \{0,1\}^n_r$ be an initial segment of the colexicographic order, and let $\mathcal{A}$ be the set system associated to A. Then we have $|\delta(A)| = |\mathcal{A}_1|$.*

Proof. The proof is by induction on the size of A; note that the statement certainly holds when $|A| = 1$. Suppose that the statement holds for A, and let B be the initial segment satisfying $|B| = |A| + 1$. Let x be the unique element in $B \setminus A$, and write $x = x_1 \dots x_n$.

We start by proving that either $\delta(B) \setminus \delta(A)$ is empty or it contains only one element which is $x_2 \dots x_n$. Indeed, suppose that $t \in \delta(B) \setminus \delta(A)$ and that t is obtained by removing the r^{th} coordinate of x. Hence $t = x_1 \dots x_{r-1} x_{r+1} \dots x_n$ and $x_r = 0$. Let $i = \min\{j : x_j = 1\}$, and consider the particular element $y = 0t$. If $i \le r$, we have $y_j = x_j$ for all $j \le i - 1$ and $y_i = 0 < 1 = x_i$. Hence we have $y <_{colex} x$. However, as $t \in \delta(\{y\}) \subseteq \delta(A)$, this contradicts the fact that $t \in \delta(B) \setminus \delta(A)$. Hence we must have $i > r$, which implies that $x_1 = \dots = x_r = 0$. In particular, it follows that $t = x_1 \dots x_{r-1} x_{r+1} x_n = x_2 \dots x_n$, and thus $\delta(B) \setminus \delta(A)$ is either empty or contains only one element $x_2 \dots x_n$.

Note that $0x_2 \dots x_n$ is the least element with respect to the colexicographic order for which the element $x_2 \dots x_n$ is contained in its zero-deletion shadow. Hence it follows that $x_2 \dots x_n \in \delta(B) \setminus \delta(A)$ if and only if $x = 0x_2 \dots x_n$. In particular, we have

$$|\delta(B)| = \begin{cases} |\delta(A)| + 1 & \text{if } x_1 = 0 \\ |\delta(A)| & \text{if } x_1 = 1 \end{cases}.$$

Since $\mathcal{B} = \mathcal{A} \cup \{R_0(x)\}$ and the set $R_0(x)$ contains 1 if and only if $x_1 = 0$, it follows that

$$|\mathcal{B}_1| = \begin{cases} |\mathcal{A}_1| + 1 & \text{if } x_1 = 0 \\ |\mathcal{A}_1| & \text{if } x_1 = 1 \end{cases}.$$

Hence we have $|\delta(B)| = |\mathcal{B}_1|$ by induction. $\square$

3.3 The main theorem

Let H be the bipartite graph on the vertex set $[k]^n \cup [k]^{n-1}$ whose edges are precisely the pairs s, t with $s \in [k]^n$ and $t \in \delta(\{s\})$. Note that $\delta(A)$ is the neighbourhood of $A \subseteq [k]^n$ in the graph H. Observe that both vertex classes of H can be partitioned as $[k]^n = \bigcup_{i=0}^{n} L_i(n)$ and $[k]^{n-1} = \bigcup_{i=0}^{n-1} L_i(n-1)$, and by definition of the zero-deletion shadow it is clear that there are edges only between $L_i(n)$ and $L_{i-1}(n-1)$, with the convention $L_{-1} = \emptyset$.

Let C be a non-trivial connected component in H, i.e. a connected component satisfying $[k]^n \cap C \subseteq L_i(n)$ for some $i > 0$. Recall that the reduced sequence of a point $y \in \delta(\{x\})$ is always the same as the reduced sequence of x. Since C is a connected component, it follows that every $z \in C$ has the same reduced sequence. Conversely, it is easy to check that for each $i > 0$, all the elements in $L_i(n) \cup L_{i-1}(n-1)$ with the same reduced sequence are also in the same connected component. Hence the non-trivial connected components of H are characterised by the reduced sequences.

Lemma 27. *For $s \in \bigcup_{i=1}^{r} \{1, \dots, k-1\}^i$ define*

$$C_s = \{x \in [k]^n : re(x) = s\}$$

and

$$D_s = \left\{ x \in [k]^{n-1} : re(x) = s \right\}.$$

Then the sets of the form $C_s \cup D_s$ are the non-trivial connected components of H. □

Note that the only points outside the non-trivial connected components are the points in $\{1, \ldots, k-1\}^n$, and their zero-deletion shadows are always empty.

Broadly speaking, we only need to understand how to minimise the size of the zero-deletion shadow inside a connected component, and to determine how to distribute the points into different connected components in order to minimise the size of the zero-deletion shadow. It turns out that inside a connected component one should choose points x so that the set system consisting of the sets $R(x)$ forms an initial segment of the colexicographic order.

Lemma 28. *Let $C \subseteq L_i(n) \cup L_{i-1}(n-1)$ be a connected component corresponding to a reduced sequence $x = x_1 \ldots x_{n-i}$. Let $B \subseteq L_i \cap C$ and let $A \subseteq L_i \cap C$ be the set of points of size $|B|$ chosen so that $\{R_0(x) : x \in A\}$ is an initial segment of the colexicographic order. Then we have $|\delta(B)| \geq |\delta(A)|$.*

Proof. Note that the order of the coordinates is preserved under the zero-deletion shadow, and recall that the reduced sequence is preserved under the zero-deletion shadow. Moreover, the exact values of the coordinates in the reduced sequence do not affect the zero-deletion shadow, and hence the behaviour of the connected component depends only on the value of i, and not on the exact values of the coordinates $x_1, \ldots, x_{n-i}$.

In particular, all connected components containing points with equal number of zeroes have the same size and they all behave in the same way under the zero-deletion shadow. Hence each connected component is identical to the connected component corresponding to the reduced sequence $x_1 = \cdots = x_{n-i} = 1$. Since this particular component is $\{0,1\}_i^n$, the result follows from Lemma 25. □

Our next aim is to understand how to fill different connected components. We will prove that it is optimal to first prioritise points in $L_i(n)$ over points in $L_{i+1}(n)$, and that it is also optimal to choose all points in one connected component on a level $L_i(n)$ before choosing points from another component on the same level.

From now on, we call the sets C_s *connected components*. That is, by a connected component we refer to the intersection of a connected component with $[k]^n$.

For $s, t \in \bigcup_{i=0}^{r} \{1, \ldots, k-1\}^i$ define the s,t*-compression* operator as follows. For $A \subseteq [k]^n$, its compression $B = C_{s,t}(A)$ is given by setting

1. $B \cap C_s$ to be the initial segment of the colexicographic order of size $\min(|A \cap (C_s \cup C_t)|, |C_s|)$
2. $B \cap C_t$ to be the initial segment of the colexicographic order of size $\max(0, |A \cap (C_s \cup C_t)| - |C_s|)$
3. $B \setminus (C_s \cup C_t) = A \setminus (C_s \cup C_t)$

It is clear that we have $|C_{s,t}(A)| = |A|$ for all s and t. As usual, we say that $A \subseteq [k]^n$ is s,t*-compressed* if $C_{s,t}(A) = A$.

In order to prove Theorem 23, we need the following two results.

Lemma 29. *For any $A \subseteq [k]^n$ and $s, t \in \{1, \ldots, k-1\}^{n-i}$ we have $|\delta(A)| \geq |\delta(C_{s,t}(A))|$.*

Lemma 30. *For any $A \subseteq [k]^n$, $s \in \{1, \ldots, k-1\}^{n-i}$ and $t \in \{1, \ldots, k-1\}^{n-i-1}$ we have $|\delta(A)| \geq |\delta(C_{s,t}(A))|$.*

In order to prove these results, we translate them to questions related to the properties of the set systems on $X^{(i)}$. We now state the versions of these questions that are concerned with set systems, and their proofs are given after the proofs of Lemma 29 and Lemma 30.

A set system $\mathcal{B} \subseteq X^{(i)}$ is said to be a *segment* if there exist initial segments $\mathcal{I}$ and $\mathcal{J}$ of the colexicographic order satisfying $\mathcal{A} = \mathcal{I} \setminus \mathcal{J}$.

Lemma 31. *The following claims are true.*

Claim 1. Let $\mathcal{A} \subseteq X^{(i)}$ be a segment and let $\mathcal{I} \subseteq X^{(i)}$ be the initial segment of the colexicographic order of size $|\mathcal{A}|$. Then we have $|\mathcal{I}_1| \geq |\mathcal{A}_1|$.

Claim 2. Let $\mathcal{I} \subseteq X^{(i)}$ and $\mathcal{J} \subseteq X^{(i+1)}$ be initial segments of the colexicographic order satisfying $|\mathcal{I}| = |\mathcal{J}|$. Then we have $|\mathcal{J}_1| \geq |\mathcal{I}_1|$.

Claim 3. Let $\mathcal{A} \subseteq X^{(i)}$ be a segment and let $\mathcal{I} = X^{(i)} \setminus \mathcal{J}$, where $\mathcal{J}$ is the initial segment of the colexicographic order chosen so that $\left|X^{(i)} \setminus \mathcal{J}\right| = |\mathcal{A}|$. Then we have $|\mathcal{A}_1| \geq |\mathcal{I}_1|$.

Claim 4. Let $\mathcal{I}_*$ and $\mathcal{J}_*$ be initial segments of the colexicographic order chosen so that $\mathcal{I} = X^{(i)} \setminus \mathcal{I}_*$ and $\mathcal{J} = X^{(i+1)} \setminus \mathcal{J}_*$ satisfy $|\mathcal{I}| = |\mathcal{J}|$. Then we have $|\mathcal{J}_1| \geq |\mathcal{I}_1|$.

Proof of Lemma 29. Let $A \subseteq [k]^n$ and $B = C_{s,t}(A)$, and note that B depends only on the sizes of $A \cap C_s$ and $A \cap C_t$. Hence we may assume that $Q = A \cap C_s$ and $R = A \cap C_t$ are initial segments of the colexicographic order by Lemma 28.

Let $S = B \cap C_s$ and $T = B \cap C_t$. Let $\mathcal{Q}$, $\mathcal{R}$, $\mathcal{S}$ and $\mathcal{T}$ be the associated set systems in $X^{(i)}$. By using Lemma 26 and the fact that $B \setminus (C_s \cup C_t) = A \setminus (C_s \cup C_t)$, it follows that the conditions $|\delta(A)| \geq |\delta(B)|$ and $|\mathcal{Q}_1| + |\mathcal{R}_1| \geq |\mathcal{S}_1| + |\mathcal{T}_1|$ are equivalent. We now split the rest of the proof into two cases based on the sizes of Q and R.

Case 1. Q and R satisfy $|Q| + |R| \leq |C_s|$.

By definition of B it follows that $T = \emptyset$ and $|S| = |Q| + |R|$. Let $\mathcal{I} = \mathcal{S} \setminus \mathcal{Q}$. Since $\mathcal{S}$ and $\mathcal{Q}$ are initial segments of the colexicographic order, it follows that $\mathcal{I}$ is a segment of size $|\mathcal{R}|$. Thus Claim 1 implies that $|\mathcal{R}_1| \geq |\mathcal{I}_1|$, and hence we have

$$|\mathcal{Q}_1| + |\mathcal{R}_1| \geq |\mathcal{Q}_1| + |\mathcal{I}_1| = |\mathcal{S}_1| = |\mathcal{S}_1| + |\mathcal{T}_1|,$$

as required.

Case 2. Q and R satisfy $|Q| + |R| > |C_s|$.

In this case we have $S = C_s$ and hence it follows that $|T| \leq |R|$. Let $\mathcal{I} = \mathcal{R} \setminus \mathcal{T}$, and note that $\mathcal{I}$ is a segment. Let $\mathcal{J} = \mathcal{S} \setminus \mathcal{Q} = X^{(i)} \setminus \mathcal{Q}$, and note that $\mathcal{J}$ is also a segment. Since $|\mathcal{S}| + |\mathcal{T}| = |\mathcal{R}| + |\mathcal{Q}|$, it follows that $|\mathcal{I}| = |\mathcal{J}|$. Thus Claim 3 implies that $|\mathcal{I}_1| \geq |\mathcal{J}_1|$. Combining this together with the definitions of $\mathcal{I}$ and $\mathcal{J}$, it follows that

$$|\mathcal{Q}_1| + |\mathcal{R}_1| = |\mathcal{Q}_1| + |\mathcal{I}_1| + |\mathcal{T}_1| \geq |\mathcal{Q}_1| + |\mathcal{J}_1| + |\mathcal{T}_1| = |\mathcal{S}_1| + |\mathcal{T}_1|,$$

which completes the proof of Lemma 29. $\square$

Proof of Lemma 30. Let $A \subseteq [k]^n$ and $B = C_{s,t}(A)$. As before, we may assume that both $A \cap C_s$ and $A \cap C_t$ are initial segments of the colexicographic order, and again we set $Q = A \cap C_s$, $R = A \cap C_t$, $S = B \cap C_s$ and $T = B \cap C_t$. Let $\mathcal{Q}$ and $\mathcal{S}$ be the associated set systems in $X^{(i)}$, and $\mathcal{R}$ and $\mathcal{T}$ be the associated set systems in $X^{(i+1)}$. Again, Lemma 26 implies that the conditions $|\delta(A)| \geq |\delta(B)|$ and $|\mathcal{Q}_1| + |\mathcal{R}_1| \geq |\mathcal{S}_1| + |\mathcal{T}_1|$ are equivalent. Again, we split the proof into two cases based on the sizes of Q and R.

***Case 1*.** Q and R satisfy $|\mathcal{Q}| + |\mathcal{R}| \leq |C_s|$.

By definition of B, it follows that $\mathcal{S}$ is an initial segment of the colexicographic order of size $|\mathcal{Q}| + |\mathcal{R}|$ on $X^{(i)}$, and we also have $\mathcal{T} = \emptyset$. Let $\mathcal{I}$ be the initial segment of the colexicographic order of size $|\mathcal{R}|$ on $X^{(i)}$, and set $\mathcal{J} = \mathcal{S} \setminus \mathcal{Q}$. Then $\mathcal{J}$ is a segment satisfying $|\mathcal{J}| = |\mathcal{R}| = |\mathcal{I}|$. Thus Claim 1 implies that we have $|\mathcal{I}_1| \geq |\mathcal{J}_1|$. On the other hand, Claim 2 implies that we have $|\mathcal{R}_1| \geq |\mathcal{I}_1|$, and these two inequalities imply that $|\mathcal{R}_1| \geq |\mathcal{J}_1|$. Hence it follows that

$$|\mathcal{R}_1| + |Q_1| \geq |\mathcal{J}_1| + |\mathcal{Q}_1| = |\mathcal{S}_1| = |\mathcal{S}_1| + |\mathcal{T}_1|,$$

as required.

***Case 2*.** Q and R satisfy $|\mathcal{Q}| + |\mathcal{R}| > |C_s|$.

By definition of B it follows that $\mathcal{S} = X^{(i)}$. Since $|\mathcal{S}| \geq |\mathcal{Q}|$, it follows that $|\mathcal{R}| \geq |\mathcal{T}|$. Hence $\mathcal{I} = \mathcal{R} \setminus \mathcal{T} \subseteq X^{(i+1)}$ is a segment satisfying $\mathcal{R} = \mathcal{I} \cup \mathcal{T}$. Let $\mathcal{I}_* \subseteq X^{(i+1)}$ be the initial segment of the colexicographic order chosen so that $\mathcal{K} = X^{(i+1)} \setminus \mathcal{I}_*$ is a segment of size $|\mathcal{I}|$. Define $\mathcal{J} = X^{(i)} \setminus \mathcal{Q} = \mathcal{S} \setminus \mathcal{Q}$. Hence $\mathcal{J}$ is a segment of size $|\mathcal{S}| - |\mathcal{Q}| = |\mathcal{R}_1| - |\mathcal{T}_1| = |\mathcal{I}|$.

Claim 3 implies that $|\mathcal{I}_1| \geq |\mathcal{K}_1|$ and Claim 4 implies that $|\mathcal{K}_1| \geq |\mathcal{J}_1|$, and hence it follows that $|\mathcal{I}_1| \geq |\mathcal{J}_1|$. Thus we have

$$|\mathcal{R}_1| + |\mathcal{Q}_1| = |\mathcal{I}_1| + |\mathcal{Q}_1| + |\mathcal{T}_1| \geq |\mathcal{J}_1| + |\mathcal{Q}_1| + |\mathcal{T}_1| = |\mathcal{S}_1| + |\mathcal{T}_1|,$$

which completes the proof of Lemma 30. $\square$

Proof of Lemma 31. We start by proving Claim 1, and then we prove that the other results follow from Claim 1.

Proof of Claim 1. Since $\mathcal{A}$ is a segment, there exist initial segments $\mathcal{I}_A$ and $\mathcal{J}_A$ of the colexicographic order so that $\mathcal{A} = \mathcal{I}_A \setminus \mathcal{J}_A$. Denote the set of points associated to these set systems by I_A and J_A respectively. Let C be obtained from J_A by adding $2n$ coordinates that equal 1 at the start of each point, and let D be obtained from I by adding $2n$ coordinates that equal 1 at the end of each point in I, where I is the set of points associated to $\mathcal{I}$. Finally, we set $B = C \cup D$.

Due to the added 1's at the start of the elements of C and at the end of the elements of D, it follows that $\delta(C)$ and $\delta(D)$ are disjoint sets. Also note that adding 1's to every point does not change the size of the zero-deletion shadow. Hence we have $|\delta(B)| = |\delta(C)| + |\delta(D)| = |\delta(I)| + |\delta(J_A)|$. On the other hand, since $\mathcal{I}$ and $\mathcal{J}_A$ are initial segments of the colexicographic order, Lemma 26 implies that we have $|\delta(I)| = |\mathcal{I}_1|$ and $|\delta(J_A)| = |(\mathcal{J}_A)_1|$. Hence it follows that

$$|\delta(B)| = |\mathcal{I}_1| + |(\mathcal{J}_A)_1|. \tag{3.4}$$

Since $\mathcal{I}_A$ is an initial segment of the colexicographic order, Lemma 26 implies that $|\delta(I_A)| = |(\mathcal{I}_A)_1|$. Since $\mathcal{I}_A$ is a disjoint union of $\mathcal{J}_A$ and $\mathcal{A}$, it follows that

$$|\delta(I_A)| = |(\mathcal{I}_A)_1| = |(\mathcal{J}_A)_1| + |\mathcal{A}_1|. \tag{3.5}$$

Since $\mathcal{I}_A$ is an initial segment of the colexicographic order, the corresponding set of points I_A has minimal zero-deletion shadow inside a connected component. Since $|\mathcal{B}| = |\mathcal{I}_A|$, it follows that

$$|\delta(B)| \geq |\delta(I_A)|. \tag{3.6}$$

Thus (3.4), (3.5) and (3.6) imply that

$$|\mathcal{I}_1| \geq |\mathcal{A}_1|, \tag{3.7}$$

which completes the proof. □

Claim 1 ⇒ Claim 3. Let $\mathcal{A}$ and $\mathcal{I}$ be defined as in the statement of Claim 3. For a set system $\mathcal{D} \subseteq X^{(r)}$, recall that $\overline{\mathcal{D}}$ is defined by setting $\overline{\mathcal{D}} = \{A^c : A \in \mathcal{D}\}$, and recall that we have $|\overline{\mathcal{D}}| = |\mathcal{D}|$ and $\overline{\mathcal{D}} \subseteq X^{(n-r)}$. Furthermore, it is easy to check that if $\mathcal{D} \subseteq X^{(r)}$ is an initial segment of the colexicographic order then so is $\overline{(X^{(r)} \setminus \mathcal{D})}$. In particular, it follows that $\overline{\mathcal{I}}$ is an initial segment of the colexicographic order.

Since $\mathcal{A}$ is a segment, there exist initial segments $\mathcal{K}$ and $\mathcal{L}$ so that $\mathcal{A} = \mathcal{K} \setminus \mathcal{L}$. This can be rewritten as $\mathcal{A} = \left(X^{(r)} \setminus \mathcal{L}\right) \setminus \left(X^{(r)} \setminus \mathcal{K}\right)$, and hence it follows that

$$\overline{\mathcal{A}} = \overline{\left(X^{(r)} \setminus \mathcal{L}\right) \setminus \left(X^{(r)} \setminus \mathcal{K}\right)} = \overline{\left(X^{(r)} \setminus \mathcal{L}\right)} \setminus \overline{\left(X^{(r)} \setminus \mathcal{K}\right)}.$$

Since $\overline{\left(X^{(r)} \setminus \mathcal{L}\right)}$ and $\overline{\left(X^{(r)} \setminus \mathcal{K}\right)}$ are initial segments of the colexicographic order, it follows that $\overline{\mathcal{A}}$ is also a segment. Hence $\overline{\mathcal{A}}$ and $\overline{\mathcal{I}}$ satisfy the conditions of Claim 1, and therefore we have

$$\left|\left(\overline{\mathcal{I}}\right)_1\right| \geq \left|\left(\overline{\mathcal{A}}\right)_1\right|. \tag{3.8}$$

Note that for any set system $\mathcal{B}$ we have $|\mathcal{B}| = |\mathcal{B}_1| + \left|\left(\overline{\mathcal{B}}\right)_1\right|$, as for every $A \in \mathcal{B}$ exactly one of the conditions $1 \in A$ and $1 \in A^c$ is satisfied. Thus it follows that

$$|\mathcal{I}| = |\mathcal{I}_1| + \left|\left(\overline{\mathcal{I}}\right)_1\right| \tag{3.9}$$

and

$$|\mathcal{A}| = |\mathcal{A}_1| + \left|\left(\overline{\mathcal{A}}\right)_1\right|. \tag{3.10}$$

Thus (3.8), (3.9) and (3.10) together with $|\mathcal{I}| = |\mathcal{A}|$ imply that we have

$$|\mathcal{A}_1| \geq |\mathcal{I}_1|, \tag{3.11}$$

as required. $\square$

Claim 1 $\Rightarrow$ Claim 2. Let $\mathcal{I}$ and $\mathcal{J}$ be defined as in the statement of Claim 2. For each $i+1 \leq j \leq n$ let $\mathcal{S}_j = \{A \setminus \{j\} : A \in \mathcal{J}, \max A = j\}$, and note that we have $\mathcal{S}_j \subseteq \{1, \ldots, j-1\}^{(i)} \subseteq X^{(i)}$. Since $\mathcal{J}$ is an initial segment of the colexicographic order, it follows that $\mathcal{S}_j$ is an initial segment of the colexicographic order on $\{1, \ldots, j-1\}^{(i)}$ for each j. Thus $\mathcal{S}_j$ is an initial segment of the colexicographic order also on $X^{(i)}$, as the definition of the colexicographic order is independent of the ground set.

Note that we can express $\mathcal{J}$ as a disjoint union $\mathcal{J} = \bigcup_{j=i+1}^{n} (\mathcal{S}_j + \{j\})$. Hence it follows that

$$|\mathcal{J}_1| = \sum_{j=i+1}^{n} \left|(\mathcal{S}_j + \{j\})_1\right| = \sum_{j=i+1}^{n} \left|(\mathcal{S}_j)_1\right|. \tag{3.12}$$

Since each $\mathcal{S}_j$ is an initial segment of the colexicographic order on $X^{(i)}$ and we have $\sum_{j=i+1}^{n} |\mathcal{S}_j| = |\mathcal{J}| = |\mathcal{I}|$, applying Claim 1 $n-i-2$ times implies the result. $\square$

Claim 2 $\Rightarrow$ Claim 4. Let $\mathcal{I}$, $\mathcal{J}$, $\mathcal{I}_*$ and $\mathcal{J}_*$ be defined as in the statement of Claim 4. Since $\mathcal{I}_*$ and $\mathcal{J}_*$ are initial segments of the colexicographic order, the observation pointed out in the proof of Claim 3 implies that $\overline{\mathcal{I}} \subseteq X^{(n-i)}$ and $\overline{\mathcal{J}} \subseteq X^{(n-i-1)}$ are initial segments of the colexicographic order as well. Thus Claim 2 implies that we have

$$\left|(\overline{\mathcal{I}})_1\right| \geq \left|(\overline{\mathcal{J}})_1\right|. \tag{3.13}$$

Combining this observation with

$$|\mathcal{I}| = |\mathcal{I}_1| + \left|(\overline{\mathcal{I}})_1\right| \tag{3.14}$$

and

$$|\mathcal{J}| = |\mathcal{J}_1| + \left|(\overline{\mathcal{J}})_1\right| \tag{3.15}$$

implies that $|\mathcal{J}_1| \geq |\mathcal{I}_1|$. $\square$

This completes the proof of Lemma 31. $\square$

We are now ready to deduce Theorem 23. For convenience, we recall the definition of the $\leq$-order. For distinct x, $y \in [k]^n$ we set $x \leq y$ if one of the following conditions holds.

1. $w_0(x) < w_0(y)$

2. $w_0(x) = w_0(y)$, $re(x) \neq re(y)$ and $re(x) \leq_c re(y)$

3. $w_0(x) = w_0(y)$, $re(x) = re(y)$ and $R_0(x) \leq_{colex} R_0(y)$

Proof of Theorem 23. Let A be a subset of $[k]^n$. For $D \subseteq [k]^n$, define

$$v(D) = \sum_{j=0}^{n} j \left|D \cap L_j(n)\right|.$$

Note that for $l \in \{1, \ldots, n\}$, $s \in \{1, \ldots, k-1\}^{n-l}$ and $t \in \{1, \ldots, k-1\}^{n-l-1}$ we have $v(D) \geq v(C_{s,t}(D))$, and the equality holds only when D is s,t-compressed.

Starting with $A_0 = A$, we construct a sequence (A_i) as follows. Given A_i, if there exist $l \in \{1, \ldots, n\}$, $s \in \{1, \ldots, k-1\}^{n-l}$ and $t \in \{1, \ldots, k-1\}^{n-l-1}$ for which A_i is not s, t-compressed, we set $A_{i+1} = C_{s,t}(A_i)$. Otherwise, we set $A_{i+1} = A_i$. Since $v(A_i)$ is a non-increasing sequence of non-negative integers, it must eventually become constant, and hence the sequence (A_i) is also eventually constant. In particular, there exists A_m that is s, t-compressed for all $s \in \{1, \ldots, k-1\}^{n-l}$ and $t \in \{1, \ldots, k-1\}^{n-l-1}$. Hence there exists i for which we have $L_j(n) \subseteq A_m$ for all $j < i$ and $L_j(n) \cap A_m = \emptyset$ for all $j > i$. Note that Lemma 28 implies that we have $|\delta(A)| \geq |\delta(A_m)|$.

Let $C_{s_1}, \ldots, C_{s_t}$ be the connected components in $L_i(n)$ so that $s_u \leq_c s_v$ whenever $u \leq v$, and define $w(D) = \sum_{j=1}^{t} j \left|D \cap C_{s_j}\right|$. It is not too hard to check that for each $u < v$ we have $w(D) \geq w\left(C_{s_u,s_v}(D)\right)$, and the equality holds only when D is s_u, s_v-compressed. Again, we construct a sequence of sets starting with $B_0 = A_m$. Given B_i, if there exist $u < v$ for which B_i is not s_u, s_v-compressed, we set $B_{i+1} = C_{s_u,s_v}(B_i)$, and otherwise we set $B_{i+1} = B_i$. As before, $w(B_i)$ is a non-increasing sequence of non-negative integers, and hence there exists M so that B_M is s_u, s_v-compressed for all $u < v$. Hence there exist i and p so that $L_j(n) \subseteq B_M$ for all $j < i$, $L_j(n) \cap B_M = \emptyset$ for all $j > i$, $C_{s_u} \subseteq B_M$ for all $u < p$ and $C_{s_u} \cap B_M = \emptyset$ for all $u > p$. Note that Lemma 29 implies that we have $|\delta(B_0)| \geq |\delta(B_M)|$, and hence it follows that $|\delta(A)| \geq |\delta(B_M)|$.

Let $D = B_M \cap C_{s_p}$ and let B be the set obtained by taking $B \cap C_{s_p}$ to be the set corresponding to the initial segment of the colexicographic order of size $|D|$, and by taking $B \setminus C_{s_p} = B_M \setminus C_{s_p}$. Thus Lemma 25 implies that we have $|\delta(B_M)| \geq |\delta(B)|$. It is not too hard to check that the set B constructed in this way is an initial segment of the $\leq$-order. Since we have proved that $|\delta(A)| \geq |\delta(B)|$, this completes the proof. $\square$

3.4 An extremal result for the generalised shadow

So far we have considered an operator which allows us to delete a coordinate that equals 0. It is natural to ask what happens if we generalise this set-up and allow the deletion of any coordinate that is in some chosen set.

Define the *δ_r-shadow* of $A \subseteq [k]^n$ to be the set of points obtained by removing exactly one coordinate that is in $\{0, \ldots, r\}$ from a point in A. In particular, we have $\delta = \delta_0$ and $\Delta = \delta_{k-1}$. Define $v_r(x) = \sum_{i=0}^{r} w_i(x)$, i.e. $v_r(x)$ is the number of coordinates of x in the set $\{0, \ldots, r\}$. Define $L_s(n) = \{x \in [k]^n : v_r(x) = s\}$ and $L_{\leq s}(n) = \bigcup_{i=0}^{s} L_i(n)$. The aim of this section is to prove that the sets $L_{\leq s}(n)$ have δ_r-shadow of a minimal size. This follows directly from the following result.

Proposition 32. *Let $A \subseteq [k]^n$ and let $A_s = A \cap L_s(n)$. Then we have*

$$|\delta_r(A)| \geq \frac{1}{n(r+1)} \sum_{s=0}^{n} s\,|A_s|\,.$$

Proof. Let r be a given integer with $0 \leq r \leq k-1$. Let $X = [k]^n$, $Y = [k]^{n-1}$, let H be defined as in Section 3.4 and let $\mathcal{H}$ be a bipartite multigraph on $X \cup Y$ whose edges are given as follows. For each $x \in X \cap L_s(n)$ there are exactly s coordinates $x_{i_1}, \ldots, x_{i_s}$ which are in $\{0, \ldots, r\}$.

Define y_j to be the element obtained by deleting the coordinate x_{i_j}. Then we certainly have $y_j \in \delta_r(\{x\})$, and some of the points y_j may be equal. Define the edges of $\mathcal{H}$ to be the edges xy_j for all $1 \leq j \leq s$ counted with multiplicities. For example, when $r = 1$ the point $x = 00121$ is connected with 0121 by two edges, and with both 0012 and 0021 by one edge.

It is easy to check that for all $y \in Y$ the degree of y is $n(r+1)$, as this corresponds to adding an element in $\{0, \ldots, r\}$ to any of the n possible places in the sequence y. Note that for all $x \in X$ we have $\Gamma_{\mathcal{H}}(x) = \delta_r(\{x\})$, and hence for any $A \subseteq X$ we have $\delta_r(A) = \Gamma_{\mathcal{H}}(A)$. By the construction of $\mathcal{H}$ we have $d(x) = s$ for all $x \in L_s(n)$, and as observed earlier we have $d(y) = n(r+1)$ for all $y \in Y$. Since the connected components of $\mathcal{H}$ are contained in the sets $L_s(n) \cup L_{s-1}(n-1)$, we have $\Gamma_{\mathcal{H}}(A) \cap L_{s-1}(n-1) = \Gamma_{\mathcal{H}}(A \cap L_s(n))$. Therefore it follows that

$$|\Gamma_{\mathcal{H}}(A)| = \sum_{s=0}^{r} |\Gamma_{\mathcal{H}}(A_s)|. \tag{3.16}$$

For a set $B \subseteq L_s(n)$ we have

$$s|B| = e(B, \Gamma_{\mathcal{H}}(B)) \leq e(\Gamma_{\mathcal{H}}(B), X) = |\Gamma_{\mathcal{H}}(B)|\, n(r+1),$$

and hence it follows that

$$|\Gamma_{\mathcal{H}}(B)| \geq \frac{s}{n(r+1)} |B|. \tag{3.17}$$

Applying (3.17) to each term in the sum (3.16), we obtain that

$$|\delta_r(A)| = |\Gamma_{\mathcal{H}} A| \geq \frac{1}{n(r+1)} \sum_{s=0}^{r} s|A_s|, \tag{3.18}$$

which completes the proof. □

Now we are ready to conclude that the sets $L_{\leq s}(n)$ have δ_r-shadow of a minimal size.

Corollary 33. *Let A be a subset of $[k]^n$ satisfying $|A| = |L_{\leq s}(n)|$. Then we have $|\delta_r(A)| \geq |\delta_r(L_{\leq s}(n))|$, and the equality holds if and only if $A = L_{\leq s}(n)$.*

Proof. Let $B = L_{\leq s}(n)$. We start by checking that the equality holds for B in (3.18). Note that we have $B_i = L_i(n)$ for all $i \leq s$ and $B_i = \emptyset$ for all $i > s$. For $i \leq s$, we have

$$|B_i| = |L_i(n)| = \binom{n}{i} (r+1)^i (k - (r+1))^{n-i}$$

and

$$|\delta_r(B_i)| = |L_{i-1}(n-1)| = \binom{n-1}{i-1} (r+1)^{i-1} (k - (r+1))^{n-i}.$$

Therefore, we have $|\delta_r(B_i)| = \frac{i}{n(r+1)} |B_i|$ for each $i \leq s$, and in fact, this also holds for each $i > s$ as in this case both sides are 0. Hence the equality holds in (3.17) for all i, and thus the equality holds in (3.18) as well.

Given a set A of a fixed size for which we have $|A_i| \leq |L_i(n)|$ for all i, it is not too hard to check that the quantity $\frac{1}{n(r+1)} \sum_{t=0}^{r} t|A_t|$ is minimised if and only if $A = L_{\leq r}(n) \cup B$ for a suitably chosen r and for any $B \subseteq L_{r+1}(n)$ of a suitable size. Given that the size of A

satisfies $|A| = |L_{\leq s}(n)|$, the quantity $\frac{1}{n(r+1)} \sum_{t=0}^{r} t\,|A_t|$ attains its minimum value uniquely when $A = L_{\leq s}(n)$.

Hence it follows that

$$|\delta_r(A)| \geq \frac{1}{n(r+1)} \sum_{t=0}^{r} t\,|A_t| \geq \frac{1}{n(r+1)} \sum_{t=0}^{s} t\,|L_t(n)| = |\delta_r(L_{\leq s}(n))|, \tag{3.19}$$

and the second inequality holds if and only if $A = L_{\leq s}(n)$. □

Chapter 4

A grid generalisation of the Kruskal-Katona theorem

4.1 Introduction

Recall that the *lower shadow* of $A \subseteq \{0,1\}^n$ is defined to be the set of points that can be obtained by flipping exactly one 1-entry to 0 from a point in A, and it is denoted by $\partial^- A$. Similarly, the *upper shadow* of A is defined to be the set of points that can be obtained by flipping exactly one 0-entry to 1 from a point in A, and it is denoted by $\partial^+ A$. For $x \in \{0,1\}^n$, recall that in the previous chapter we defined the rank of x to be the sum of its coordinates. However, for the purposes of this chapter, it turns out to be more convenient to define the rank $w(x)$ of a $\{0,1\}$-sequence x by setting $w(x) = |\{i : x_i \geq 1\}|$. Even though this notion of rank coincides with the previous notion for $\{0,1\}$-sequences, these two notions will differ for general k. We choose to also call this new notion the rank given that it plays similar role with respect to the shadow operator considered in this chapter. As before, we define $\{0,1\}^n_r = \{x \in \{0,1\}^n : w(x) = r\}$.

For a given r, recall that the Kruskal-Katona theorem [25, 28] states that for a set $A \subseteq \{0,1\}^n_r$ of a given size, the size of the lower shadow of A is minimised when A is chosen to be the initial segment of the colexicographic order. Given $A \subseteq \{0,1\}^n_r$, recall that $\overline{A}$ is defined by setting $\overline{A} = \{x^c : x \in A\}$. Recall from Chapter 2 that we have $\left|\overline{A}\right| = |A|$, $\overline{A} \subseteq \{0,1\}^n_{n-r}$ and $\partial^+ A = \overline{\partial^- \overline{A}}$, and hence it follows that $|\partial^+ A| = \left|\partial^- \overline{A}\right|$. Thus there is a close connection between the lower shadow and the upper shadow, and in particular the Kruskal-Katona theorem implies that the size of the upper shadow of A is minimised when A is chosen to be an initial segment of the lexicographic order.

There are many natural generalisations of the lower shadow and the upper shadow for points in $[k]^n$, and one such generalisation can be obtained in the following way. Define the *d-shadow* of a point $x \in [k]^n$ to be the set of points obtained by flipping one of the coordinates of x that is in $\{1, \ldots, k-1\}$ to 0. Denote the d shadow of a point x by $d(\{x\})$. For $A \subseteq [k]^n$, define the d-shadow of A by setting $d(A) = \bigcup_{x \in A} d(\{x\})$. For example, we have $d(\{012\}) = \{002, 010\}$ and $d(\{000\}) = \emptyset$. It is clear that this operator agrees with the lower shadow operator when $k = 2$.

Define the *rank* of a point $x \in [k]^n$ to be $w(x) = |\{i : x_i \geq 1\}|$. For $0 \leq r \leq n$ set $[k]^n_r = \{x \in [k]^n : w(x) = r\}$. Note that the rank of a point in $d(\{x\})$ is one lower than the rank of x.

There is a superficial resemblance to a result of Clements as we now describe. Define the d^+*-shadow* of a point $x \in [k]^n$ by setting $d^+(\{x\})$ to be the set of points obtained from x by changing one of the coordinates of x which equals 0 to any number in $\{1, \dots, k-1\}$, and we set $d^+(A) = \bigcup_{x \in A} d^+(\{x\})$. Again, it is clear that the rank of a point in $d^+(\{x\})$ is one larger than the rank of x.

Clements [13] found an order on $[k]_r^n$ whose initial segments minimise the size of the d^+-shadow. Recall that the ordinary lower shadow and upper shadow can be related to each other by using the fact that $\partial^+ A = \overline{\partial^- \overline{A}}$. However, for $k \geq 3$, it is clear that there is no similar natural relation between the d^+-shadow and the d-shadow. That is, given the Clements' result for the d^+-shadow, there seems to be no way to deduce results for the d-shadow.

There is also a superficial resemblance to the Clements-Lindström Theorem [14]. Let $k_1, \dots, k_n$ be integers such that $1 \leq k_1 \leq \dots \leq k_n$, and let F be the set of all integer points $(a_1, \dots, a_n)$ with $0 \leq a_i \leq k_i$ for all i. Define the shadow operator Γ by setting

$$\Gamma((a_1, \dots a_n)) = \{(a_1 - 1, a_2 \dots, a_n), (a_1, a_2 - 1, \dots, a_n) \dots, (a_1, a_2 \dots, a_n - 1)\} \cap F,$$

and $\Gamma(A) = \bigcup_{a \in A} \Gamma(a)$ for $A \subseteq F$. Let F_r be the set of points $(a_1, \dots, a_n) \in F$ with $\sum_{i=1}^n a_i = r$. Generalise the colexicographic order by writing $(a_1, \dots, a_n) <_c (b_1, \dots, b_n)$ if there exists i such that $a_j = b_j$ for all $j > i$ and $a_i < b_i$. The Clements-Lindström theorem states that the initial segments of the colexicographic order minimise the size of the Γ-shadow on F_r.

The aim of this chapter is to find an order on $[k]_r^n$ whose initial segments minimise the size of the d-shadow among the subsets of a given size. In fact, we do this by first solving the unrestricted version, i.e. we find an order on $[k]^n$ whose initial segments minimise the size of the d-shadow among the subsets of a given size. Once we have proved the result for $[k]^n$, the result for $[k]_r^n$ follows easily. We have recently learnt that our result may also be deduced from a result of Frankl, Füredi and Kalai [19], and of London [32]; in this chapter we provide a new proof of their result.

We start by defining the order whose initial segments minimise the size of the d-shadow for the unrestricted version. For each i define $R_i(x) = \{j : x_j = i\}$. For a fixed k, define the order $\leq$ on $[k]^n$ by setting $x \leq y$ if $x = y$ or one of the following conditions holds.

1. $|R_0(x)| > |R_0(y)|$
2. $|R_0(x)| = |R_0(y)|$ and for the largest index i satisfying $R_i(x) \neq R_i(y)$ we have $\max(R_i(x) \Delta R_i(y)) \in R_i(y)$.

As usual, we say that $x < y$ if $x \leq y$ and $x \neq y$. Now we are ready to state the unrestricted version of our theorem.

Theorem 34. *Let A be a subset of $[k]^n$ and let B be an initial segment of the $\leq$-order on $[k]^n$ of size $|A|$. Then we have $|d(A)| \geq |d(B)|$.*

The proof of Theorem 34 is an inductive proof. As the proof of Theorem 34 is fairly long, we split the proof it into four subsections. In the first subsection, we introduce certain codimension-1 compression operators, and we prove that they cannot increase the size of the shadow. In particular, because of the compression operators, it suffices to prove Theorem 34 for those sets A that are stable under the compression operators.

Compression operators have been much utilised previously, see e.g. [7, 27]. For example, there are very straightforward proofs using compression operators for Harper's vertex-isoperimetric inequality on the hypercube and for the edge-isoperimetric inequality on the hypercube. In these examples, it is straightforward to prove the desired isoperimetric inequality for the sets that are stable under the compression operators. However, proving the inequality for the sets that are stable under the compression operators is highly nontrivial in our main theorem. In particular, the main part of the proof consists of dealing with the sets that are stable under the compression operators.

Even though $n = 1$ is the base case in the proof, it turns out to be convenient to consider the case $n = 2$ individually as well. Hence we consider the case $n = 2$ in the second subsection. This special case turns out to be reasonably straightforward when restricted to the sets that are stable under the compression operators.

We start the proof of the general case $n \geq 3$ by making some observations on the structure of the sets that are stable under the compression operators. As an example, we prove that it suffices to restrict our attention to the sets A for which there exists r such that A contains all the points with at least $r+1$ zeroes and no point with at most $r-1$ zeroes, and which are also stable under the compression operators. Let D be the set of points in A with exactly r zeroes. Since the d-shadow operator preserves the rank, it suffices to focus on analysing $d(D)$. The fourth subsection is dedicated to analysing $d(D)$ by using the previous and some new structural observations. At this stage of the proof, we need to split the proof into multiple subcases depending on the size and the structure of A.

Denote the restriction of the $\leq$-order on $\mathbb{N}_r^n = \{x = x_1 \dots x_n : w(x) = r\}$ also by $\leq$, where we define $\mathbb{N} = \{0, 1, \dots\}$. This order is well-defined, as it is easy to check that $[m]_r^n$ is an initial segment of the restriction of the $\leq$-order on $[k]_r^n$ for all $k \geq m$, and these orders coincide on $[m]_r^n$. Now we can state our main theorem.

Theorem 35. *Let A be a subset of $[k]_r^n$ and let B be an initial segment of the $\leq$-order on $[k]_r^n$ of size $|A|$. Then we have $|d(A)| \geq |d(B)|$.*

We end this section by introducing some notation. As before, we write $X = \{1, \dots, n\}$, $X^{(r)} = \{A \subseteq X : |A| = r\}$ and $X^{(\leq r)} = \{A \subseteq X : |A| \leq r\}$. For convenience, we often write

$$B_r = [k]_{n-r}^n = \{x \in [k]^n : |R_0(x)| = r\}$$

and $B_{\geq r} = \bigcup_{i=r}^n B_i$ for the set of points containing exactly r coordinates that are zeroes and for the set of points containing at least r coordinates that are zeroes respectively. Note that B_r depends on the values of n and k, but since their values are often clear from the context, the dependence is not highlighted. For $x \in [k]^n$, we write $d(x)$ instead of $d(\{x\})$.

For $a_i \in [k]$ and for positive integers $t_i \in \mathbb{N}$ we define $(t_1 \cdot a_1)(t_2 \cdot a_2) \dots (t_r \cdot a_r)$ to represent the element $a_1 \dots a_1 a_2 \dots a_2 \dots a_r \dots a_r$ which has t_1 a_1's immediately followed by t_2 a_2's, and so on. For example, we have $(3 \cdot 0)(2 \cdot 4)56 = 0004456$.

For $y = y_1 \dots y_{n-1} \in [k]^{n-1}$, $s \in X$ and $t \in [k]$ define

$$t_s y = y_1 \dots y_{s-1} t y_s \dots y_{n-1},$$

and for $Y \subseteq [k]^{n-1}$ we write $t_s Y = \{t_s y : y \in Y\}$. Note that we have $t_s x \leq t_s y$ if and only if $x \leq y$.

Finally, define the *binary order* $\leq_{bin}$ on $\mathcal{P}(\mathbb{N})$ by setting $X \leq_{bin} Y$ if $X = Y$ or $\max(X \Delta Y) \in Y$. We write $X <_{bin} Y$ if $X \leq_{bin} Y$ and $X \neq Y$. Note that the second condition in the definition of the $\leq$-order is equivalent to saying that for the largest index i satisfying $R_i(x) \neq R_i(y)$ we have $R_i(x) <_{bin} R_i(y)$.

4.2 Proof of Theorem 34

For convenience, we say that a set $A \subseteq [k]^n$ is *extremal* if the size of $d(A)$ is minimal among the subsets of $[k]^n$ of the same size. The proof of Theorem 34 is an inductive proof, and note that the result is trivial when $n = 1$. Throughout this section we assume that Theorem 34 holds for $[k]^{n-1}$, and our aim is to prove it for $[k]^n$.

4.2.1 The compression operators

For $A \subseteq [k]^n$, $t \in [k]$ and $s \in X$ define

$$A_{s,t} = \left\{ y \in [k]^{n-1} : t_s y \in A \right\}.$$

Let $B_{s,t} \subseteq [k]^{n-1}$ be the initial segment of the $\leq$-order of size $|A_{s,t}|$, and set $C_{s,t} = t_s B_{s,t}$. Define the *C_s-compression* of A by setting $C_s(A) = \bigcup_{t=0}^{k-1} C_{s,t}$. We start by proving that the C_s-compression operators cannot increase the size of the d-shadow.

Claim 1. For all $A \subseteq [k]^n$ and $s \in X$ we have $|C_s(A)| = |A|$ and $|d(A)| \geq |d(C_s(A))|$.

Proof of Claim 1. For a given s, note that the sets $C_{s,t}$ are pairwise disjoint for $t \in [k]$, as every $x \in C_{s,t}$ satisfies $x_s = t$. Since $|C_{s,t}| = |A_{s,t}|$ for all $t \in [k]$, it follows that $|C_s(A)| = |A|$.

Note that we have

$$d(C_s(A)) = d\left(\bigcup_{t=0}^{k-1} t_s B_{s,t}\right) = \left(\left(\bigcup_{t=1}^{k-1} 0_s B_{s,t}\right) \cup 0_s d(B_{s,0})\right) \cup \left(\bigcup_{t=1}^{k-1} t_s d(B_{s,t})\right), \tag{4.1}$$

and similarly we have

$$d(A) = d\left(\bigcup_{t=0}^{k-1} t_s A_{s,t}\right) = \left(\left(\bigcup_{t=1}^{k-1} 0_s A_{s,t}\right) \cup 0_s d(A_{s,0})\right) \cup \left(\bigcup_{t=1}^{k-1} t_s d(A_{s,t})\right). \tag{4.2}$$

Observe that the k sets

$$\left(\bigcup_{t=1}^{k-1} 0_s B_{s,t}\right) \cup 0_s d(B_{s,0}),\ 1_s d(B_{s,1}), \dots, (k-1)_s\, d(B_{s,k-1})$$

are pairwise disjoint as their s^{th} coordinates are distinct. Hence we have

$$\left|d\left(\bigcup_{t=0}^{k-1} t_s B_{s,t}\right)\right| = \left|\left(\bigcup_{t=1}^{k-1} 0_s B_{s,t}\right) \cup 0_s d\left(B_{s,0}\right)\right| + \sum_{t=1}^{k-1} \left|t_s d\left(B_{s,t}\right)\right|, \tag{4.3}$$

and similarly we have

$$\left|d\left(\bigcup_{t=0}^{k-1} t_s A_{s,t}\right)\right| = \left|\left(\bigcup_{t=1}^{k-1} 0_s A_{s,t}\right) \cup 0_s d\left(A_{s,0}\right)\right| + \sum_{t=1}^{k-1} \left|t_s d\left(A_{s,t}\right)\right|. \tag{4.4}$$

Since Theorem 34 holds for $[k]^{n-1}$, it follows that for any s and t we have

$$\left|t_s d\left(A_{s,t}\right)\right| = \left|d\left(A_{s,t}\right)\right| \geq \left|d\left(B_{s,t}\right)\right| = \left|t_s d\left(B_{s,t}\right)\right|. \tag{4.5}$$

Note that the image of an initial segment under the d-shadow is also an initial segment. Since initial segments are nested, we have

$$\left|\left(\bigcup_{t=1}^{k-1} 0_s B_{s,t}\right) \cup 0_s d\left(B_{s,0}\right)\right| = \max\left(\left|B_{s,1}\right|, \ldots, \left|B_{s,k-1}\right|, \left|d\left(B_{s,0}\right)\right|\right). \tag{4.6}$$

Combining the trivial estimate

$$\left|\left(\bigcup_{t=1}^{k-1} 0_s A_{s,t}\right) \cup 0_s d\left(A_{s,0}\right)\right| \geq \max\left(\left|A_{s,1}\right|, \ldots, \left|A_{s,k-1}\right|, \left|d\left(A_{s,0}\right)\right|\right) \tag{4.7}$$

with (4.5), (4.6) and the fact that $\left|A_{s,i}\right| = \left|B_{s,i}\right|$ for all $0 \leq i \leq k-1$, it follows that

$$\left|\left(\bigcup_{t=1}^{k-1} 0_s A_{s,t}\right) \cup 0_s d\left(A_{s,0}\right)\right| \geq \left|\left(\bigcup_{t=1}^{k-1} 0_s B_{s,t}\right) \cup 0_s d\left(B_{s,0}\right)\right|. \tag{4.8}$$

Thus pairing up the terms in (4.3) and (4.4) in the natural way and applying (4.5) and (4.8), it follows that

$$\left|d\left(A\right)\right| \geq \left|d\left(C_s\left(A\right)\right)\right|, \tag{4.9}$$

which completes the proof. □

We say that $T \subseteq [k]^n$ is *compressed* if for every $s \in X$ we have $C_s\left(T\right) = T$. We now make the standard observation that it suffices to prove Theorem 34 for compressed sets.

Claim 2. Let A be a subset of $[k]^n$. Then there exists a compressed set $B \subseteq [k]^n$ of size $|A|$ for which we have $\left|d\left(A\right)\right| \geq \left|d\left(B\right)\right|$.

Proof of Claim 2. Consider a sequence $\left(A_m\right)$ with $A_0 = A$ obtained as follows. Given A_m, if there exists $s \in X$ so that $C_s\left(A_m\right) \neq A_m$, we set $A_{m+1} = C_s\left(A_m\right)$. Otherwise, we set $A_{m+1} = A_m$.

Let K_i be the i^{th} set in $[k]^n$ with respect to the $\leq$-order. As in Chapter 3, define $f\left(A\right) = \sum_{i=1}^{k^n} i\mathbb{I}\left\{K_i \in A\right\}$, where $\mathbb{I}\left\{K_i \in A\right\}$ denotes the indicator function of the event $K_i \in A$. By the construction of the compression operator C_s, it is easy to verify that we have $f\left(C_s\left(A\right)\right) \leq f\left(A\right)$ for all $s \in X$, and for a given s the equality holds if and only if we have $C_s\left(A\right) = A$. Since $f\left(A\right)$

is always a non-negative integer, it follows that the sequence $f(A_m)$ is eventually constant. Thus there exists r for which the condition $f(C_s(A_r)) = f(A_r)$ is satisfied for all s, and hence A_r is compressed.

By Claim 1, for all $0 \le i \le r-1$ we have $|d(A_i)| \ge |d(A_{i+1})|$. Hence it follows that $|d(A)| \ge |d(A_r)|$, and thus we may take $B = A_r$. $\square$

From now on, let $A \subseteq [k]^n$ denote an arbitrary compressed set, and let C denote the initial segment of size $|A|$ on $[k]^n$. By Claim 2, it suffices to prove that $|d(A)| \ge |d(C)|$.

4.2.2 The special case $n = 2$.

Before moving on to the general case, we prove that Theorem 34 holds when $n = 2$. This turns out to be convenient as $n = 2$ is too small in one part of the general argument.

Claim 3. Theorem 34 holds for a compressed set $A \subseteq [k]^2$.

Proof of Claim 3. The claim is trivial if $|A| = 1$. Let C denote the initial segment of $[k]^2$ of size $|A|$. If $2 \le |A| \le 2k-1$, then C is a subset of $B_{\ge 1}$, and hence it follows that $d(C) = \{00\}$. Thus we evidently have $|d(A)| \ge |d(C)|$.

Now suppose that $|A| \ge 2k$. Write A as $A = A_0 \cup X$, where $A_0 = A \cap B_{\ge 1}$ and

$$X = A \setminus A_0 = \{x_1x_2 \in A : x_1 \neq 0 \text{ and } x_2 \neq 0\}.$$

Since $|B_{\ge 1}| = 2k-1$, it follows that $|X| \ge |A| - 2k + 1$, and in particular X is non-empty.

Let $x_1, \ldots, x_r \in \{1, \ldots, k-1\}$ and $y_1, \ldots, y_s \in \{1, \ldots, k-1\}$ be chosen so that

$$d(X) = \{0x_1, \ldots, 0x_r, y_10, \ldots, y_s0\}.$$

Then we certainly have

$$X \subseteq \{y_jx_i : 1 \le j \le s,\ 1 \le i \le r\},$$

which implies that $|X| \le rs$. Since for non-negative integers r and s we have $r + s \ge \left\lceil \sqrt{4rs} \right\rceil$, it follows that $|d(X)| \ge \left\lceil \sqrt{4|X|} \right\rceil$.

Since A is compressed and $|A| > 1$, it follows that A contains a point $x_1x_2 \neq 00$ with $x_1 = 0$ or $x_2 = 0$. In particular, we must have $00 \in d(A)$. Hence it follows that $d(A) = \{00\} \cup d(X)$. As $|X| \ge |A| - 2k + 1$, we have

$$|d(A)| \ge 1 + \left\lceil \sqrt{4(|A| - 2k + 1)} \right\rceil.$$

If $r^2 + 2k \le |A| \le r^2 + r + 2k - 1$ for some $1 \le r \le k-2$, it is easy to verify that C satisfies $C \subseteq B_{\ge 1} \cup \{x_1x_2 : 1 \le x_1 \le r+1,\ 1 \le x_2 \le r\}$. In particular, we have

$$|d(C)| \le 1 + (r+1) + r = 1 + \left\lceil \sqrt{4(|A| - 2k + 1)} \right\rceil,$$

and hence it follows that $|d(A)| \ge |d(C)|$. If $r^2 + r + 2k \le |A| \le (r+1)^2 + 2k - 1$ for some $1 \le r \le k-2$, it is easy to verify that C satisfies $C \subseteq B_{\ge 1} \cup \{x_1x_2 : 1 \le x_1, x_2 \le r+1\}$. Again, we have

$$|d(C)| \leq 1 + 2(r+1) = 1 + \left\lceil \sqrt{4(|A| - 2k + 1)} \right\rceil.$$

Thus in either case we have $|d(A)| \geq |d(C)|$. □

4.2.3 General observations

From now on, we assume that $n \geq 3$, and our aim is to prove that Theorem 34 holds for a compressed set $A \subseteq [k]^n$. As in the previous section, the proof is trivial when $|A| = 1$ or $2 \leq |A| \leq n(k-1) + 1$, as in these cases the d-shadow of the initial segment has size 0 or 1 respectively. Thus we may assume that $|A| \geq n(k-1) + 2$.

We say that $B \subseteq [k]^n$ is a *down-set* if for any points $y \in B$ and $x \in [k]^n$ satisfying $x_j \leq y_j$ for all j we have $x \in B$. In this subsection, we make some observations on the structure of a compressed set A. We start by proving that a compressed set is also a down-set, and in fact, this follows from a slightly stronger statement. Secondly, recall that for an initial segment C there exists r satisfying $B_{\geq r+1} \subseteq C \subseteq B_{\geq r}$. For a compressed set the same conclusion does not necessarily hold, but as a second structural claim we prove that for a compressed set A there exists r which satisfies $d(B_{\geq r+1}) \subseteq d(A) \subseteq d(B_{\geq r})$. Since we only consider the size of the d-shadow of A, in a sense this is 'equally good' for our purposes as having $B_{\geq r+1} \subseteq A \subseteq B_{\geq r}$. Note that the second observation allows us to focus only on $d(A \cap B_r)$.

Before proving that A and $d(A)$ are down-sets, we start with a straightforward observation that turns out to be even more useful. Given points x, y with $x \leq y$ and $y \in A$, suppose that there exists an index i satisfying $x_i = y_i$, and for convenience set $t = x_i$. Since A is compressed, it follows that $A_{i,t}$ is an initial segment. Let x' and y' be the points obtained from x and y by removing their i^{th} coordinate, and note that we have $y' \in A_{i,t}$. Since $x \leq y$ and $x_i = y_i$, it follows that $x' \leq y'$. Since A is compressed and $y' \in A_{i,t}$, it follows that $x' \in A_{i,t}$ and hence we also have $x \in A$. In summary, we have proved that

$$x \leq y,\ y \in A,\ x_i = y_i \text{ for some } i \Rightarrow x \in A. \tag{4.10}$$

This turns out to be a very useful fact that we will use throughout this chapter. As a simple consequence, we now prove that A and $d(A)$ are down-sets.

Claim 4. Let $A \subseteq [k]^n$ be a compressed set. Then both A and $d(A)$ are down-sets.

Proof of Claim 4. Let $y \in A$ and let $x \in [k]^n$ be a point satisfying $x_i \leq y_i$ for all i. Let $z \in [k]^n$ be obtained by taking $z_1 = x_1$ and $z_s = y_s$ for all $s \geq 2$, and note that we certainly have $x \leq z \leq y$. Since $y_2 = z_2$ and $x_1 = z_1$, by using (4.10) we first obtain that $z \in A$, and applying (4.10) again we obtain that $x \in A$. This completes the proof of the first part.

Let $y \in d(A)$ and let $x \in [k]^n$ be a point satisfying $x_i \leq y_i$ for all i. Choose a point $v \in A$ for which we have $y \in d(v)$, and let a be the unique index for which we have $v_a \neq 0$ and $y_a = 0$. Let u be the point obtained by setting $u_a = v_a$ and $u_j = x_j$ for all $j \neq a$. Then for any $j \neq a$ we have $u_j = x_j \leq y_j = v_j$. Since we also have $u_a = v_a$, the first part implies that $u \in A$. Since $x_a \leq y_a = 0$, we must have $x_a = 0$. Hence it follows that $x \in d(u) \subseteq d(A)$, which completes the proof. □

Before moving on to the second structural fact, we introduce some notation that will be used throughout the rest of this section. For $x = x_1 \dots x_n \in \mathbb{N}^n$, set $m(x) = \max(x_1, \dots, x_n)$ and recall that for all $i \in \mathbb{N}$ we defined $R_i(x) = \{j : x_j = i\}$. Define

$$c(x) = \left(m(x), R_{m(x)}(x), |R_0(x)|\right).$$

That is, the first coordinate of $c(x)$ is $\max(x_1, \dots, x_n)$, the second coordinate is the set of all positions where this maximal value is attained and the last coordinate is the number of x_i's that equal 0. Define the *component* of x by setting $C_x = \{y \in [k]^n : c(y) = c(x)\}$.

Note that we have $x \in [k]^n$ if and only if $C_x \subseteq [k]^n$, as the maximal value among the coordinates is the same for any $y \in C_x$. Moreover, since every $y \in C_x$ also shares the same positions of the coordinates that attain the maximal value and has the same number of coordinates that equal 0, it follows that for any other class C_z either all the points in C_x occur before the points in C_z with respect to the $\leq$-order, or all the points in C_x occur after the points in C_z. Hence we can order the classes inside $[k]^n$ as $C_1, \dots, C_m$ such that whenever $i \neq j$, for points $x \in C_i$ and $y \in C_j$ we have $x \leq y$ if and only if $i < j$.

Note that for all $x, y \in C_i$ there exists r satisfying $x_r = y_r$. Indeed, any $r \in R_{m(x)}(x) = R_{m(y)}(y)$ works. Hence (4.10) implies that for any i we either have $A \cap C_i = \emptyset$, or there exists $y_i \in C_i$ such that $A \cap C_i = \{x \in C_i : x \leq y_i\}$.

For fixed s and t, the classes of the form $(s,\ A,\ t)$ for $A \in X^{(\leq n-t)}$ occur consecutively with respect to the $\leq$-order. Furthermore, from the definition of the $\leq$-order it is easy to verify that the classes of the form $(s,\ A,\ t)$ for given s and t occur in the order induced by the binary order on $X^{(\leq n-t)}$. In particular, if A_i and A_{i+1} are two consecutive sets under the binary order satisfying $|A_i| \leq n-t$ and $|A_{i+1}| \leq n-t$, then $(s,\ A_i, t)$ immediately precedes $(s,\ A_{i+1},\ t)$ in the order of classes.

Recall that B_s is defined to be the set of points with exactly s coordinates that equal 0, i.e.

$$B_s = \{x \in [k]^n : |R_0(x)| = s\},$$

and $B_{\geq s} = \bigcup_{i=s}^n B_i$. Note that if X is an initial segment of the $\leq$-order, there exists r such that $B_{\geq r+1} \subseteq X \subset B_{\geq r}$. Our next aim is to prove that for any compressed set A there exists r for which we have $d(B_{\geq r+1}) \subseteq d(A) \subseteq d(B_{\geq r})$.

Claim 5. Let A be a compressed set and let $0 \leq p \leq n$ be the minimal index satisfying $A \cap B_p \neq \emptyset$. Then we have $d(B_{\geq p+1}) \subseteq d(A) \subseteq d(B_{\geq p})$. In particular, if r is chosen such that $|B_{\geq r+1}| < |A| \leq |B_{\geq r}|$, it suffices to prove Theorem 34 for compressed sets A which satisfy $B_{\geq r+1} \subseteq A \subseteq B_{\geq r}$.

Proof of Claim 5. Let p be the minimal index for which we have $A \cap B_p \neq \emptyset$, and let u be the minimal point under the $\leq$-order in $A \cap B_p$. Since A is a down-set and $u \in B_p$, it follows that every coordinate of u must be either 0 or 1. Hence there exists a set X of size $n-p$ for which we have $u_i = \mathbb{I}\{i \in X\}$.

Let $x \in B_{\geq p+2} = d(B_{\geq p+1})$. If $R_0(x) \cap X \neq \emptyset$, choose $i \in R_0(x) \cap X$ and consider the point y obtained by taking $y_j = x_j$ for $j \neq i$ and $y_i = 1$. Since $i \in X$, it follows that $u_i = 1 = y_i$. On the other hand, since $|R_0(y)| = p+2-1 = p+1 > p = |R_0(u)|$, it follows that $y \leq u$. Hence (4.10) implies that $y \in A$, and hence it follows that $x \in d(A)$.

If $R_0(x) \cap X = \emptyset$, choose any $i \in R_0(x)$ and again consider the point y obtained by taking $y_j = x_j$ for $j \neq i$ and $y_i = 1$. Let $j \in R_0(x) \setminus \{i\}$, and note that such j exists as $|R_0(x)| \geq p+2 \geq 2$. Since $j \notin X$, it follows that $u_j = 0 = x_j = y_j$. Similarly as in the first case, we have $|R_0(y)| > |R_0(u)|$ and hence it follows that $y \leq u$. Hence (4.10) implies that $y \in A$, and therefore we have $x \in d(A)$, which completes the proof of the first part.

For the second part, suppose that A' is a compressed set satisfying $|B_{\geq r+1}| < |A'| \leq |B_{\geq r}|$ for some r. Let C denote the initial segment of the $\leq$-order of size $|A'|$, and assume that Theorem 34 holds for any compressed set $A \subseteq [k]^n$ satisfying the conditions $|A| = |A'|$ and $B_{\geq r+1} \subseteq A \subseteq B_{\geq r}$. Let p denote the least integer for which we have $A' \cap B_p \neq \emptyset$. Since $|A'| > |B_{\geq r+1}|$, it follows that $p \leq r$.

If $p < r$, the first part implies that we have $d(B_{\geq r}) \subseteq d(B_{\geq p+1}) \subseteq d(A')$, and since $C \subseteq B_{\geq r}$ it certainly follows that $|d(A')| \geq |d(C)|$. If $p = r$, the choice of p implies that we have $A' \subseteq B_{\geq r}$, and the first part implies that $d(B_{\geq r+1}) \subseteq d(A')$. Let $D' = A' \cap B_r$, and note that the last two observations imply that $d(A') = d(B_{\geq r+1}) \cup d(D')$. Let $C' = C \cap B_r$, and note that since $|A'| = |C|$ and $B_{\geq r+1} \subseteq C$, it follows that $|C'| \leq |D'|$. Let D be the set of $|C'|$ smallest points in D' under the $\leq$-order, and let $A = B_{\geq r+1} \cup D$. It is easy to see that A is compressed and we have $|A| = |C| = |A'|$. By the construction of A it follows that $B_{\geq r+1} \subseteq A \subseteq B_{\geq r}$, and since we have $D \subseteq D'$ and $d(B_{\geq r+1}) \subseteq d(A')$, it follows that $d(A) \subseteq d(A')$. In particular, we have $|d(A')| \geq |d(A)|$, which implies the result. □

4.2.4 Rest of the proof

Let A be a compressed set, let r be chosen so that $|B_{\geq r+1}| < |A| \leq |B_{\geq r}|$ and let C be the the initial segment of the $\leq$-order of size $|A|$. By Claim 5 we may assume that we have $B_{\geq r+1} \subseteq A \subseteq B_{\geq r}$. From now on, we set $D = A \cap B_r$. Since $|d(A)| = |B_{\geq r+2}| + |d(D)|$ and $|d(C)| = |B_{\geq r+2}| + |d(C \cap B_r)|$, it suffices to focus on comparing $d(D)$ with $d(C \cap B_r)$.

We split the proof into two cases depending on whether we have $|A| \leq |B_{\geq 1}|$ or $|A| > |B_{\geq 1}|$.

Case 1. $|A| \leq |B_{\geq 1}|$.

Since $|A| \leq |B_{\geq 1}|$, it follows that $A = B_{\geq r+1} \cup D$ for some $D \subseteq B_r$ and $r \geq 1$. Recall that Theorem 34 is trivial when $|A| \leq n(k-1)+1$, and hence we may assume that $r \leq n-2$.

Note that each of the sets $B_{\geq r+1} \cup ([s]^n \cap B_r)$ is an initial segment of the $\leq$-order, and recall that $[s]^n = \{0, \ldots, s-1\}^n$. We start by comparing the sets $d(D)$ and $d([s]^n \cap B_r)$ for an appropriately chosen s.

Claim 6. Let $s = \max\{m(x) : x \in D\} = \max\{x_i : x_1 \ldots x_n \in D\}$. Then we have $d([s]^n \cap B_r) \subseteq d(A)$. In particular, if r and s are chosen so that

$$|B_{\geq r+1} \cup ([s]^n \cap B_r)| < |A| \leq |B_{\geq r+1} \cup ([s+1]^n \cap B_r)|,$$

it suffices to prove Theorem 34 for compressed sets A satisfying

$$B_{\geq r+1} \cup ([s]^n \cap B_r) \subseteq A \subseteq B_{\geq r} \cup ([([s+1]^n \cap B_r)]).$$

Proof of Claim 6. When $s = 1$, the claim is equivalent to the condition $B_{\geq r+1} \subseteq d(A)$, which is certainly true by the choice of A. Now suppose that $s \geq 2$, and let x be the least point under the $\leq$-order in D with $m(x) = s$. Since A is a down-set and $s \geq 2$, the minimality of x implies that there exists a unique index i satisfying $x_i = s$.

Let $v \in d([s]^n \cap B_r)$. Since $|R_0(v)| = r + 1 \geq 2$, it follows that there exists $j \neq i$ for which we have $v_j = 0$. If $x_j \neq 0$, consider the point z obtained by setting

$$z_t = \begin{cases} v_t & \text{if } t \neq j \\ x_j & \text{if } t = j \end{cases}.$$

Otherwise, pick any $p \neq j$ with $v_p = 0$, and consider the point z obtained by setting

$$z_t = \begin{cases} v_t & \text{if } t \neq p \\ 1 & \text{if } t = p \end{cases}.$$

Since $v_t \in [s]$ for all t and j is chosen so that $x_j \neq s$, in either case we have $z \in [s]^n$. In both cases we certainly have $v \in d(z)$ by the construction of z, and we have $|R_0(x)| = |R_0(z)|$. Since $m(x) > m(z)$, it follows that $z \leq x$.

In the first case we have $x_j = z_j$, and hence (4.10) implies that $z \in A$. In the second case we have $z_j = v_j = x_j = 0$, so again (4.10) implies that $z \in A$. Hence in either case we have $v \in d(A)$, which completes the proof of the first part.

In order to prove the second part, suppose that A' is a compressed set satisfying

$$|B_{\geq r+1} \cup ([s]^n \cap B_r)| < |A'| \leq |B_{\geq r+1} \cup ([s+1]^n \cap B_r)|$$

for some r and s. Recall that by Claim 5 we may assume that $B_{\geq r+1} \subseteq A' \subseteq B_{\geq r}$. Let C be the initial segment of the $\leq$-order of size $|A'|$, and note that we have $C \subseteq B_{\geq r+1} \cup ([s+1]^n \cap B_r)$.

Let $t = \max(m(x) : x \in A' \cap B_r)$. Since $|A'| > |B_{\geq r+1} \cup ([s]^n \cap B_r)|$, we must have $t \geq s$. If $t \geq s + 1$, the first part implies that

$$d(B_{\geq r+1} \cup ([s+1]^n \cap B_r)) \subseteq d(A').$$

In particular, it follows that $d(C) \subseteq d(A')$, and thus we have $|d(A')| \geq |d(C)|$.

If $t = s$, it follows that $A' \subseteq B_{\geq r+1} \cup ([s+1]^n \cap B_r)$. Let $X = \{x \in A \cap B_r : m(x) = s\}$, and note that we have $X \neq \emptyset$. Let Y be the set of k least points in X under the $\leq$-order, where $k = |A'| - |B_{\geq r+1} \cup ([s]^n \cap B_r)|$, and consider A defined by $A = B_{\geq r+1} \cup ([s]^n \cap B_r) \cup Y$. It is clear that A is compressed, and we have $|A| = |A'| = |C|$.

By the construction of A it follows that $B_{\geq r+1} \cup ([s]^n \cap B_r) \subseteq A \subseteq B_{\geq r+1} \cup ([s+1]^n \cap B_r)$. By Claim 5 and the first part, we have $d(B_{\geq r+1} \cup ([s]^n \cap B_r)) \subseteq d(A')$. In particular, it follows that $d(B_{\geq r+1} \cup ([s]^n \cap B_r)) \cup d(X) \subseteq d(A')$. Since $d(Y) \subseteq d(X)$, we must have $d(A) = d(B_{\geq r+1} \cup ([s]^n \cap B_r)) \cup d(Y) \subseteq d(A')$. Hence it follows that $|d(A')| \geq |d(A)|$, which completes the proof. □

Let $s = \max(m(x) : x \in D)$. From now on, we suppose that A is a compressed set satisfying $B_{\geq r+1} \cup ([s]^n \cap B_r) \subset A \subseteq B_{\geq r+1} \cup ([s+1]^n \cap B_r)$.

We split our proof into two subcases based on whether $s = \max(m(x) : x \in D)$ satisfies $s = 1$ or $s \geq 2$. When $s = 1$, Theorem 34 is equivalent to the Kruskal-Katona Theorem, and hence the case $s \geq 2$ is the main part of the proof.

Case 1.1. $s = 1$.

Since $s = 1$, it follows that $D \subseteq \{0,1\}^n \cap B_r$, and note that we have $\{0,1\}^n \cap B_r = \{0,1\}^n_{n-r}$. It is easy to check that the colexicographic order and the $\leq$-order coincide on $\{0,1\}^n_{n-r}$. Hence Kruskal-Katona theorem implies that Theorem 34 holds when $s = 1$.

Case 1.2. $s \geq 2$.

Since $s \geq 2$, it follows that for any point x changing a value of a non-zero coordinate to 1 cannot increase the set $R_s(x)$. As a consequence of Claim 6, from now on, we only need to focus on the points in D that contain s as a coordinate.

Let $\mathcal{A} = X^{(\leq n-r)} \setminus \{\emptyset\}$, and for $X \in \mathcal{A}$ denote the class C_i corresponding to the tuple (s, X, r) by C_X. Note that $\mathcal{A}$ characterizes all the classes containing points with exactly r zeroes and s as a maximal coordinate. As noted before, the order of the classes under the $\leq$-order is the order induced on $\mathcal{A}$ by the binary order.

Let $T \in \mathcal{A}$ be the largest set under the binary order satisfying $C_T \cap D \neq \emptyset$. We start by proving that for any $S \in \mathcal{A}$ with $S <_{bin} T$ we have $d(C_S) \subseteq d(D)$. If $S \cap T \neq \emptyset$, the claim follows easily as $C_S \subseteq D$ holds by (4.10). We start by proving the claim when T contains only one point, as in this case such an easy argument does not exist and the stronger conclusion $C_s \subseteq D$ may not always hold. In Claim 8 we deduce that for any $S \in \mathcal{A}$ with $S <_{bin} T$ we have $d(C_S) \subseteq d(D)$ regardless of the size of T.

Claim 7. Let $X \in \mathcal{A}$ be a set satisfying $C_X \cap D \neq \emptyset$ and $|X| = 1$. Then for any $S \in \mathcal{A}$ with $S <_{bin} X$ we have $d(C_S) \subseteq d(D)$.

Proof of Claim 7. Let i be chosen so that $X = \{i\}$, and define the particular point

$$a = s_i(((n-r-1) \cdot 1)(r \cdot 0)).$$

Note that a is the least point in C_X, and hence (4.10) implies that $a \in D$.

Let $x \in d(C_S)$. Since $D \subseteq B_r$ for some $r \geq 1$, it follows that $|R_0(x)| = r + 1 \geq 2$. Hence there exist distinct elements l and m satisfying $x_l = x_m = 0$. In particular, we may assume that $m \neq i$.

Let y be the point obtained by taking $y_j = x_j$ for $j \neq m$ and $y_m = 1$, and let z be the point obtained by taking $z_j = x_j$ for $j \neq l$ and $z_l = 1$. Note that we have $R_s(y) \subseteq S$ and $R_s(z) \subseteq S$, and recall that $R_s(a) = X$. Since $S <_{bin} X$, these conditions imply that $y \leq a$ and $z \leq a$. Note that by the construction of the elements y and z we have $y_m = 1$ and $z_m = x_m = 0$. Since $m \neq i$, it follows that $a_m \in \{0,1\}$, and hence we either have $a_m = y_m$ or $a_m = z_m$. Thus (4.10) implies that we have $y \in D$ or $z \in D$, and in either case it follows that $x \in d(D)$. □

Claim 8. Let $T \in \mathcal{A}$ be the largest set under the binary order satisfying $C_T \cap D \neq \emptyset$. Then for any $S \in \mathcal{A}$ with $S <_{bin} T$ we have $d(C_S) \subseteq d(D)$.

Proof of Claim 8. The result follows immediately from Claim 7 when $|T| = 1$. Hence we may assume that $|T| > 1$, and define $T_1 = \{\max T\}$. Note that we have $T_1 \cap T \neq \emptyset$ and $T_1 <_{bin} T$. Hence (4.10) implies that $C_{T_1} \subseteq D$.

Let $S \in \mathcal{A}$ be a set satisfying $S <_{bin} T$. If $\max S = \max T$, then (4.10) implies that $C_S \subseteq D$, and in particular it follows that $d(C_S) \subseteq d(D)$. If $\max S < \max T$, it follows that $S <_{bin} T_1$. By applying Claim 7 for the set T_1 which contains only one point, we obtain that $d(C_S) \subseteq d(D)$, which completes the proof. $\square$

Now we are ready to finish the proof of Case 1.2. Let C be the initial segment of the $\leq$-order of size $|A|$. Recall that r and s are chosen so that $B_{\geq r+1} \cup ([s]^n \cap B_r) \subseteq A \subseteq B_{\geq r+1} \cup ([s+1]^n \cap B_r)$. Since C is an initial segment, it follows that there exist $T_1 \in \mathcal{A}$ and $x \in T_1$ so that

$$C = B_{\geq r+1} \cup ([s]^n \cap B_r) \cup \left(\bigcup_{S \in \mathcal{A},\, S <_{bin} T_1} C_S \right) \cup \{y : y \leq x, y \in C_{T_1}\}.$$

Since $|A| = |C|$, we must have $T_1 \leq_{bin} T$. If $T_1 < T$, we have $d(C_S) \subseteq d(A)$ for all S with $S \in \mathcal{A}$ and $S \leq_{bin} T_1$ by Claim 8. In particular, it follows that $d(C) \subseteq d(A)$, and therefore we have $|d(A)| \geq |d(C)|$.

Now suppose that $T_1 = T$. Let $z \in C_T$ be chosen so that we have $A \cap C_T = \{y : y \leq z, y \in C_T\}$. Since C is the initial segment of the $\leq$-order of size $|A|$, it follows that $x \leq z$. Hence we have $d(\{y : y \leq x, y \in T_1\}) \subseteq d(A)$, and it follows that $d(C) \subseteq d(A)$. This completes the proof of Case 1.2.

Case 2. $|A| > |B_{\geq 1}|$.

Recall that A is of the form $A = B_{\geq 1} \cup D$ for some $D \subseteq B_0 = \{1, \ldots, k-1\}^n$. Note that Theorem 34 holds when $|A| = |B_{\geq 1}| + 1$ as $|d(x)| = n$ for any $x \in B_0$. From now on, we will assume that $|A| > |B_{\geq 1}| + 1$. Hence it follows that $|D| > 1$, and thus $s = \max(m(x) : x \in D)$ satisfies $s \geq 2$.

Ideally, we would like to use an approach that is similar to the one used in the proof of Case 1. However, it turns out that some difficulties may arise with the appropriate versions of Claims 6 and 8. We start with a preliminary result which states that (4.10) also holds for points inside $d(D)$. Then we move on to prove an appropriate version of Claim 6. When $T \neq \{1\}$, we again have $d(\{1, \ldots, s-1\}^n) \subseteq d(A)$. When $T = \{1\}$, one specific point from $d(\{1, \ldots, s-1\}^n)$ might not be contained in $d(A)$.

Claim 9. Let $y \in d(D)$, and let x be a point satisfying the conditions $|R_0(x)| = 1$ and $x \leq y$. Furthermore, suppose that there exists an index i satisfying $x_i = y_i$. Then we have $x \in d(D)$.

Proof of Claim 9. Let x and y be points satisfying the conditions described above, and note that we may assume that $x < y$. Since $y \in d(D)$, there exists $v \in D$ for which $y \in d(v)$. Thus there exists b satisfying $v_j = y_j$ for $j \neq b$, where $y_b = 0$ and $v_b \neq 0$. Since A is a down-set, we may assume that $v_b = 1$. Note that b is the unique index for which we have $y_b = 0$ as $|R_0(y)| = 1$.

Since $|R_0(x)| = 1$, there exists a unique index a with $x_a = 0$. Let u be the element obtained by taking $u_j = x_j$ for $j \neq a$ and $u_a = 1$, and observe that we have $x \in d(u)$.

Our aim is to prove that $u \in D$. Let t be the largest index satisfying $R_t(x) \neq R_t(y)$, and note that for this choice of t we have $\max(R_t(x) \Delta R_t(y)) \in R_t(y)$. Observe that we must have $t \geq 1$, as for any point z the sets $R_t(z)$ are disjoint and their union equals $\{1, \ldots, n\}$.

First note that by the construction of the points u and v it follows that $R_s(u) = R_s(x)$ and $R_s(v) = R_s(y)$ for $s \geq 2$. In particular, it follows that $R_s(u) = R_s(v)$ for all $s > t$.

We start by proving that $t = 1$ implies $u = v$. Indeed, by the previous observation we have $R_s(u) = R_s(v)$ for all $s > 1$, and we also have $R_0(u) = R_0(v) = \emptyset$. Since for any point c the sets $R_s(c)$ are disjoint and their union equals $\{1, \ldots, n\}$, we must also have $R_1(u) = R_1(v)$. In particular, it follows that $u = v$.

If $t \geq 2$, it follows that $R_s(u) = R_s(x)$ and $R_s(v) = R_s(y)$ for all $s \geq t$. In particular, it follows that $\max(R_t(u) \Delta R_t(v)) \in R_t(v)$, and thus we have $u < v$.

Recall that there exists an index i satisfying $x_i = y_i$. If $i \notin \{a, b\}$, it follows that $u_i = x_i$ and $v_i = y_i$ by the construction of the points u and v. In particular, we have $u_i = v_i$, and thus (4.10) implies that $u \in D$.

Now suppose that $i \in \{a, b\}$. Since $x_a = 0$ and $y_b = 0$, we must have $x_i = y_i = 0$. Since both $R_0(x)$ and $R_0(y)$ are sets containing only one point, it follows that $a = b = i$. Thus by the construction of u and v it follows that $u_i = v_i = 1$. As before, (4.10) implies that $u \in D$, which completes the proof. □

Claim 10.

1. If $T \neq \{1\}$ we have $d(\{1, \ldots, s-1\}^n) \subseteq d(A)$.

2. If $T = \{1\}$ we have $d(\{1, \ldots, s-1\}^n) \setminus \{0((n-1) \cdot (s-1))\} \subseteq d(A)$.

Proof of Claim 10. We start with the case $T \neq \{1\}$, and recall that we have $s \geq 2$. Let $j = \max(T)$, which by assumption is at least 2. Define the particular point $a = s_j((n-1) \cdot 1)$ which is the least point in $C_{\{j\}}$. Let z be a point in $C_T \cap D$. Since $a \leq z$ and $a_j = z_j = s$, (4.10) implies that we have $a \in D$. Define the particular point $b = s_{j-1}((n-1) \cdot 1)$, and note that b is well-defined as $j > 1$. Since $2(n-1) > n$, there exists an index i satisfying $a_i = b_i = 1$ by the pigeonhole principle. Thus (4.10) implies that $b \in D$.

Let $x \in d(\{1, \ldots, s-1\})$, and let i be the unique index satisfying $x_i = 0$. Let y be obtained by taking $y_r = x_r$ for $r \neq i$ and $y_i = 1$. Since $s \geq 2$, it follows that $y \in \{1, \ldots, s-1\}^n$, and hence we have $y \leq a$ and $y \leq b$.

We have $a_i = y_i = 1$ whenever $i \neq j$, and we have $b_i = y_i = 1$ whenever $i = j$. Since $a, b \in D$, in either case we have $y \in D$ by (4.10). This completes the proof of the first part.

Now suppose that $T = \{1\}$. For convenience, define the points $w = 0((n-1) \cdot (s-1))$ and $a = s((n-1) \cdot 1)$. Note that a is the least point in C_T, and thus we have $a \in D$. For every $2 \leq i \leq n$ define the points $b_i = 1_i((n-1) \cdot (s-1))$. Then $(b_i)_i = 1 = a_i$ and $b_i \leq a$ for all i, so (4.10) implies that we have $b_i \in d(D)$ for all i. For $2 \leq i \leq n$ define the points $c_i = 0_i((n-1) \cdot (s-1))$, and note that we have $c_i \in d(b_i)$ for all i. Hence it follows that $c_i \in d(D)$ for all $2 \leq i \leq n$.

Let x be a point satisfying $x \in d(\{1, \ldots, s-1\}^n)$ and $x \neq w$. If $x_i = 0$ for some $i \geq 2$, we have $x_l \leq (c_i)_l$ for every l. Since $d(A)$ is a down-set by Claim 4, it follows that $x \in d(D)$.

Now suppose that x is a point with $x_1 = 0$ and $x \neq w$. Since $x \neq w$, it follows that $R_{s-1}(x)$ is a proper subset of $\{2, \ldots, n\}$. Note that $R_{s-1}(c_2) = \{1, 3, \ldots, n\}$ is larger than any proper subset of $\{2, \ldots, n\}$ under the binary order, and hence we have $x \leq c_2$. Let $k \in \{2, \ldots, n\} \setminus R_{s-1}(x)$, and consider v obtained by taking $v_l = (c_2)_l$ for $l \neq k$ and $v_k = x_k$. Then we have $x \leq v \leq c_2$, and hence applying Claim 9 twice implies that $x \in d(D)$. Thus we have $d(\{1, \ldots, s-1\}^n) \setminus \{w\} \subseteq d(D)$, which completes the proof. □

Define the particular points $w = 0((n-1) \cdot (s-1))$ and $a_i = i((n-1) \cdot (s-1))$ for $1 \leq i \leq s-1$. It is easy to verify that we have $w \in d(x)$ for $x \in [s]^n \cap B_0$ if and only if $x = a_i$ for some i. In order to deal with the case $T = \{1\}$, we prove that adding suitably many consecutive points to an initial segment must increase the size of the d-shadow. This is done in Claim 12, but first we need a preliminary result. For $x \in [k]^n$, define the point x^+ to be the successor of x, i.e. x^+ is the least point under the $\leq$-order satisfying $x < x^+$.

Claim 11. Let $x \in D$ and suppose that we also have $y = x^+ \in D$. Then the following claims are true.

1. There exists a unique index i satisfying $y_i > x_i$.
2. If $y_j < x_j$ for some j then we have $y_j = 1$.
3. If $y_j \geq 2$ for every j, we must have $y_j > x_i$ for every j, where i is the unique index satisfying $y_i > x_i$.

Proof of Claim 11. Let t be the largest index satisfying $R_t(x) \neq R_t(y)$, and since $x < y$ we must have $R_t(x) <_{bin} R_t(y)$. Let $i = \max(R_t(x) \Delta R_t(y))$, and since $i \in R_t(y)$ it follows that $y_i > x_i$.

Consider the point z obtained by taking $z_j = \min(x_j, y_j)$ for every $j \neq i$ and $z_i = y_i$. By the construction of the element z we have $z_j \leq y_j$ for each j, and hence we have $z \leq y$. For all $s > t$ we have $R_s(x) = R_s(y)$, and hence by the construction of z it follows that $R_s(x) = R_s(z)$ for all $s > t$. Similarly, it is easy to see that we have $R_t(x) \cap R_t(y) \subseteq R_t(z)$. Since we also have $i \in R_t(z)$, it follows that $\max(R_t(x) \Delta R_t(z)) = i \in R_t(z)$, which implies that $x < z$. Combining these two observations, we obtain that $x < z \leq y$, and since $y = x^+$ it follows that $y = z$. Thus for any $j \neq i$ we have $y_j = \min(x_j, y_j) \leq x_j$, which completes the proof of the first part.

Let $X = \{j : y_j < x_j\}$, and let u be the element obtained by taking $u_j = 1$ for every $j \in X$ and $u_j = y_j$ for every $j \notin X$. Since $y_j > 0$ for every j, we must have $u_j \leq y_j$ for every j, and hence we certainly have $u \leq y$. Let t and i be defined as in the previous part. As in the previous part, it is easy to check that we have $R_s(x) = R_s(u)$ for all $s > t$ and $\max(R_t(u) \Delta R_t(x)) = i$. Therefore we have $x < u$. Combining these two observations we obtain that $x < u \leq y$, and since $y = x^+$ it follows that $y = u$. Hence the condition $y_j < x_j$ implies that $y_j = 1$, which proves the second part.

Suppose that we have $y_j \geq 2$ for every j. Combining the first and the second part, it follows that there exists a unique i so that $y_j = x_j$ for all $j \neq i$ and $y_i > x_i$. Since y is the successor of x, it evidently follows that $y_i = x_i + 1$. Our aim is to prove that x_i is strictly smaller than any other x_j.

First assume that there exists k satisfying $x_k < x_i$. Consider the point v obtained by taking $v_j = x_j$ for $j \neq k$ and $v_k = x_k + 1$, and note that we certainly have $x < v$. For convenience, set $a = x_i$ and $b = x_k$. By the construction of v it follows that $R_s(x) = R_s(y) = R_s(v)$ for all $s \geq a+2$. Since $b < a$, it follows that $R_{a+1}(y) = R_{a+1}(x) \cup \{i\}$, while we have $R_{a+1}(v) = R_{a+1}(x)$. Hence we also have $v < y$, which contradicts the fact that y is the successor of x.

Again, let $a = x_i$, and note that when $a = 1$ the claim follows evidently. Now suppose that for all j we have $x_j \geq a$ and $|R_a(x)| \geq 2$. Let v be the point obtained by taking $v_j = x_j$ for all $j \notin R_a(x)$, $v_i = a+1$ and $v_j = 1$ for all $j \in R_a(x) \setminus \{i\}$. Note that for all $s \geq a+2$ we have $R_s(x) = R_s(v) = R_s(y)$. However, we also have $R_{a+1}(y) = R_{a+1}(v) = R_{a+1}(x) \cup \{i\}$, $R_a(y) = R_a(x) \setminus \{i\}$ and $R_a(v) = \emptyset$ since $a > 1$. Since $R_a(x)$ contains at least two elements, it follows that $x < v < y$, which contradicts the fact that $y = x^+$.

Hence for each $j \neq i$ we have $x_j > x_i$. Since $y_j \neq 1$ for all j, the second part implies that we have $y_j \geq x_j$ for all j. Hence for all $j \neq i$ we have $y_j \geq x_j > x_i$, and by the choice of i we have $y_i > x_i$. This completes the proof of the third part. $\square$

Claim 12. Let $x_1, \ldots, x_{L-1} \in \{1, \ldots, L-1\}^n$ be consecutive points under the $\leq$-order. Let X and Y be the initial segments of the $\leq$-order on $[k]^n$ defined by $X = \{y : y \leq x_{L-1}\}$ and $Y = \{y : y \leq x_1\}$. Then we have $|d(X)| = |d(Y)|$ if and only if $x_i = (i)((n-1) \cdot (L-1))$ for all $1 \leq i \leq L-1$.

Proof of Claim 12. Let $x \in \{1, \ldots, L-1\}^n$, and consider the initial segment Z defined by $Z = \{y : y < x\}$. We start by proving that $|d(Z \cup \{x\})| = |d(Z)|$ if and only if $R_1(x) = \emptyset$. Suppose that there exists an element $z \in d(Z \cup \{x\}) \setminus d(Z)$, and let j be the unique index satisfying $z_j = 0$. Let $a_1, \ldots, a_{L-1}$ be the points obtained by taking $(a_i)_k = z_k$ for $k \neq j$, and $(a_i)_j = i$. It is clear that we have $a_1 < \cdots < a_{L-1}$. Furthermore, for an element $u \in \{1, \ldots, L-1\}^n$ we have $z \in d(u)$ if and only if $u \in \{a_1, \ldots, a_{L-1}\}$. In particular, it follows that $z \in d(Z \cup \{x\}) \setminus d(Z)$ if and only if $Z \cap \{a_1, \ldots, a_{L-1}\} = \emptyset$ and $x \in \{a_1, \ldots, a_{L-1}\}$. Since Z is an initial segment and x is the least element in Z^c under the $\leq$-order, we must have $x = a_1$. Hence the condition $|d(Z \cup \{x\})| = |d(Z)|$ implies that $R_1(x) = \emptyset$. Conversely, if $x_i = 1$ for some i, consider the element u obtained by taking $u_j = x_j$ for $j \neq i$ and $u_i = 0$. By using a similar argument, it is easy to see that we have $u \in d(x) \setminus d(Z)$.

Let X and Y satisfy the conditions of the claim and suppose that we have $|d(X)| = |d(Y)|$. Then the previous observation implies that we have $R_1(x_i) = \emptyset$ for all $2 \leq i \leq L-1$.

Let $m(i)$ denote the value of the smallest coordinate of the point x_i. Since $R_1(x_i) = \emptyset$ for $2 \leq i \leq L-1$, the third part of Claim 11 implies that we have $(x_{i+1})_j > m(i)$ for all j. In particular, it follows that $m(i+1) > m(i)$, i.e. m is strictly increasing. Since $R_1(x_2) = \emptyset$, we have $m(2) \geq 2$, and hence we must also have $m(i) \geq i$ for all i. In particular, it follows that $m(L-1) \geq L-1$, and thus we must have $x_{L-1} = (n \cdot (L-1))$. Since the points x_i occur consecutively under the $\leq$-order, it is easy to verify that we have $x_i = (i)((n-1) \cdot (L-1))$ for all $1 \leq i \leq L-1$. $\square$

As a consequence of the following claim, it suffices to prove Theorem 34 only for compressed sets A satisfying the condition $B_{\geq 1} \cup ([s]^n \cap B_0) \subseteq A \subseteq B_{\geq 1} \cup ([s+1]^n \cap B_0)$ for some s.

Claim 13. Let A be a compressed set with $|A| \geq |B_{\geq 1}| + 2$, and let $s \geq 2$ be chosen so that $|B_{\geq 1} \cup ([s]^n \cap B_0)| < |A| \leq |B_{\geq 1} \cup ([s+1]^n \cap B_0)|$. Then there exists a compressed set B of size $|A|$ satisfying $B_{\geq 1} \cup ([s]^n \cap B_0) \subseteq B \subseteq B_{\geq 1} \cup ([s+1]^n \cap B_0)$ for which we have $|d(A)| \geq |d(B)|$.

Proof of Claim 13. Let $q = \max(m(x) : x \in D)$, and note that the condition on the size of A implies that we have $q \geq s$. Let T be the largest point under the binary order so that the class C_T has a non-empty intersection with A.

First suppose that $T = \{1\}$, and let C be the initial segment of the $\leq$-order of size $|A|$. Define the particular points $w = 0((n-1) \cdot (q-1))$ and $w_i = i((n-1) \cdot (q-1))$ for $1 \leq i \leq q-1$. Thus Claim 10 implies that we have $d([s]^n) \setminus \{w\} \subseteq d(A)$.

We start by verifying that $C \cap C_T \subseteq A \cap C_T$. If $q > s$, the inclusion is trivial as we have $C \cap C_T = \emptyset$. If $q = s$, it follows that $A \setminus (B_{\geq 1} \cup ([s]^n \cap B_0)) = A \cap C_T$, and similarly we have $C \setminus (B_{\geq 1} \cup ([s]^n \cap B_0)) = C \cap C_T$ (as $T = \{1\}$ is the first non-empty set under the binary order). Since C is an initial segment of size $|A|$, we have $|A \cap C_T| \geq |C \cap C_T|$. As $A \cap C_T$ is of the form $\{y \in C_T : y \leq x\}$ for some $x \in C_T$, the inclusion follows.

First consider the case when we also have $w \in d(A)$. Then Claim 10 implies that we have $d([s]^n) \subseteq d(A)$, and since we also have $d(C \cap C_T) \subseteq d(A)$, it follows that $d(C) \subseteq d(A)$. Thus we may take $B = C$.

Now suppose that we have $w \notin d(A)$. We first deal with the easy case $q > s$. Note that we have $d(C) \subseteq d(B_{\geq 1} \cup ([s+1]^n \cap B_0))$. Since $q \geq s+1$, Claim 10 implies that $d([s+1]^n \cap B_0) \setminus \{w\} \subseteq d(A)$, and hence it follows that $d(C) \setminus d(A) \subseteq \{w\}$. Since $q \geq s+1$, A contains a point $u \in ([q]^n \setminus [s]^n) \cap B_0$. As $n \geq 2$ and $R_0(u) = \emptyset$, $d(u)$ contains a point having a coordinate which is at least $s+1$. In particular, it follows that $d(u) \not\subseteq d(C)$, and hence we have $|d(A) \setminus d(C)| \geq 1$. Combining this with the earlier observation $|d(C) \setminus d(A)| \leq 1$, it follows that $|d(A)| \geq |d(C)|$, and hence we may take $B = C$.

Now suppose that we have $w \notin d(A)$ and $q = s$. By using the previous observation $C \cap C_T \subseteq A \cap C_T$ and Claim 10, it follows that $d(C) \setminus d(A) = \{w\}$. Since $w \notin d(A)$ and $q = s$, it follows that $\{w_1, \ldots, w_{s-1}\} \subseteq C \setminus A$, and in particular we have $|(C \setminus A) \cap (B_{\geq 1} \cup ([s]^n \cap B_0))| \geq s-1$. Since $A \setminus (B_{\geq 1} \cup ([s]^n \cap B_0)) = A \cap C_T$ and $C \setminus (B_{\geq 1} \cup ([s]^n \cap B_0)) = C \cap C_T$, it follows that $|A \cap C_T| \geq |C \cap C_T| + s - 1$.

Let x be the largest point in A and y be the largest point in C under the $\leq$-order. Define $X = \{z : z \leq x\}$, and note that $C = \{z : z \leq y\}$ as C is an initial segment. Note that we have $x, y \in C_T$, and hence we have $|X| \geq |C| + (s-1)$. Thus Claim 12 with $L = s$ implies that we have $|d(X)| > |d(C)|$, unless $x = (2)((n-1) \cdot (s))$ and $y = (n \cdot s)$. However, these points are not in C_T since $T = \{1\}$, and thus there exists an element $u \in d(X) \setminus d(C)$. Since $u \in d(\{z : x < z \leq y\}) \subseteq d(C_T \cap A) \subseteq d(A)$, it follows that $|d(A) \setminus d(C)| \geq 1$. Combining this with the earlier observation $|d(C) \setminus d(A)| = 1$, it follows that $|d(A)| \geq |d(C)|$.

Finally, consider the case when we have $T \neq \{1\}$. If $q > s$, Claim 10 implies that we have $d([s+1]^n \cap B_0) \subseteq d(A)$, and thus we can take $B = C$. If $q = s$, let $X = A \setminus (B_{\geq 1} \cup ([s]^n \cap B_0))$ and $k = |A| - |B_{\geq 1} \cup ([s]^n \cap B_0)|$. Let Y be the set of k least points in X under the $\leq$-order, and let $B = B_{\geq 1} \cup ([s]^n \cap B_0) \cup Y$. We certainly have $|A| = |B|$, and since $q = s$ it follows that $B \subseteq B_{\geq 1} \cup ([s+1]^n \cap B_0)$. Furthermore, as A is compressed, it follows that B is compressed as well. Since $T \neq \{1\}$, Claim 10 implies that we have $d([s]^n \cap B_0) \subseteq d(A)$. Since $Y \subseteq X$, it follows that $d(B) \subseteq d(A)$, and thus B satisfies the required conditions. $\square$

From now on, we assume that A is a compressed set for which there exists $s \geq 2$ so that $B_{\geq 1} \cup ([s]^n \cap B_0) \subseteq A \subseteq B_{\geq 1} \cup ([s+1]^n \cap B_0)$. As before, we set T to be the largest set under the binary order for which we have $C_T \cap A \neq \emptyset$. Our next aim is to prove an appropriate version of Claim 8.

Claim 14.

1. If $|T| \neq 1$, then for all S with $S <_{bin} T$ we have $C_S \subseteq A$.
2. Let $T = \{j\}$ and $U = \{1, \ldots, j-1\}$ where $j \neq 1$. Then for all S with $S <_{bin} U$ we have $d(C_S) \subseteq d(A)$. Define the particular point $w_j = 0_j\left(((j-1) \cdot s)\left((n-j) \cdot (s-1)\right)\right)$. Then we also have $d(C_U) \setminus \{w_j\} \subseteq d(A)$.

Proof of Claim 14. We start by proving the first part. Let $j = \max(T)$ and set $U = \{1, \ldots, j-1\}$. Since $|T| \neq 1$ it follows that $U \cap T \neq \emptyset$, and hence (4.10) implies that we have $C_U \subseteq A$. Given any $S \in \mathcal{A}$ with $S <_{bin} T$, note that we must either have $S \subseteq U$ or $j \in S$. If $S \subseteq U$, we also have $S \leq_{bin} U$ and since $C_U \subseteq A$, (4.10) implies that we also have $C_S \subseteq A$. If $j \in S$, we certainly have $S \cap T \neq \emptyset$. Since $S <_{bin} T$, (4.10) implies that $C_S \subseteq A$. In particular, for all $S <_{bin} T$ we have $C_S \subseteq A$, which completes the proof of the first part.

Now suppose that we have $T = \{j\}$ for some j with $j \notin \{1, n\}$. The assumption $j < n$ is used in the proof when we are considering the $(j+1)^{th}$ coordinate of a point. Again, let $U = \{1, \ldots, j-1\}$. Define the particular points $a = s_j\left((n-1) \cdot 1\right)$ and $b = \left((j-1) \cdot s\right)\left((n-j+1) \cdot 1\right)$. Note that a is the least point under the $\leq$-order in C_T and b is the least point in C_U. Since $C_T \cap A \neq \emptyset$, (4.10) implies that we have $a \in A$. Since $U <_{bin} T$ and $a_{j+1} = b_{j+1} = 1$, (4.10) implies that we have $b \in A$, and thus we also have $C_U \cap A \neq \emptyset$. Let S be a set satisfying the conditions $S <_{bin} T$ and $S \neq U$. Since $S \subset U$ and $C_U \cap A \neq \emptyset$, (4.10) implies that we have $C_S \subseteq A$.

Let $x \in d(C_U) \setminus \{w_j\}$ and let i be the unique index satisfying $x_i = 0$. If $i \neq j$, consider the point y obtained by taking $y_t = x_t$ for $t \neq i$ and $y_i = 1$. Since $j \neq i$, it follows that $y_i = a_i = 1$. We also have $R_s(y) \subseteq U$, and hence it follows that $y \leq a$. Since $a \in A$, (4.10) implies that we have $y \in A$, and hence we also have $x \in d(A)$.

Now suppose that $i = j$. Since $x \in d(C_U) \setminus \{w_j\}$ and $i = j$, it follows that $x_t = s$ for all $t \leq j-1$, $x_j = 0$, $1 \leq x_t \leq s-1$ for all $t \geq j+1$, and there exists $k \geq j+1$ satisfying $x_k \leq s-2$. Let u be the point obtained by setting $u_t = x_t$ for $t \neq j$ and $u_j = 1$, and let v be obtained by setting $v_t = x_t$ for all $t \notin \{j, k\}$, $v_j = s-1$ and $v_k = 1$. It is easy to see that we have $x \in d(u)$. Since $R_s(u) = R_s(v)$ and $R_{s-1}(v) = R_{s-1}(u) \cup \{j\}$, it follows that $u < v$. Since $k \neq j$, it follows that $v_k = a_k = 1$, and hence (4.10) implies that we have $v \in A$. Since $u_1 = v_1 = s$, (4.10) implies that $u \in A$ and hence we have $x \in d(A)$.

Finally, suppose that $T = \{n\}$. Let $U = \{1, \ldots, n-1\}$ and $V = \{2, \ldots, n-1\}$. Note that for any $S <_{bin} T$ we either have $S = U$ or $S \leq_{bin} V$. Define the particular point a by setting $a = ((n-1) \cdot 1)\, s$. Since a is the least point in C_T, it follows that $a \in A$.

Given a set S satisfying the condition $S <_{bin} T$ and a point $x \in d(C_S)$, and let i be the unique index satisfying $x_i = 0$. Note that we may certainly assume that $x \notin d(C_{S'})$ for any proper subset S' of S, with the convention $C_\emptyset = [s]^n \cap B_0$. It is easy to see that this is equivalent to assuming that $i \notin S$.

If $S = U$, the only possible point is $x = ((n-1) \cdot s)\, 0$ which is precisely the point w_n. Hence we are done in this case. Now suppose that $S = V$, and hence we have $i \in \{1, n\}$. Let u and v be the points $u = (1)\,((n-2) \cdot s)\,(s-1)$ and $v = (s-1)\,((n-2) \cdot s)\,(1)$. Since $a_1 = u_1 = 1$ and $u < a$, (4.10) implies that $u \in A$. Note that we have $R_s(u) = R_s(v)$ and $R_{s-1}(v) <_{bin} R_{s-1}(u)$, and hence it follows that $v < u$. Since $n \geq 3$, we have $u_2 = v_2 = s$, and hence (4.10) implies that we have $v \in A$.

For all $1 \leq j \leq s-2$ and $1 \leq k \leq s-2$ define the points $w_{j,k} = (j)\,((n-2) \cdot s)\,(k)$. Since $R_{s-1}(w_{j,k}) = \emptyset$, it follows that $w_{j,k} < u$. As $(w_{j,k})_2 = u_2 = s$, (4.10) implies that we have $w_{j,k} \in A$ for all $1 \leq j, k \leq s-2$.

Since $0 \in \{x_1, x_n\}$, we either have $x = (0)\,((n-2) \cdot s)\,(c)$ or $x = (d)\,((n-2) \cdot s)\,(0)$ for some $1 \leq c, d \leq s-1$. If $c = s-1$ or $d = s-1$, we have $x \in d(u)$ or $x \in d(v)$ respectively. If $c \leq s-2$ or $d \leq s-2$, it is easy to see that we have $x \in d(w_{1,c})$ or $x \in d(w_{d,1})$ respectively. In particular, it follows that $x \in d(A)$. This completes the proof when $S = V$.

In order to deal with the sets S with $S <_{bin} V$, we first consider the case when we have $n \geq 4$. As noticed in the previous case, we have $C_V \cap A \neq \emptyset$. If $S \cap V \neq \emptyset$, (4.10) implies that we also have $C_S \subseteq A$ and hence the conclusion certainly follows. The only non-empty set S satisfying $S <_{bin} V$ together with $S \cap V = \emptyset$ is $S = \{1\}$. However, as $n \geq 4$, it follows that $\{1\} <_{bin} \{1,2\} <_{bin} V$, and since $\{1\} \cap \{1,2\} \neq \emptyset$ and $\{1,2\} \cap V \neq \emptyset$, we also have $C_{\{1\}} \subseteq D$. This completes the proof when $n \geq 4$.

Now suppose that $n = 3$. Then $V = \{2\}$, and thus the only non-empty set S satisfying $S <_{bin} V$ is $S = \{1\}$. Let $x \in d\left(C_{\{1\}}\right)$ and let i be the unique index satisfying $x_i = 0$. Recall that we may assume that $i \in \{2, 3\}$. Let y be the point obtained by setting $y_j = x_j$ for $j \neq i$ and $y_i = 1$. Our aim is to prove that $y \in D$.

If $i = 2$ and $y_3 = s-1$, we have $y_3 = u_3$. Since $y \leq u$, (4.10) implies that we have $y \in D$. If $i = 2$ and $y_3 = c$ for some $1 \leq c \leq s-2$, then $y_3 = c = (w_{1,c})_3$, and since $y \leq w_{1,c}$ it follows that $y \in D$. Finally, if $i = 3$, we have $y_3 = v_3 = 1$, and since $y < v$ it follows that $y \in D$. Thus in all three cases we have $y \in D$, which completes the proof. $\square$

Now we have the necessary tools to finish the proof of Theorem 34. Let A be a set satisfying $B_{\geq 1} \cup ([s]^n \cap B_0) \subseteq A \subseteq B_{\geq 1} \cup ([s+1]^n \cap B_0)$, and let C be the initial segment of the $\leq$-order of size $|A|$. Since A and C have the same size, we must have $B_{\geq 1} \cup ([s]^n \cap B_0) \subseteq C \subseteq B_{\geq 1} \cup ([s+1]^n \cap B_0)$. Let T_1 be the largest set under the binary order satisfying $C_{T_1} \cap C \neq \emptyset$, and since C is an initial segment we certainly have $T_1 \leq_{bin} T$.

First suppose that we have $T_1 <_{bin} T$, and set

$$X = B_{\geq 1} \cup ([s]^n \cap B_0) \cup \bigcup_{S \leq_{bin} T_1} C_S.$$

Note that we cannot have $T = \{1\}$ as $\{1\}$ is the least set under the binary order in $\mathcal{A}$.

If $|T| \geq 2$, for any $S \in \mathcal{A}$ with $S \leq_{bin} T_1$ we have $S <_{bin} T$, and hence Claim 14 implies that we have $C_S \subseteq A$. Hence it follows that $X \subseteq A$, and since $C \subseteq X$ we must have $C \subseteq A$. Since A and C are sets of the same size, it follows that $A = C$. However, this contradicts the assumption $T_1 <_{bin} T$, so this case cannot actually occur when $|T| \geq 2$.

If $T = \{j\}$ for some $j \neq 1$, define the particular point $w_j = 0_j\left((j-1)\cdot s\right)\left((n-j)\cdot(s-1)\right)$. Claim 14 implies that for all $S <_{bin} \{1,\ldots,j-1\}$ we have $d(C_S) \subseteq d(A)$, and for $U = \{1,\ldots,j-1\}$ we have $d(C_U) \setminus \{w_j\} \subseteq d(A)$. Since $T_1 <_{bin} T$, it follows that $T_1 \leq_{bin} U$, and hence we have $d(X) \setminus \{w_j\} \subseteq d(A)$. In particular, it follows that $|d(X) \setminus d(A)| \leq 1$.

Pick any point $v \in C_T$, and consider an element $u \in d(v)$ obtained by flipping the first coordinate to 0. Since $j \neq 1$, it follows that $u_j = s$. In particular, we must have $u \notin d(X)$ as $T_1 \leq_{bin} U$. Hence it follows that $u \in d(A) \setminus d(X)$, and hence we have $|d(A) \setminus d(X)| \geq 1$. In particular, it follows that $|d(A)| \geq |d(X)|$, and since $C \subseteq X$ this completes the proof of this case.

Hence we are only left with the case when $T_1 = T$. It turns out to be convenient to split the rest of the argument into two subcases based on the size of T. When $T = \{1\}$, there are no non-empty sets S satisfying $S <_{bin} T$. Hence we can deal with the cases $T = \{1\}$ and $|T| \geq 2$ at the same time.

Case 2.1. $|T| \geq 2$ or $T = \{1\}$.

Let C be an initial segment of size $|A|$, and let s, T and T_1 be chosen as before. Recall that we have $T_1 = T$. We prove that under these assumptions we must have $A = C$.

Let X be defined by $X = B_{\geq 1} \cup ([s]^n \cap B_0) \cup \left(\bigcup_{S <_{bin} T} C_S\right)$. Claim 14 implies that for all $S <_{bin} T$ we have $C_S \subseteq A$, and hence it follows that $X \subseteq A$. On the other hand, by the choice of T we have $A \setminus X = A \cap C_T$ and $C \setminus X = C \cap C_T$. Since $|A| = |C|$ and both of these sets contain X as a subset, it follows that $|A \cap C_T| = |C \cap C_T|$. Since both sets $A \cap C_T$ and $C \cap C_T$ are of the form $\{y \in C_T : y \leq x\}$ for some element $x \in C_T$, we must have $A \cap C_T = C \cap C_T$. Hence it follows that $A = C$.

Case 2.2. $T = \{j\}$ where $2 \leq j \leq n$.

As before, let C be the initial segment of size $|A|$, let s, T and T_1 be chosen as before, and recall that we have $T_1 = T = \{j\}$ for some $2 \leq j \leq n$. Define the particular points $w = 0_j\left((j-1)\cdot s\right)\left((n-j)\cdot(s-1)\right)$ and $w_i = i_j\left((j-1)\cdot s\right)\left((n-j)\cdot(s-1)\right)$ for $1 \leq i \leq s-1$. Note that if we have $w \in d(x)$ for some element $x \in C_S$ with $S \leq_{bin} T$, it follows that $x = w_i$ for some i. As before, let $X = B_{\geq 1} \cup ([s]^n \cap B_0) \cup \bigcup_{S <_{bin} T} C_S$.

First consider the easy case when we have $w \in d(A)$. Then Claim 10 implies that we have $d(C_S) \subseteq d(A)$ for all S satisfying $S <_{bin} T$, and hence we have $d(X) \subseteq d(A)$. Since $X \subseteq C$ and we have $C \setminus X \subseteq C_T$ and $A \setminus X \subseteq C_T$, it follows that $|A \cap C_T| \geq |C \cap C_T|$. By using the same argument as in the proof of Case 2.1, it follows that $C \cap C_T \subseteq A \cap C_T$, and thus it follows that $d(C) \subseteq d(A)$.

Now suppose that $w \notin d(A)$, and hence we must have $\{w_1,\ldots,w_{s-1}\} \cap A \neq \emptyset$. Since $\{w_1,\ldots,w_{s-1}\} \subseteq X \subseteq C$, it follows that $|C \setminus X| \geq |A \setminus X| + s - 1$, and thus it follows that $|A \cap C_T| \geq |C \cap C_T| + s - 1$. Let u and v be chosen so that $A \cap C_T = \{z \in C_T : z \leq u\}$ and $C \cap C_T = \{z \in C_T : z \leq v\}$. Let W be the initial segment obtained by setting $W = \{z : z \leq u\}$, and note that we have $C = \{z : z \leq v\}$. Since $|C \setminus X| \geq |A \setminus X| + s - 1$, it follows that $|W| \geq |C| + s - 1$.

Since $n \geq 3$, it follows that $(1)((n-1) \cdot s) \notin C_T$, and hence must have $v \neq (1)((n-1) \cdot s)$. Since $|W| \geq |C| + s - 1$, Claim 12 with $L = s + 1$ implies that there exists an element $a \in d(W) \setminus d(C)$. Hence we have $a \in d(A \cap C_T)$, and thus it follows that $|d(A) \setminus d(C)| \geq 1$. Together with the previous observation $|d(C) \setminus d(A)| = 1$, it follows that $|d(A)| \geq |d(C)|$.

This completes the proof of Theorem 34. $\square$

4.3 Minimal d-shadow for a given rank

Recall that we defined $[k]^n_r = \{x \in [k]^n : w(x) = r\}$ to be the set of points with exactly r non-zero coordinates, and consider the restriction of the $\leq$-order on $[k]^n_r$. Since $|R_0(x)| = n - r$ holds for every point $x \in [k]^n_r$, it follows that for distinct x and y we have $x \leq y$ if and only if $\max(R_j(x) \Delta R_j(y)) \in R_j(y)$, where j is the largest index satisfying $R_j(x) \neq R_j(y)$.

Note that for all $m \leq k$, $[m]^n_r$ is an initial segment of the $\leq$-order on $[k]^n_r$, and the restrictions of the $\leq$-order on $[m]^n_r$ and on $[k]^n_r$ coincide on $[m]^n_r$. Thus the $\leq$-order naturally extends to an order on $\mathbb{N}^n_r = \{x = x_1 \dots x_n : x_i \in \mathbb{N}, w(x) = r\}$, which we will also denote by $\leq$.

The notion of the d-shadow is still sensible for the subsets of $\mathbb{N}^n_r$ as well. If $A \subseteq \mathbb{N}^n_r$, we certainly have $d(A) \subseteq \mathbb{N}^n_{r-1}$. We now use Theorem 34 to deduce that among the subsets of $\mathbb{N}^n_r$ of a given finite size, the initial segment of the $\leq$-order minimises the size of the d-shadow.

Theorem 35. *Let A be a finite subset of $\mathbb{N}^n_r$ and let C be the initial segment of the $\leq$-order on $\mathbb{N}^n_r$ of size $|A|$. Then we have $|d(A)| \geq |d(C)|$.*

Proof. Let k be the size of A. Since A contains points of length n and with exactly r non-zero coordinates, by relabeling the elements of the ground set if necessary we may assume that $A \subseteq [nk]^n_r$. Let $B = [nk]^n_{\leq r-1} \cup A$, where $[nk]^n_{\leq r-1} = \{x \in [nk]^n : |R_0(x)| \geq n - (r-1)\} = B_{\geq n-r+1}$. Let X be the initial segment of size $|B|$ on $[nk]^n$. Then X can be written as $X = [nk]^n_{\leq r-1} \cup C$, where C is the initial segment of the $\leq$-order of size $|A|$ on $[nk]^n_r$. Thus C is also the initial segment of the $\leq$-order of size $|A|$ on $\mathbb{N}^n_r$.

Note that we have

$$|d(B)| = \left|[nk]^n_{\leq r-1}\right| + |d(A)|$$

and

$$|d(X)| = \left|[nk]^n_{\leq r-1}\right| + |d(C)|,$$

as

$$d(A) \cap [nk]^n_{\leq r-1} = d(C) \cap [nk]^n_{\leq r-1} = \emptyset.$$

Since $|B| = |X|$ and X is an initial segment of the $\leq$-order on $[nk]^n$, Theorem 34 implies that we have $|d(B)| \geq |d(X)|$. Hence it follows that $|d(A)| \geq |d(C)|$, which completes the proof. $\square$

Chapter 5

The Toucher-Isolator game

5.1 Introduction

Consider the following game, called the 'Toucher-Isolator' game, introduced by Dowden, Kang, Mikalački and Stojaković [17]. The two players, Toucher and Isolator, claim edges of a given graph G alternately with Toucher having the first move. At the end of the game when all the edges of G have been claimed, a vertex is said to be *touched* if it is incident with at least one of the edges claimed by Toucher, and *isolated* otherwise. In particular, all the edges incident with an isolated vertex must be claimed by Isolator. The aim of Toucher is to minimise the number of isolated vertices and the aim of Isolator is to maximise the number of isolated vertices. Hence Toucher-Isolator game is a 'quantitative' Maker-Breaker type of game. Define the value of the game $u(G)$ to be the number of isolated vertices at the end of the game when both players play under optimal strategies.

In [17], the authors studied the size of $u(G)$ for graphs G belonging to certain families. For a tree T, they gave bounds for $u(T)$ in terms of the degree sequence of T. As a consequence, they proved that if T is a tree with n vertices, then

$$\frac{n+2}{8} \leq u(T) \leq \frac{n-1}{2}. \tag{5.1}$$

If T is a star with $n \geq 3$ vertices, it is easy to verify that $u(T) = \left\lfloor \frac{n-1}{2} \right\rfloor$ regardless of how Toucher and Isolator play the edges. Hence the upper bound in (5.1) is tight.

Let P_n be the path on n vertices and C_n be the cycle of length n. The authors proved in [17] that when $T = P_n$, the general bound (5.1) can be improved to

$$\frac{3}{16}(n-2) \leq u(P_n) \leq \frac{n+1}{4},$$

and as an easy consequence it follows that

$$\frac{3}{16}(n-3) \leq u(C_n) \leq \frac{n}{4}.$$

These bounds imply that the asymptotic proportion of isolated vertices is between $3/16$ and $1/4$ in both cases. The authors asked in [17] what the correct asymptotic proportion of isolated vertices is and suggested that the correct answer could be $1/5$ for cycles and paths. In this chapter we prove that this is the correct asymptotic proportion, and in fact, we find the exact values of $u(C_n)$ and $u(P_n)$ for all n.

Theorem 36. *Let $n \geq 3$. When both players play optimally, there will be $\lfloor \frac{n+3}{5} \rfloor$ isolated vertices on P_n and $\lfloor \frac{n+1}{5} \rfloor$ isolated vertices on C_n.*

Our other main result is to improve their lower bound of (5.1) for general trees. We prove that paths are the 'worst' for Isolator among the trees with n vertices.

Theorem 37. *Let $n \geq 3$ and let T be a tree with n vertices. When both players play optimally, there will be at least $\lfloor \frac{n+3}{5} \rfloor$ isolated vertices on T.*

Although this chapter is self-contained, for a general background on Maker-Breaker type of games see Beck [4]. There are many other papers dealing with achievement games on graphs – see for example a classical paper of Chvátal and Erdős [12], and subsequent papers [3, 21].

We start by outlining the proof of Theorem 36. For convenience, in the proof of Theorem 36 we work with the 'dual version' of the game which is played on the vertices of a path or a cycle. In the dual version, Isolator is aiming to claim as many pairs of consecutive vertices as possible, and Toucher is aiming to minimise the number of such pairs claimed by Isolator. In addition, in the dual version played on a path, claiming an endpoint of the path increases the score by one as well.

Since Isolator is trying to maximise the score, it seems sensible for her to start by claiming some suitably chosen vertex i, and then trying to claim as long a block of consecutive elements as possible. Now suppose she has claimed a block of length t and she cannot proceed in this way. This means that Toucher has claimed the points next to the endpoints of this block, or one of the endpoints of the block is also an endpoint of the path. Removing this block together with exactly one of its endpoints claimed by Toucher leaves a shorter path with no elements claimed by Isolator and $t-1$ elements claimed by Toucher. This motivates us to define a 'delayed' notion of the dual Toucher-Isolator game played on a path, where at the start of the game Toucher is allowed to claim k points, and then the players claim elements alternately.

It turns out that following the idea described above, we can prove a good enough lower bound for the delayed version of the Toucher-Isolator game. However, one has to be slightly careful with the choice of the initial move whenever a new block is started. This is especially important near the start of the game when the endpoints of the path are not yet claimed. Using such an approach, we can prove the lower bound of Theorem 36.

By analysing the proof of the lower bound, one can observe that allowing Isolator to claim multiple 'long blocks' would allow Isolator to achieve a better score than the one stated in Theorem 36. This suggests that Toucher should primarily claim an element next to the element claimed by Isolator on her previous move. As a result of such a pair of moves, the initial path splits into two disjoint paths. However, the endpoints of the paths obtained during the process may behave in different ways, as the vertex next to an endpoint before splitting the board may have been claimed by Toucher or Isolator. Hence it turns out to be useful to define three different games played on paths which essentially only differ at the endpoints of the path. Given that after every pair of moves the path splits into two disjoint paths, it is natural to analyse all three types of boards simultaneously by considering games G played on boards that are disjoint unions of these three types of paths. By proving an upper bound for such generalised boards, we prove the upper bound of Theorem 36.

We now outline the proof of Theorem 37. The approach is very similar to the proof of the lower bound for $u(P_n)$ in Theorem 36. At the early stages of the game, it seems natural for Isolator to claim edges having a leaf as an endpoint, as claiming such an edge instantly increases the score by one. When $T = P_n$, this naturally corresponds to claiming blocks of consecutive elements near the endpoints at the start of the game. Note that after claiming such an edge, the isolated leaf does not affect the rest of the game, and hence it may be discarded (together with the edge that was occupied by Isolator). During the process, the other endpoint of the edge may become a leaf, and in such case Isolator can continue claiming isolated leaves.

Suppose that at some point Isolator has no such move available. Let T_1 be the tree obtained as a result of the process and let C_1 be the set of edges claimed by Toucher. Hence for every leaf $v \in T_1$ there exists an edge $e \in C_1$ whose endpoint v is. Such a situation is very similar to the delayed game introduced in the proof of Theorem 36, in which Toucher is allowed to claim a certain number of edges at the start of the game and in which isolating endpoints does not increase the score.

However, before we can define a delayed game that is good enough for our purposes, the structure of T_1 and C_1 may need to be modified. Let v be a leaf in T_1, and let w be the unique neighbour of v. Similarly to leaves that are already isolated, the leaves that are already touched are useless for the rest of the game, so it would be tempting to just delete them. However, during the process we must keep in mind that w is already touched, even though the edge vw is deleted during the process. Thus it will be convenient to declare a set $X_1 \subseteq V(T_1)$ of 'additional' touched vertices for the delayed game.

It turns out that such a simple deletion is fine whenever $d_{T_1}(w) = 2$. Indeed, in this case w becomes a leaf as its neighbour v is removed from the graph. When $d_{T_1}(w) \geq 3$, we need to modify the structure of T_1 in a slightly different way. In this case, we duplicate w into two vertices joined by an edge claimed by Toucher and split the original neighbours of w between these two vertices so that both of these new vertices have at least one neighbour. As a result, we may restrict ourselves to those delayed games where X_1 is exactly the set of all leaves in T_1.

When analysing the delayed game, we start by modifying the structure of the tree locally even further. If T contains touched edges that are close to each other or a touched edge that is close to a leaf, it seems intuitive that Toucher can prevent Isolator from increasing the score near such local configuration. Hence it would be convenient to delete such local configurations. After the deletion of such configurations, we move on to analysing structures that are suitable for Isolator. The structures we are looking for are local configurations that are similar to paths, that is, neighbouring vertices of degree 1 or 2 that are not yet touched. If there are no such configurations in T, then the vertices of degree 1 or 2 must be spread out or they are already touched. Since every leaf of T is declared to be touched at the start of the delayed game, in both cases T contains many vertices that are initially touched, which implies the result by induction.

The plan of this chapter is following. In Section 5.2.1 we prove the lower bound for $u(P_n)$. Of course, this lower bound could also be deduced from Theorem 37, but in this special case a much simpler proof is available. Furthermore, some important ideas that are used in the proof of Theorem 37 are already introduced in the proof of the lower bound. In Section 5.2.2 we prove the upper bound for $u(P_n)$, and deduce Theorem 36. We start Section 5.3. by defining a suitable notion that allows us to reduce the tree without changing the game too much, and in Section 5.3.1 we deal with the first phase of the game in which Isolator is claiming only leaves. In Section 5.3.2 we move on to analysing the specific delayed version of the game where leaves are initially counted as touched vertices, and Toucher is allowed to claim a certain number of edges at the start of the game. We then use such delayed games to deal with the second phase of the game, and we deduce Theorem 37 from these results.

5.2 Toucher-Isolator game on a cycle and a path

Recall that in this section we work with the dual version of the Toucher-Isolator game. There is a natural connection between the original and the dual version when the board is a cycle, but we need to be more careful with paths due to irregular behaviour near the endpoints. Define the game $F(n)$ played on the elements of $\{1, \dots, n\}$ with two players Isolator and Toucher claiming elements in alternating turns with the first move given to Isolator. Define the *score* of this game to be the number of pairs $\{i, i+1\}$ for which both i and $i+1$ are claimed by Isolator, and as usual Isolator is aiming to maximise the score and Toucher is aiming to minimise the score. Let $\alpha(n)$ be the score attained at the end of the game when both Isolator and Toucher play optimally.

Consider the dual version of the Toucher-Isolator game played on a cycle with vertex set $\{1, \dots, n\}$, and recall that the first move is given to Toucher. Due to the symmetry of the cycle, we may assume that Toucher claims the element n on her first move. Hence after the first move, the available pairs that can increase the score are the pairs $\{1, 2\}, \dots, \{n-2, n-1\}$. These correspond to the pairs that can increase the score of the game $F(n-1)$, and since Isolator has the next move, it follows that the subsequent game is equivalent to the game $F(n-1)$. In particular, we have $u(C_n) = \alpha(n-1)$.

The situation with the dual version on a path is slightly more complicated due to irregular behaviour at the endpoints. Define the games $G(n)$ and $H(n)$ both played on $\{1, \dots, n\}$, with the players Isolator and Toucher claiming elements in alternating turns with the first move given to Isolator. On $G(n)$, we increase the score by one for each pair $\{i, i+1\}$ with both i and $i+1$ claimed by Isolator, and the score is also increased by 1 if Isolator claims the element 1. In a sense, this can be viewed as a game on the board $\{0, \dots, n\}$ with 0 assigned to Isolator initially. Similarly on $H(n)$, we increase the score by one for each pair $\{i, i+1\}$ with both i and $i+1$ claimed by Isolator, and additionally the score is increased by 1 for each element in the set $\{1, n\}$ claimed by Isolator. Again, this can be viewed as a game on the board $\{0, \dots, n+1\}$ with both 0 and $n+1$ assigned to Isolator initially. Define $\beta(n)$ and $\gamma(n)$ to be the scores of these games when both players play optimally.

Let B denote a Toucher-Isolator game played on a board that is of the form $F(n)$, $G(n)$ or $H(n)$ for some n. If Toucher plays her first move adjacent to Isolator's first move, the board B splits into at most two disjoint boards, both of which are of the form $F(m)$, $G(m)$ or $H(m)$ for some m. However, note that these two boards are not in general of the same form or size, and not necessarily of the same form as the original board. Hence in order to find an upper bound for the value of the game, it turns out to be useful to consider games G played on a board that is a disjoint union of $F(l_1), \ldots, F(l_r)$, $G(m_1), \ldots, G(m_s)$ and $H(n_1), \ldots, H(n_t)$.

Finally, define the game $H_b(n)$ in the same way as $H(n)$, but with the first move given to Toucher, and let $\gamma_b(n)$ be the score of this game when both players play optimally. It is easy to see that the Toucher-Isolator game and $H_b(n-1)$ are equivalent, as the $n-1$ edges naturally correspond to the $n-1$ elements of $H_b(n-1)$. Hence it follows that $u(P_n) = \gamma_b(n-1)$, and thus it suffices to find the value of γ_b.

5.2.1 The lower bound

We start by focusing on the game $F(n)$ and finding the value of $\alpha(n)$. Since Isolator is trying to maximise the score, it seems sensible for her to start by claiming some suitably chosen i, and then trying to claim as long a block of consecutive elements as possible. As long as $i \notin \{1, n\}$, she can certainly guarantee a block of length at least 2. Now suppose that she has claimed a block of length t, and that she cannot proceed in this way. This means that Toucher must have claimed the points next to the endpoints of the block (or one of the endpoints is 1 or n). Removing this block together with exactly one of its endpoints claimed by Toucher leaves a path with $n-t-1$ elements containing exactly $t-1$ elements claimed by Toucher, and no elements claimed by Isolator.

This motivates the definition of the following game, which can be viewed as a delayed version of $F(n)$. Let $F(n,k)$ be the game played on $\{1, \ldots, n\}$, where at the start of the game Toucher is allowed to claim k points, and then the players claim elements alternately with the first move given to Isolator. The score of the delayed game is defined in the same way as the score of $F(n)$, and hence $F(n)$ and $F(n,0)$ are identical games. Let $\alpha(n,k)$ be the score attained when both players play optimally. We start by proving the following lower bound for $\alpha(n,k)$, which is later used to deduce the lower bound for $\gamma_b(n)$.

Lemma 38. *For all $n \geq 2$ and $k \leq n$ we have $\alpha(n,k) \geq \left\lfloor \frac{n-3k+2}{5} \right\rfloor$.*

Proof. Suppose that Toucher claims the elements $s_1, \ldots, s_k$ on her first move. These elements split the path into $k+1$ (possibly empty) intervals of lengths $l_0, \ldots, l_k$, with $l_i = s_{i+1} - s_i - 1$ (where $s_0 = 0$ and $s_{k+1} = n+1$). By symmetry we may assume that l_0 is the longest interval.

If $l_0 \leq 2$, then $n \leq k + 2 \cdot (k+1) = 3k+2$, and hence it follows that $\left\lfloor \frac{n-3k+2}{5} \right\rfloor \leq 0$. Thus the claim evidently holds in this case, and hence we may assume that $l_0 \geq 3$. We treat the cases $l_0 \geq 4$ and $l_0 = 3$ individually. In both cases the proof follows the same idea, however the choice of the initial move is slightly different for $l_0 = 3$ since an interval with only 3 elements is 'too short' for the general argument.

Case 1. $l_0 \geq 4$.

The aim for Isolator is to build a long block of consecutive elements inside the interval. Initially, she claims the element 3. Assuming she has already claimed exactly the elements in $\{t, \ldots, t+r\}$ for some t and r, she claims one of the elements $t+r+1$ or $t-1$, if possible. If not, she stops.

Consider the point when the process terminates and suppose that at the point when the process terminates, the elements claimed by Isolator are exactly the elements in $\{t, \ldots, t+r\}$.

Since this set contains the element 3, we must have $t+r \geq 3$ and $t \in \{1,2,3\}$. Furthermore, the element $t+r+1$ must be claimed by Toucher, and either $t=1$ or the element $t-1$ is claimed by Toucher. Since $l_0 \geq 4$, it follows that the elements 2 and 4 are not claimed after Isolator's first move. Since Toucher cannot claim both of these elements on her first move, it follows that Isolator can always guarantee that $r \geq 1$.

Let $T_1 = \{t+r+2, \ldots n\}$ and let b be the number of elements claimed by Toucher in T_1. Note that Toucher has claimed $k+r+1$ elements in total and one of them must be $t+r+1$. Furthermore, if $t > 1$ then one of them must be $t-1$ as well. Hence it follows that $b \leq k+r$, and if $t \geq 2$ we also have $b \leq k+r-1$. Also note that Isolator has not claimed any elements in T_1.

Note that Isolator has increased her score by r by claiming the elements $\{t, \ldots, t+r\}$, and this is the only contribution arising outside T_1. Thus the total score that Isolator can attain is at least $r + \alpha(n-t-r-1, b)$. By induction, it follows that the score is at least

$$r + \left\lfloor \frac{n-t-r-1-3b+2}{5} \right\rfloor. \tag{5.2}$$

If $t=1$, it follows that $b \leq k+r$. Furthermore, the condition $t+r \geq 3$ implies that $r \geq 2$. Hence (5.2) implies that Isolator can guarantee that the score is at least

$$\left\lfloor \frac{n-3k+r+1-t}{5} \right\rfloor \geq \left\lfloor \frac{n-3k+2}{5} \right\rfloor.$$

If $t \geq 2$, it follows that $b \leq k+r-1$. Recall that we always have $t \leq 3$ and $r \geq 1$. Hence (5.2) implies that Isolator can guarantee that the score is at least

$$\left\lfloor \frac{n-3k-t+r+4}{5} \right\rfloor \geq \left\lfloor \frac{n-3k-3+1+4}{5} \right\rfloor = \left\lfloor \frac{n-3k+2}{5} \right\rfloor.$$

Hence in either case we have $\alpha(n,k) \geq \left\lfloor \frac{n-3k+2}{5} \right\rfloor$.

Case 2. $l_0 = 3$.

Again, Isolator is aiming to claim as long a block of consecutive elements in $\{1,2,3\}$ as possible. Initially, she claims the element 2, and on the subsequent moves she claims one of the remaining elements of the set, if they are available. Given that Toucher cannot pick both 1 and 3 on her first move, Isolator can always pick at least two of these elements. In particular, at the end of such process exactly one of the following is true.

1. Isolator has claimed all three elements in $\{1,2,3\}$.

2. Isolator has claimed two consecutive elements in $\{1,2,3\}$ and Toucher has claimed the third element in $\{1,2,3\}$.

In both cases, consider the game played on $T_1 = \{5, \ldots, n\}$. Let a be the number of elements Isolator has claimed in $\{1,2,3\}$. Note that in both cases Toucher claims all the other elements in $\{1,2,3,4\}$ not claimed by Isolator, and thus Toucher claims $4-a$ elements in $\{1,2,3,4\}$. Since Toucher claims in total $a+k$ elements, it follows that she claims $k+2a-4$ elements in T_1. Since Isolator has not yet claimed any elements in T_1, it follows that on T_1 Isolator can increase the score by $\alpha(n-4, k+2a-4)$. Since she has achieved a score of $a-1$ outside T_1 with her block of a consecutive elements, it follows that the total score achieved is $a-1+\alpha(n-4, k+2a-4)$.

By induction, it follows that the score achieved is at least

$$a-1+\left\lfloor \frac{n-4-3(k+2a-4)+2}{5} \right\rfloor = \left\lfloor \frac{n-3k-a+5}{5} \right\rfloor.$$

Since $a \in \{2,3\}$, it follows that

$$\alpha(n,k) \geq \left\lfloor \frac{n-3k+2}{5} \right\rfloor.$$

Thus Lemma 38 follows by induction. $\square$

Lemma 39. *For all $n \geq 2$ we have $\gamma_b(n) \geq \lfloor \frac{n+4}{5} \rfloor$, and $\gamma_b(1) = 0$.*

Proof. When $n = 1$, the claim is trivial as the only move is given to Toucher. Now consider the case when we have $n \geq 2$.

At the start of the game, Isolator is aiming to claim as long a block of consecutive elements as possible near both of the endpoints. Once this is no longer possible, she starts using the same strategy as in Lemma 38. We start by describing this initial process formally.

Suppose that after Isolator's k^{th} move the set of elements claimed by Isolator is of the form $\{1, \dots t\} \cup \{n-k+t+1, \dots, n\}$ for some $t \in \{0, \dots, k\}$, with the convention that $\{1, \dots, t\} = \emptyset$ when $t = 0$ and $\{n-k+t+1, \dots, n\} = \emptyset$ when $t = k$. Note that this certainly holds when $k = 0$, as Isolator has not claimed any elements before her first move. If at least one of the elements $t+1$ or $n-k+t$ is not yet claimed before Isolator's $(k+1)^{th}$ move, then Isolator claims one of these elements which is still available, and thus the set of vertices claimed by Isolator is of this form also after $k+1$ moves. If both $t+1$ and $n-k+t$ are claimed by Toucher, the process stops.

The process terminates trivially, as Toucher must claim an element during the game. Let k and $t \in \{0, \dots, k\}$ be chosen so that the set of vertices claimed by Isolator at the end of the process is $\{1, \dots t\} \cup \{n-k+t+1, \dots, n\}$. Note that we must have $k \geq 1$, as Toucher cannot claim both of the elements 1 and n on her first move.

Let $T = \{t+2, \dots, n-k+t-1\}$, and note that by the choice of k and t it follows that Isolator has not claimed any elements in T. Since the process has terminated at this stage, it follows that Toucher must have claimed both of the elements $t+1$ and $n-k+t$. Since Toucher started the game, she has claimed $k+1$ elements in total, and thus $k-1$ of these elements must be in T.

Note that any increment of the score arising outside T occurs from the sets $\{1, \dots, t\}$ and $\{n-k+t+1, \dots, n\}$. On the other hand, since Isolator has not claimed any elements in T and Toucher has claimed $k-1$ elements in T, the rest of the game on T is identical to the game $F(n-k-2, k-1)$. Hence Isolator can increase the score by at least $\alpha(n-k-2, k-1)$ during the rest of the game played on T.

It is easy to check that the contribution of the score arising from the sets $\{1, \dots, t\}$ and $\{n-k+t+1, \dots, n\}$ is exactly $t+(k-t) = k$ for any choice of t. Thus by Lemma 38, it follows that Isolator can guarantee that the score is at least

$$k + \alpha(n-k-2, k-1) \geq k + \left\lfloor \frac{n-k-2-3(k-1)+2}{5} \right\rfloor = \left\lfloor \frac{n+k+3}{5} \right\rfloor .$$

Since $k \geq 1$, Isolator can always guarantee that the score is at least $\lfloor \frac{n+4}{5} \rfloor$, which completes the proof. □

5.2.2 The upper bound

In this section, all congruences are considered modulo 5 unless otherwise stated, and in such cases we omit (mod 5) from the notation. Furthermore, we write $n \equiv 0$ or 1 instead of ($n \equiv 0$ or $n \equiv 1$), and $n \not\equiv 0$ and 1 instead of ($n \not\equiv 0$ and $n \not\equiv 1$).

Lemma 40. *Suppose that T is a Toucher-Isolator game played on a disjoint union of the boards $F(l_1), \ldots, F(l_r)$, $G(m_1), \ldots, G(m_s)$ and $H(n_1), \ldots, H(n_t)$ with Isolator having the first move. Let $f(\underline{l}; \underline{m}; \underline{n})$ be the score of this game when both players play optimally. Let $N_1 = |\{i : l_i \equiv 3 \text{ or } 4\}|$, $N_2 = |\{i : m_i \equiv 0 \text{ or } 1\}|$, $N_3 = |\{i : n_i \neq 2 \text{ and } n_i \equiv 2 \text{ or } 3\}|$, $N_4 = |\{i : n_i = 2\}|$ and $N_5 = |\{i : n_i = 1\}|$. Let $\epsilon \in \{0, 1\}$ be chosen so that $N_5 \equiv \epsilon$ (mod 2). Then we have*

$$f(\underline{l}; \underline{m}; \underline{n}) \leq \sum_{i=1}^{r} \left\lfloor \frac{l_i+2}{5} \right\rfloor + \sum_{i=1}^{s} \left\lfloor \frac{m_i+5}{5} \right\rfloor + \sum_{i=1}^{t} \left\lfloor \frac{n_i+8}{5} \right\rfloor - N_4 + \epsilon - \left\lfloor \frac{N_1+N_2+N_3+\epsilon}{2} \right\rfloor . \quad (5.3)$$

By looking at the proof of Lemma 38, it is reasonable for Toucher to claim one of the points next to the point claimed by Isolator on her first move, as in this case Toucher can restrict the length of the interval created by Isolator. Such a first pair of moves splits the original board into two new boards, which motivates the idea of considering unions of disjoint boards. It might be tempting to say that Toucher can always follow Isolator to the board where she plays her next move, and hence proceed by using an inductive proof. However, sometimes Toucher may gain an 'additional move' if one of these boards has no sensible moves left (i.e. the component is $F(1)$ or $F(2)$).

Ignoring such additional moves completely would make the proof much shorter, but the bound obtained would not even be asymptotically good enough. Since Isolator is free to alternate between these two boards, she has some control on the time of the game when Toucher is given the additional move. In particular, we cannot assume that the additional moves are given at the start of the game, which was the case in Section 5.2.1. In order to keep track of such additional moves, we need to consider arbitrary disjoint unions of boards.

We start by briefly outlining the structure of the proof and explaining the motivation behind the upper bound in (5.3). The proof is an inductive argument, and we chose to apply the induction on the sum of the lengths of the paths. The aim is to prove that for any possible Isolator's initial move, there is a move for Toucher that can be used to show that (5.3) holds by induction. This move will in general depend on the position of the initial move modulo 5, however, we have to be slightly careful if the initial move is close to an endpoint of the board. For the same reason, one has to be careful with small components as well.

Since there are 3 possible types of boards, 5 possible locations of the initial move (modulo 5), and two possible cases for the size of the initial length of the component (depending on whether the initial length affects the values of N_1, N_2 or N_3 or not), it follows that there are in some sense 30 cases to be considered. In addition, we have to individually consider some of the situations when the initial move is near to an endpoint of the board. Fortunately, some of these cases can be treated simultaneously, and in general the techniques used are very similar.

In a sense, the most difficult part is rather to come up with a suitable upper bound in (5.3) which is strong enough for an inductive argument to work than the proof itself. Once a suitable upper bound is obtained, identifying the possible 'response moves' for Toucher is a reasonably easy task. Finally, the proof itself is mathematically not challenging, but it is tedious.

Why should we choose this particular upper bound in (5.3)? For $B = F(l)$, $G(m)$ or $H(n)$ (when $n \geq 3$) it turns out that Toucher can always guarantee that the score is at most $\left\lfloor \frac{l+2}{5} \right\rfloor$, $\left\lfloor \frac{m+5}{5} \right\rfloor$ or $\left\lfloor \frac{n+8}{5} \right\rfloor$ respectively. This explains the first three terms of the upper bound. Moreover, if $l \equiv 3$ or 4, $m \equiv 0$ or 1 or $n \equiv 2$ or 3, it turns out that Toucher has a strategy which allows her to force Isolator to either play the last non-trivial move, or Isolator can only attain a score which is strictly less than the previously stated bound. Hence the quantity $N_1 + N_2 + N_3$ is measuring the number of such 'additional moves'. Given such an additional move, Toucher can make another component of the board slightly shorter, which either reduces the score by one, or guarantees that she will also gain an additional move from the new board which she would not have otherwise gained.

However, one has to be careful with small values of n. Indeed, it turns out that on $H(2)$ Isolator can only increase the score by 1 (instead of 2), and Toucher cannot gain an additional move. This is the reason behind the $-N_4$-term. Also on $H(1)$, Isolator can increase the score by 2 (instead of 1), and Toucher gains an additional turn. Note that if the number of components of the form $H(1)$ is even, then Toucher can always claim a point on another component that is of the form $H(1)$. If the number of components of the form $H(1)$ is odd, she can follow a pairing strategy until the number of such boards decreases to 1. When the element on the last board of the form $H(1)$ has been claimed by Isolator, Toucher has to use the additional move elsewhere. Hence only the parity of N_5 affects the bound.

In a sense, dealing with the boards of the form $H(n)$ is the most difficult task due to irregular behaviour at both endpoints, and in particular, when n is small. Hence we start the proof by considering the boards of the form $H(n)$, and during the proof we also introduce some standard arguments that can be used when dealing with the boards of the form $F(l)$ or $G(m)$. In those cases, we do not always give full justification.

Note that the bound (5.3) may not always be tight, but by following a similar argument as presented in Section 5.2.1, one could verify that the bound is tight when applied to a single board of the form $F(l)$, $G(m)$ or $F(n)$, which is good enough for our purposes. The reason why the bound is not necessarily tight is the fact that sometimes Toucher could have a better place to play her additional move, rather than the 'worst case scenario' that is considered in the proof.

For convenience, define

$$g(\underline{l};\ \underline{m};\ \underline{n}) = \sum_{i=1}^{r} \left\lfloor \frac{l_i+2}{5} \right\rfloor + \sum_{i=1}^{s} \left\lfloor \frac{m_i+5}{5} \right\rfloor + \sum_{i=1}^{t} \left\lfloor \frac{n_i+8}{5} \right\rfloor - N_4 + \epsilon - \left\lfloor \frac{N_1+N_2+N_3+\epsilon}{2} \right\rfloor,$$

$$y(\underline{l};\ \underline{m};\ \underline{n}) = \sum_{i=1}^{r} \left\lfloor \frac{l_i+2}{5} \right\rfloor + \sum_{i=1}^{s} \left\lfloor \frac{m_i+5}{5} \right\rfloor + \sum_{i=1}^{t} \left\lfloor \frac{n_i+8}{5} \right\rfloor$$

and

$$z(\underline{l};\ \underline{m};\ \underline{n}) = -N_4 + \epsilon - \left\lfloor \frac{N_1+N_2+N_3+\epsilon}{2} \right\rfloor.$$

For later purposes, it is convenient to observe that we may rewrite z as

$$z(\underline{l};\ \underline{m};\ \underline{n}) = -N_4 - \left\lfloor \frac{N_1+N_2+N_3-\epsilon}{2} \right\rfloor. \tag{5.4}$$

Proof of Lemma 40. Define $N = \sum_{i=1}^{r} l_i + \sum_{i=1}^{s} m_i + \sum_{i=1}^{t} n_i$. The proof is by induction on N, and it is easy to check that the claim holds for all possible configurations when $N = 1$ or $N = 2$. Suppose that the claim holds whenever we have $N \leq M - 1$ for some $M \geq 3$ and suppose that $\underline{l}$, $\underline{m}$, $\underline{n}$ are chosen so that $\sum_{i=1}^{r} l_i + \sum_{i=1}^{s} m_i + \sum_{i=1}^{t} n_i = M$.

We now split the proof into several cases depending on Isolator's first move. In each case, let $S(T)$ be the maximum score that Isolator can attain given her first move and given that Toucher plays optimally.

Case 1. Isolator claims an element on $H(n_t)$ on her first move.

For convenience, we set $n = n_t$. The game $H(n)$ is played on $\{1, \ldots, n\}$, and since both endpoints of the board are symmetric, we may assume that Isolator claims an element j satisfying $j \leq \lceil \frac{n}{2} \rceil$ on her move. We prove that apart from small values of n, claiming one of $j-1$ or $j+1$ is a suitable choice for Toucher, where the choice is made depending on the value of $j \pmod 5$, as indicated in Table 5.1. If $j \geq 3$, it is easy to see that after such first pair of moves the $H(n)$-component of the board splits into disjoint union of two boards $H(a)$ and $G(b)$ for some a, b with $n = a + b + 2$. However, since the boards $H(1)$ and $H(2)$ behave in a different way compared to other boards of the form $H(n)$, it turns out to be convenient to consider the cases $j = 1$, $j = 2$ and $(j, n) = (3, 5)$ individually.

Indeed, if $4 \leq j \leq \lceil \frac{n}{2} \rceil$, the board splits into components $H(a)$ and $G(b)$ with $a \geq 3$. If $j = 3$, Toucher claims the element 2 as indicated in Table 5.1. Hence the original board splits into two boards $G(1)$ and $H(n-3)$, and the second one of these is of the form $H(1)$ or $H(2)$ only when $n = 5$, as $j \leq \lceil \frac{n}{2} \rceil$. Hence $j = 1$, $j = 2$ and $(j, n) = (3, 5)$ are the only special cases which could change the number of boards of the form $H(1)$ or $H(2)$.

Denote the new set of parameters obtained after the first pair of moves by $\underline{l}'$, $\underline{m}'$ and $\underline{n}'$, and let s_i denote the increment of the score caused by Isolator's first move. Throughout the proof it is convenient to define the quantities $d_1 = z(\underline{l};\ \underline{m};\ \underline{n}) - z(\underline{l}';\ \underline{m}';\ \underline{n}')$ and $d_2 = y(\underline{l};\ \underline{m};\ \underline{n}) - y(\underline{l}';\ \underline{m}';\ \underline{n}')$. Note that we have $g(\underline{l};\ \underline{m};\ \underline{n}) = d_1 + d_2 + g(\underline{l}';\ \underline{m}';\ \underline{n}')$.

The inductive hypothesis implies that we have $S(T) \leq g\left(\underline{l'};\ \underline{m'};\ \underline{n'}\right) + s_i$. Since our aim is to prove $S(T) \leq g(\underline{l};\ \underline{m};\ \underline{n})$, it suffices to prove that we always have $d_1 + d_2 \geq s_i$. Indeed, we will prove that for all possible Isolator's initial moves there exists a move for Toucher which satisfies $d_1 + d_2 \geq s_i$.

We start with the general case $j \geq 3$ and $n \geq 6$, and we deal with the special cases later.

Table 5.1: Choices for Toucher's first move depending on the value of j

	$F(n)$	Condition on a or b	$G(n)$	Condition on a or c	$H(n)$	Condition on a or b
$j \equiv 0$	$j+1$	$b \equiv 4$	$j-1$	$a \equiv 3$	$j-1$	$b \equiv 3$
$j \equiv 1$	$j-1$	$a \equiv 4$	$j-1$	$a \equiv 4$	$j-1$	$b \equiv 4$
$j \equiv 2$	$j+1$	$b \equiv 1$	$j+1$	$c \equiv 1$	$j+1$	$a \equiv 1$
$j \equiv 3$	$j-1$	$a \equiv 1$	$j-1$	$a \equiv 1$	$j-1$	$b \equiv 1$
$j \equiv 4$	$j-1$	$a \equiv 2$	$j+1$	$c \equiv 3$	$j+1$	$a \equiv 3$

***Case 1.1*.** $n \geq 6$, $j \geq 3$.

In this case we have $s_i = 0$, so it suffices to prove that we have $d_1 + d_2 \geq 0$. It is easy to see that N_1, N_4 and ϵ are unaffected in this case. Since N_2 certainly cannot increase and N_3 can decrease by at most 1, it follows that $d_1 \geq -1$.

Note that we have $d_2 = \left\lfloor \frac{n+8}{5} \right\rfloor - \left\lfloor \frac{a+8}{5} \right\rfloor - \left\lfloor \frac{b+5}{5} \right\rfloor$. By using the trivial upper and lower bounds $x - 1 \leq \lfloor x \rfloor \leq x$ and the fact that $n = a + b + 2$, it follows that $d_2 \geq \frac{n+3}{5} - \frac{a+b+13}{5} = \frac{-8}{5}$. Since d_2 is an integer, it follows that $d_2 \geq -1$. We now split the proof into two cases based on the values of n and j in order to improve our bounds on d_1 and d_2.

***Case 1.1.1*.** $n \equiv 2$ or 3.

We start by improving the bound on d_2. Since $n \equiv 2$ or 3, it follows that $\left\lfloor \frac{n+8}{5} \right\rfloor \geq \frac{n+7}{5}$. Hence by using the trivial bounds for the other terms, we obtain that $d_2 \geq \frac{n+7}{5} - \frac{a+b+13}{5} = \frac{-4}{5}$. Since d_2 is an integer, it follows that $d_2 \geq 0$.

First suppose that $a \equiv 2$ or 3. Then N_3 cannot decrease, so in fact, we have $d_1 \geq 0$. Hence we have $d_1 + d_2 \geq 0$.

Now suppose that $b \equiv 0$ or 1. Then N_3 decreases by at most 1 and N_2 increases by 1. Hence the sum $N_2 + N_3$ certainly cannot decrease. Thus we also have $d_1 \geq 0$, and it follows that $d_1 + d_2 \geq 0$.

Finally, suppose that $a \not\equiv 2$ and 3 and $b \not\equiv 0$ and 1. Then we have $\left\lfloor \frac{a+8}{5} \right\rfloor + \left\lfloor \frac{b+5}{5} \right\rfloor \leq \frac{a+6}{5} + \frac{b+3}{5} = \frac{a+b+9}{5}$. Note that the equality holds if and only if $a \equiv 4$ and $b \equiv 2$, but by Table 5.1 it follows that this can never happen. Hence the inequality must be strict, and it follows that $d_2 > \frac{n+7}{5} - \frac{a+b+9}{5} = 0$. Hence we must have $d_2 \geq 1$, and combining this with the trivial bound $d_1 \geq -1$, it follows that $d_1 + d_2 \geq 0$.

***Case 1.1.2*.** $n \not\equiv 2$ and 3.

Since $n \not\equiv 2$ and 3, it follows that N_3 cannot decrease. Hence we must have $d_1 \geq 0$.

First suppose that $a \equiv 2$ or 3 and $b \equiv 0$ or 1. Then both N_2 and N_3 increase by 1, and hence it follows that $d_1 \geq 1$. Combining this with the trivial bound $d_2 \geq -1$, it follows that $d_1 + d_2 \geq 0$.

Now suppose that $a \not\equiv 2$ and 3 or $b \not\equiv 0$ and 1. In both cases we can improve the upper bound of $\lfloor \frac{a+8}{5} \rfloor + \lfloor \frac{b+5}{5} \rfloor$ to $\lfloor \frac{a+8}{5} \rfloor + \lfloor \frac{b+5}{5} \rfloor \leq \frac{a+b+11}{5}$ by following the argument presented in the proof of Case 1.1.1. Note that the equality holds if and only if ($a \equiv 4$ and $b \equiv 0$) or ($a \equiv 2$ and $b \equiv 2$). However, both of these cases are impossible, as they are not compatible with Toucher's moves indicated in Table 5.1. Hence the inequality must be strict, and thus we have $d_2 > \frac{n+4}{5} - \frac{a+b+11}{5} = -1$. Thus it follows that $d_2 \geq 0$, and hence we have $d_1 + d_2 \geq 0$.

Case 1.2. $j = 1$.

Here we split the proof into three cases based on the size of n. First, we consider the case $n \geq 3$, which should be viewed as the main part of the argument. Then we consider the cases $n = 2$ and $n = 1$ individually, as these behave in a slightly different way as these boards are small. The case $n = 1$ turns out to be very tedious and lengthy, and it does not really contain any interesting ideas either. In a sense, as Toucher is forced to play a move on another component of the board, our aim is to prove that even in the worst case an additional move has a certain positive effect for Toucher.

Case 1.2.1. $n \geq 3$.

Suppose that Toucher claims the element 2. Since Isolator claimed the element 1 on the board $H(n)$ satisfying $n \geq 3$, it follows that $s_i = 1$. Hence it suffices to prove that by claiming the element 2 Toucher can guarantee that $d_1 + d_2 \geq 1$ holds. First of all, note that the board $H(n)$ is replaced with $G(n-2)$ after such pair of moves, which is non-empty as $n \geq 3$. Since $n \equiv 2$ or 3 if and only if $n - 2 \equiv 0$ or 1 and $n \geq 3$, it follows that N_3 decreases by 1 if and only if N_2 increases by 1. In particular, it follows that $d_1 = 0$. On the other hand, it is easy to see that for any n we have $d_2 = \lfloor \frac{n+8}{5} \rfloor - \left\lfloor \frac{(n-2)+5}{5} \right\rfloor = 1$. Hence we always have $d_1 + d_2 = 1$.

Case 1.2.2. $n = 2$.

Suppose that Toucher claims the element 2. Since the board $H(2)$ has only two elements, it follows that all elements of the board have been claimed after this pair of moves. Note that we certainly have $s_i = 1$ and $d_2 = \lfloor \frac{2+8}{5} \rfloor = 2$. On the other hand, it is clear that N_1, N_2, N_3 and ϵ remain unaffected while N_4 decreases by 1. Hence we have $d_1 = -1$, and thus it follows that $d_1 + d_2 = 1$.

Case 1.2.3. $n = 1$.

Since $n = 1$, it follows that $s_i = 2$. First suppose that we have $N_5 > 1$, and that Toucher claims the only element on another board that is of the form $H(1)$. Hence N_5 decreases by 2 and ϵ remains unaffected, and thus we have $d_1 = 0$. We also have $d_2 = 2\lfloor \frac{1+8}{5} \rfloor = 2$, and hence it follows that $d_1 + d_2 = s_i$.

Otherwise, we have $N_5 = 1$, and hence it follows that $\epsilon = 1$. Since the total number of points on T is strictly more than 1, there exists another component B of T.

First suppose that $B = H(2)$ and that Toucher claims the element 1. Then N_2 increases by 1, N_4 decreases by 1 and ϵ changes from 1 to 0. Hence $N_1 + N_2 + N_3 - \epsilon$ increases by 2 and

$-N_4$ increases by 1, so we have $d_1 = 0$. Note that $d_2 = \left\lfloor \frac{1+8}{5} \right\rfloor + \left\lfloor \frac{2+8}{5} \right\rfloor - \left\lfloor \frac{1+5}{5} \right\rfloor = 2$, and hence it follows that $d_1 + d_2 = s_i$.

Now suppose that $B = H(m)$ with $m \geq 3$. Suppose that Toucher claims the element 1, and hence B is replaced with $G(m-1)$. Then N_4 remains unaffected, N_3 decreases by at most 1 and N_2 increases by at most one. Since ϵ changes from 1 to 0, it follows that $N_1 + N_2 + N_3 - \epsilon$ cannot decrease, and hence we have $d_1 \geq 0$. Note that we have $d_2 = \left\lfloor \frac{1+8}{5} \right\rfloor + \left\lfloor \frac{m+8}{5} \right\rfloor - \left\lfloor \frac{m+4}{5} \right\rfloor$, and thus we have $d_2 \geq 1$.

If $m \equiv 1$, then $m - 1 \equiv 0$ and thus N_2 increases by 1 but N_3 does not decrease. Hence $N_1 + N_2 + N_3 - \epsilon$ increases by 2, and thus we have $d_1 \geq 1$. If $m \not\equiv 1$, it certainly follows that $\left\lfloor \frac{m+8}{5} \right\rfloor - \left\lfloor \frac{m+4}{5} \right\rfloor = 1$, and hence we have $d_2 \geq 2$. Hence in either case we have $d_1 + d_2 \geq 2$.

Next suppose that $B = G(m)$ and suppose that Toucher claims the element 1. Hence B is replaced with $F(m-1)$. As above, it is easy to deduce that N_4 remains unaffected and $N_1+N_2+N_3-\epsilon$ cannot decrease, and hence we have $d_1 \geq 0$. We also have $d_2 = \left\lfloor \frac{1+8}{5} \right\rfloor + \left\lfloor \frac{m+5}{5} \right\rfloor - \left\lfloor \frac{m+1}{5} \right\rfloor$, and thus it follows that $d_2 \geq 1$.

If $m \equiv 4$, then $m - 1 \equiv 3$ and hence N_1 increases by 1 but N_2 does not decrease. As ϵ changes from 1 to 0, we can similarly deduce that $d_1 \geq 1$. Otherwise, it is easy to see that we have $d_2 \geq 2$. Hence in either case we have $d_1 + d_2 \geq 2$. Note that the same argument also applies even when $m = 1$ (with the convention that $F(0)$ is the empty board).

Finally, suppose that $B = F(m)$ and suppose that Toucher claims the element $m-2$. Hence B is replaced with disjoint union of boards $F(m-3)$ and $F(2)$. The board $F(2)$ can be discarded as on this board Toucher can follow a pairing strategy and avoid any increment of the score. Again, we know that N_1 cannot decrease by more than 1, and since ϵ changes from 1 to 0 it follows that $d_1 \geq 0$. We also have $d_2 = \left\lfloor \frac{1+8}{5} \right\rfloor + \left\lfloor \frac{m+2}{5} \right\rfloor - \left\lfloor \frac{m-1}{5} \right\rfloor$, and hence we have that $d_2 \geq 1$.

If $m \equiv 3$, 4 or 5, we certainly have $d_2 \geq 2$. If $m \equiv 1$ or 2, then $m - 3 \equiv 3$ or 4, and hence N_1 increases by 1. Hence $N_1 + N_2 + N_3 - \epsilon$ increases by 2, and thus we have $d_1 \geq 1$. In either case, it follows that $d_1 + d_2 \geq 2$.

Case 1.3. $j = 2$.

Since $j \leq \left\lceil \frac{n}{2} \right\rceil$, it follows that $n \geq 3$. Hence we split the proof into cases based on whether we have $n \geq 5$, $n = 4$ or $n = 3$.

Case 1.3.1. $n \geq 5$.

Suppose that Toucher claims the element 1. Hence it follows that $s_i = 0$, and the board $H(n)$ is replaced with $H(n-2)$ after such pair of moves. Since $n - 2 \geq 3$, it follows that N_4 and ϵ remain unchanged.

First suppose that $n \not\equiv 2$ and 3. Note that in this case N_3 cannot decrease, and hence it follows that $d_1 \geq 0$. We also have $d_2 = \left\lfloor \frac{n+8}{5} \right\rfloor - \left\lfloor \frac{n+6}{5} \right\rfloor \geq 0$, and therefore it follows that $d_1 + d_2 \geq 0$.

Now suppose that $n \equiv 2$ or 3. In this case N_3 decreases by at most 1, and hence it follows that $d_1 \geq -1$. We again have $d_2 = \left\lfloor \frac{n+8}{5} \right\rfloor - \left\lfloor \frac{n+6}{5} \right\rfloor$, and since $n \equiv 2$ or 3, it follows that $d_2 \geq 1$. Thus in either case we have $d_1 + d_2 \geq 0$.

Case 1.3.2. $n = 4$.

Again suppose that Toucher claims the element 1. Hence $s_i = 0$, and since 4 is not congruent to 2 or 3 modulo 5, it follows that N_1, N_2, N_3 and ϵ remain unaffected. On the other hand, N_4 increases by 1 as after this pair of moves the board becomes $H(2)$. Hence we have $d_1 = 1$. We also have $d_2 = \lfloor \frac{4+8}{5} \rfloor - \lfloor \frac{2+8}{5} \rfloor = 0$, and thus it follows that $d_1 + d_2 = 1 > 0$.

Case 1.3.3. $n = 3$.

Again suppose that Toucher claims the element 1, and thus we have $s_i = 0$. After this pair of moves the board becomes $H(1)$, and it is easy to verify that $d_2 = \lfloor \frac{3+8}{5} \rfloor - \lfloor \frac{1+8}{5} \rfloor = 1$.

It is clear that N_1, N_2 and N_4 remain unchanged. It is easy to observe that N_3 decreases by 1 and ϵ is replaced with $1 - \epsilon$. Hence, in the worst case, $N_3 - \epsilon$ decreases by 2, and thus by (5.4) it follows that $d_1 \geq -1$. Therefore we have $d_1 + d_2 \geq 0$.

Case 1.4. $n = 5$ and $j = 3$.

Suppose that Toucher claims the element 2. Hence the board $B(5)$ splits into two boards of the form $G(1)$ and $H(2)$, and we have $s_i = 0$. Hence N_4 increases by 1, N_2 increases by 1 and N_1, N_3 and ϵ remain unaffected. Thus it follows that $d_1 \geq 1$. Since $d_2 = \lfloor \frac{5+8}{5} \rfloor - \lfloor \frac{2+8}{5} \rfloor - \lfloor \frac{1+5}{5} \rfloor = -1$, we have $d_1 + d_2 \geq 0$.

This completes the proof of Case 1.

Case 2. Isolator claims an element on $G(m_s)$ on her first move.

For convenience, we set $n = m_s$. The game $G(n)$ is played on $\{1, \ldots, n\}$, and note that in this case the board is not symmetric. Recall that claiming the element 1 increases the score by 1, but claiming the element n does not.

Assume that Isolator claims the element j on her first move. As before, we prove that claiming $j - 1$ or $j + 1$ is a suitable choice for Toucher, and this choice is again determined by the value of $j \pmod 5$. We use the same notation as before, however in this case there are two options on how the board might split: the board either splits into two components that are of the form $G(a)$ and $G(b)$ if Toucher claims the element $j - 1$, or into two components that are of the form $H(c)$ and $F(d)$ if Toucher claims the element $j + 1$. This time we only need to consider the boundary cases $j = 1$, $j = 2$ and $j = n$ individually, and note that hence we may assume that $n \geq 4$. We start by checking the special cases, and we skip some of the details when they are identical or very similar to the arguments used in the proof of Case 1.

Case 2.1. $j = 1$.

The proof is essentially identical to the proof of Case 1.2.1. Indeed, suppose Toucher claims the element 2. After the first pair of moves the board becomes $F(n-2)$ and we have $s_i = 1$. We have $d_2 = \lfloor \frac{n+5}{5} \rfloor - \lfloor \frac{n}{5} \rfloor = 1$, and as in the proof of Case 1.2.1 N_2 decreases by 1 if and only if N_1 increases by 1. Hence it follows that $d_1 = 0$, and thus we have $d_1 + d_2 = 1$.

Case 2.2. $j = 2$.

Suppose that Toucher claims the element 1. After the first pair of moves, the board becomes $G(n-2)$ and we have $s_i = 0$. Note that N_2 can decrease by at most 1, and hence it follows that $d_1 \geq -1$. We also have $d_2 = \lfloor \frac{n+5}{5} \rfloor - \lfloor \frac{n+3}{5} \rfloor$, and thus we certainly have $d_2 \geq 0$.

If $n \equiv 0$ or 1, it is easy to verify that we have $d_2 = 1$, and hence it follows that $d_1 + d_2 \geq 0$. Otherwise, N_2 cannot decrease, and hence we have $d_1 \geq 0$. Thus it follows that $d_1 + d_2 \geq 0$ in this case as well.

Case 2.3. $j = n$.

Suppose that Toucher claims the element $n - 1$. After the first pair of moves, the board becomes $G(n-2)$ and we have $s_i = 0$. Hence the proof follows by using the same steps as in the previous case.

Case 2.4. $3 \leq j \leq n - 1$.

Suppose that Toucher chooses the appropriate move as indicated in Table 5.1 depending on the value of $j \pmod 5$. Note that depending on the value of j, the board may split into two components of the form $G(a)$ and $G(b)$ or two components of the form $H(c)$ and $F(d)$. We now split the proof into 4 subcases depending on the value value of $n \pmod 5$ and depending on how the board splits into two components. As in the proof of Case 1.1, it is easy to deduce that we have the trivial lower bounds $d_1 \geq -1$ and $d_2 \geq -1$.

Case 2.4.1. $n \equiv 0$ or 1.

Note that regardless of how the board splits into two components, in either case we have $d_2 \geq \left\lfloor \frac{n+5}{5} \right\rfloor - \frac{n+8}{5}$ by using the trivial upper bound $\lfloor x \rfloor \leq x$. Since $n \equiv 0$ or 1, it follows that $d_2 \geq \frac{-4}{5}$ and thus we have $d_2 \geq 0$.

Case 2.4.1.1. $j \equiv 0$, 1 or 3.

In this case, Toucher claims the element $j - 1$, and hence the board splits into components of the form $G(a)$ and $G(b)$. Hence it follows that $d_2 = \left\lfloor \frac{n+5}{5} \right\rfloor - \left\lfloor \frac{a+5}{5} \right\rfloor - \left\lfloor \frac{b+5}{5} \right\rfloor$. Note that from Table 5.1 we can conclude that we have $a \equiv 1$, 3 or 4.

First suppose that $a \equiv 1$ or $b \equiv 0$ or 1. Then N_2 certainly does not decrease, so we have $d_1 \geq 0$, and hence it follows that $d_1 + d_2 \geq 0$.

Otherwise, we have $a \equiv 3$ or 4 and $b \equiv 2$, 3 or 4. Hence we have $\left\lfloor \frac{a+5}{5} \right\rfloor + \left\lfloor \frac{b+5}{5} \right\rfloor \leq \frac{a+2}{5} + \frac{b+3}{5} = \frac{n+3}{5}$. Since $n \equiv 0$ or 1, we have $d_2 \geq \frac{n+4}{5} - \frac{n+3}{5} > 0$, and therefore it follows that $d_2 \geq 1$. Hence we have $d_1 + d_2 \geq 0$.

Case 2.4.1.2. $j \equiv 2$ or 4.

In this case Toucher claims the element $j + 1$, and the board splits into two components of the form $H(c)$ and $F(d)$. Since $j > 2$, it follows that $j \geq 4$ and thus we have $c \geq 3$. Hence N_4 and ϵ remain unaffected. Again, we could split the proof into cases depending on whether one of $c \equiv 2$ or 3 or $d \equiv 3$ or 4 holds or not. The details are the same as in the proof of Case 2.4.1.1, and hence we omit the proof.

Case 2.4.2. $n \not\equiv 0$ and 1.

Now regardless of how the board splits into two components, we can deduce that $d_1 \geq 0$ as none of the N_i's can decrease. Again, the rest of the proof is similar to the proof of Case 1.1.2 (with appropriate modifications similar to those done in Case 2.4.1). Hence we omit the details.

This completes the proof of Case 2.

Case 3. Isolator claims an element on $F(l_r)$ on her first move.

For convenience, we set $n = l_r$. The game $F(n)$ is played on $\{1, \ldots, n\}$, and the board is again symmetric. Hence we may assume that Isolator claims an element j satisfying $j \leq \left\lceil \frac{n}{2} \right\rceil$ on her first move. In this case, the only special case that needs to be considered is $j = 1$. Again, for $j \geq 2$ claiming either $j-1$ or $j+1$ is a suitable choice for Toucher, and this choice is determined by the value of $j \pmod 5$ as indicated in Table 5.1. Apart from the case $j = 1$, the board always splits into two boards of the form $F(a)$ and $G(b)$ for some a and b with $n = a + b + 2$. We use the same notation as in the earlier cases.

Case 3.1. $j = 1$.

Suppose that Toucher claims the element 2. Then $s_i = 0$ and after the first pair of moves the board is replaced with $F(n-2)$. Hence $d_2 = \left\lfloor \frac{n+2}{5} \right\rfloor - \left\lfloor \frac{n}{5} \right\rfloor$, which is certainly always non-negative. Since N_1 decreases by at most 1, it follows that $d_1 \geq -1$.

If $n \equiv 3$ or 4, we have $d_2 \geq 1$ and hence it follows that $d_1 + d_2 \geq 0$. Otherwise, N_1 cannot decrease and hence we have $d_1 \geq 0$. Thus in either case it follows that $d_1 + d_2 \geq 0$.

Case 3.2. $j \neq 1$ and $n \equiv 3$ or 4.

The proof is identical to the proof of Case 1.1.1.

Case 3.3. $j \neq 1$ and $n \not\equiv 3$ and 4.

The proof is identical to the proof of Case 1.1.2.

This completes the proof of Claim 3, and hence Lemma 40 holds by induction. $\square$

Recall from the Introduction that $H_b(n)$ is the game played on the same board as $H(n)$, but with Toucher having the first move. Also recall that we have $u(P_n) = \gamma_b(n-1)$ and $u(C_n) = \alpha(n-1)$. We now deduce Theorem 36 from our earlier results.

Proof of Theorem 36. Let $n \geq 3$. Lemma 39 implies that we have $u(P_n) = \gamma_b(n-1) \geq \left\lfloor \frac{n+3}{5} \right\rfloor$. In order to prove the upper bound, consider the game $H_b(n-1)$ and suppose that Toucher claims the element $n-1$ on her first move. After the initial move, the subsequent game coincides with the game $G(n-2)$. Hence it follows that $\gamma_b(n-1) \leq f(\emptyset; n-2; \emptyset)$, and thus Lemma 40 implies that we have $\gamma_b(n) \leq \left\lfloor \frac{(n-2)+5}{5} \right\rfloor$. Therefore for all $n \geq 3$ we have $u(P_n) = \left\lfloor \frac{n+3}{5} \right\rfloor$.

Recall that we have $u(C_n) = \alpha(n-1)$. Hence Lemma 38 implies that $u(C_n) \geq \left\lfloor \frac{(n-1)+2}{5} \right\rfloor$, and Lemma 40 implies that $u(C_n) \leq f(n-1; \emptyset; \emptyset) = \left\lfloor \frac{(n-1)+2}{5} \right\rfloor$. Therefore for all $n \geq 3$ we have $u(C_n) = \left\lfloor \frac{n+1}{5} \right\rfloor$. In particular, for both $G = P_n$ and $G = C_n$ the asymptotic proportion of isolated vertices is $1/5$ when both players play optimally. $\square$

5.3 Toucher-Isolator game on trees

Recall that it turned out to be good for Isolator to claim consecutive edges near the endpoints when playing the Toucher-Isolator game on a path. When playing on a tree, it seems natural for Isolator to start by claiming edges that have a leaf as an endpoint, as claiming such an edge immediately increases the number of isolated vertices. Since the leaf isolated as a result of such move does not affect the subsequent game, the edge claimed by Isolator can be discarded. After discarding the edge, the other endpoint of the edge may also become a leaf, and in such case Isolator can continue claiming untouched leaves.

Now suppose that at some point of this process there is no move for Isolator which would instantly increase the number of isolated vertices. Let T_1 be the tree obtained as a result of the process and let C_1 be the set of edges claimed by Toucher. Thus for every leaf $v \in T_1$ there exists an edge $e \in C_1$ whose endpoint v is, as Isolator cannot increase the number of isolated vertices. Such a situation is very similar to the delayed game for paths introduced in the proof of Theorem. This motivates us to define a delayed version of the Toucher-Isolator game that is more suitable for trees.

Let T be a tree, let C and D be disjoint subsets of the edges of T, and let X be a subset of the vertices of T. Define the *delayed game* $F(T, C, D, X, s)$ to be the Toucher-Isolator game played on the edges of T with the edges in C and D given to Toucher and Isolator respectively at the start of the game, and with the first move given to player specified by the parameter $s \in \{\mathrm{i}, \mathrm{t}\}$. Define the *score* of the game to be the number of isolated vertices in $V(T) \setminus X$ at the end of the game, and denote the score by $\alpha(T, C, D, X, s)$. Thus X should be viewed as a set of vertices that are additionally declared to be touched at the start of the game, as they cannot increase the score even if they are isolated.

For our purposes, we mostly focus on certain subclasses of such delayed games, and hence some of the parameters can be omitted as they will be clear from the context. First of all, we use $F(T)$ to denote the ordinary Toucher-Isolator game on T, i.e. $F(T, \emptyset, \emptyset, \emptyset, \mathrm{t})$, and similarly we use $\alpha(T)$ to denote the score of $F(T)$. However, apart from this special case, it is more convenient to choose Isolator to be the player having the first move in the delayed version of the game, and hence s should be taken to be Isolator if it is omitted from the notation, with $F(T)$ being an exception. Similarly C and D should be taken to be empty if they are omitted from the notation. We often either have $X = \emptyset$ or $X = L$, where L is the set of the leaves in T. Hence we write $F(T, C, X) = F(T, C, \emptyset, X, \mathrm{i})$, $F(T, C) = F(T, C, \emptyset, \emptyset, \mathrm{i})$ and $F(T, C, L) = F(T, C, \emptyset, L, \mathrm{i})$ to simplify our notation.

Since some of the results used in the proof of Theorem 37 are proved by induction, it is convenient to introduce a suitable reduction operation that allows us to reduce the tree without increasing the score of the game. Our reduction operator is defined for the games of the form $F(T, C, D, X, s)$, and in general s is taken to be Isolator.

First we need to introduce some notation. As usual, let E and V denote the set of edges and vertices of T respectively, and let C, D and X be defined as before. Let $\hat{E} = E \setminus (C \cup D)$ be the set of edges that are not given to Toucher or Isolator at the start of the game, let I be the set of vertices in $V \setminus X$ that are isolated by the edges in D, and let O be the set of vertices in $V \setminus X$ that are touched by an edge in C. The vertices in O are called *occupied* and the vertices in $O \cup X$ are called *touched*. Finally, we set $U = V \setminus (I \cup X \cup O)$ and the vertices in U are called *unoccupied*. Note that U is the set of those vertices that could still increase the score, and in a sense they are the only interesting vertices left in the tree.

The definition of the reduction operation is quite tedious, but the ideas behind it are fairly simple, and we start by outlining these ideas. Suppose that v_1 and v_2 are two touched vertices and let e be an edge of the form uv_1 that is not in C. Let T_1 be the graph obtained by replacing the edge uv_1 with uv_2 in T and suppose that T_1 is also a tree. This operation changes the structure of T, but in a sense it does not affect the game at all. First of all, note that the process only affects the vertices u, v_1 and v_2. However, in both T and T_1 the vertices v_1 and v_2 are already touched, only one edge with u as its endpoint is affected during the process and in both trees the other endpoint of this edge is touched. Note that it does not matter which particular vertex the other endpoint is, as long as in both trees the other endpoint is touched. Hence, in fact, the game is not affected at any vertex during the process. One can also perform similar operations to leaves that are an endpoint of an edge in C and whose only neighbour has degree at least 3.

Recall that $\hat{E} = E \setminus (C \cup D)$. Let $e \in \hat{E}$, and note that neither of the endpoints of e is in I. Define the *endpoint pattern* of e to be $P(e) \in \{1, 2, 3\}$, where $P(e) = 1$ if both of the endpoints of e are unoccupied, $P(e) = 2$ if exactly one of the endpoints is unoccupied and $P(e) = 3$ if neither of the endpoints is unoccupied. Let T_1 and T_2 be trees with appropriate sets C_i, D_i and X_i. We say that a function $f : \hat{E} \to \hat{E}_2$ *preserves the type of the endpoints* if for every $e \in \hat{E}_1$, e and $f(e)$ have the same endpoint pattern. Finally, for a vertex $v \in V(T)$ define $\mathcal{E}(v)$ to be the collection of the edges whose endpoint v is.

We say that $F(T_1, C_1, D_1, X_1, s)$ is a *reduction* of $F(T_2, C_2, D_2, X_2, s)$ if $D_1 = \emptyset$ and if there exist injections $f_E : \hat{E}_1 \to \hat{E}_2$ and $f_V : U_1 \to U_2$ so that f_E preserves the type of the endpoints and we have $\mathcal{E}(f_V(v)) = f_E(\mathcal{E}(v))$ for all $v \in U_1$, where $f_E(A) = \bigcup_{e \in A} f_E(a)$ for $A \subseteq \hat{E}_1$. The first condition is intuitively clear, and the second condition implies that the neighbourhood of an unoccupied vertex is preserved. The second condition guarantees that the process of isolating an unoccupied vertex is the same in both T_1 and T_2. For convenience, we just say that T_1 is a reduction of T_2 if $F(T_1, C_1, D_1, X_1, s)$ is a reduction of $F(T_2, C_2, D_2, X_2, s)$, as the other parameters are clear from the context.

In all of our applications, T_1 is obtained by deleting some vertices from T_2 or by changing endpoints of a small number of edges. If only deletion of vertices is used in the process, we usually take f_E and f_V to be the identity maps. For convenience, if f_E and f_V are taken to be the identity maps we simply say that T_1 is a reduction of T_2 (without explicitly specifying that the maps are taken to be identity maps). If the endpoints of some edges are changed, we often still take f_V to be the identity map and we take $f_E(e) = e$ for most of the edges, apart from several exceptions involving the edges whose endpoints were changed. In such case we only specify the map f_E on such exceptional edges, and in general for any unspecified vertices $v \in U_1$ and edges $e \in \hat{E}_1$ we set $f_V(v) = v$ and $f_E(e) = e$.

Our first aim is to prove that such a reduction cannot increase the score of the game, when the effect of those vertices that are already isolated is taken into account. This essentially follows by copying the optimal strategy on T_1 to a strategy on T_2 by using the function f_E.

Lemma 41. *Let T_1 and T_2 be trees with appropriate sets C_i, D_i and X_i with $D_1 = \emptyset$ and suppose that T_1 is a reduction of T_2. Let I_2 be the set of isolated vertices in T_2. Then we have $\alpha(T_2, C_2, D_2, X_2, s) \geq |I_2| + \alpha(T_1, C_1, D_1, X_1, s)$.*

Proof. Let S_1 be a strategy on T_1 which guarantees that Isolator can isolate at least $\alpha(T_1, C_1, D_1, X_1, s)$ vertices. Consider the strategy S_2 on T_2 obtained as follows. If on her move Toucher claims an edge $e \in \hat{E}_2$ for which e is in the image of f_E, suppose that the edge

$f_E^{-1}(e)$ is assigned to Toucher on the game T_1. If she claims an edge $e \in \hat{E}_2$ that is not in the image of f_E, then assign an arbitrary edge to Toucher on T_1. On a given turn of Isolator, suppose that she should claim an edge $g \in \hat{E}_1$ on T_1 according to the strategy S_1. We define the strategy S_2 by insisting that Isolator claims the edge $f_E(g)$ on T_2 in such case. Once all the edges on T_1 are claimed, Isolator plays an arbitrary edge on T_2 on her move under the strategy S_2.

By following this strategy, at the end of the game Isolator has isolated $\alpha(T_1, C_1, D_1, X_1, s)$ vertices on T_1. Since for all $v \in U_1$ we have $f_E(\mathcal{E}(v)) = \mathcal{E}(f_V(v))$, it follows that for each isolated vertex $v \in U_1$ the appropriate vertex $f_V(v)$ is also isolated, and all of these vertices are distinct as f_V is an injection. In addition, all the vertices in I_2 are also isolated by definition, and note that we have $I_2 \cap U_2 = \emptyset$. Hence it follows that $\alpha(T_2, C_2, D_2, X_2, s) \geq |I_2| + \alpha(T_1, C_1, D_1, X_1, s)$. $\square$

5.3.1 First phase of the game

At the start of the game we say that the game is in the *first phase*, and after a given move of Toucher the game remains in the first phase if there exists an unoccupied vertex v for which $\mathcal{E}(v)$ contains exactly one edge that is not already claimed by Isolator. Otherwise, the game moves to the *second phase*, and note that this transition always occurs after Toucher's move. In particular, the game is in the first phase as long as Toucher can increase her score on every move by claiming a suitable edge - it turns out that choosing an arbitrary edge among the available candidates will be good enough for our purposes.

Let C and D be the set of edges occupied by Toucher and Isolator when the game moves from the first phase to the second phase. Recall that for a tree T we write L for the set of leaves in T. Our first aim is to show that there exists a reduction T' of T with $D' = \emptyset$, $X' = L'$ and for which $|T'| - 3|C'| - 3|L'|$ is not too small. This is done in Lemma 42. Note that the game on T' corresponds to the delayed game $F(T', C', L')$, as $X' = L'$, $D' = \emptyset$ and since Isolator has the first move in the second phase. Thus in order to analyse the second phase, we need a lower bound for $\alpha(T, C, L)$. In Lemma 43 we prove a lower bound for $\alpha(T, C, L)$ that depends on the value of $|T| - 3|C| - 3|L|$.

Lemma 42. *Let T be a tree with $n \geq 3$ vertices. Suppose that Isolator has the next move and that the game is in the first phase. Let Y be the set of those edges e in E that are not yet played for which Isolator can immediately isolate a new vertex by claiming e on her current move.*

Suppose that on each of her move Isolator claims an arbitrarily chosen edge from Y. Let r be the number of edges Isolator claims during the first phase, and let C and D be the set of edges claimed by Toucher and Isolator at the end of the first phase. Then there exists a reduction T' of the game $F(T, C, D)$ with $X' = L'$, $D' = \emptyset$ and $|T'| - 3|L'| - 3|C'| \geq |T| - 5r - 4$.

Proof. Let $C = \{e_1, \ldots, e_{r+1}\}$ and $D = \{f_1, \ldots, f_r\}$ be the set of edges claimed by Toucher and Isolator respectively at the end of the first phase, ordered in a way that f_i is claimed before f_j for $i < j$. Let v_i be the vertex isolated by claiming the edge f_i.

We start by verifying that $T \setminus \{v_1, \ldots, v_i\}$ is a tree for all i, and that claiming f_i cannot isolate both of its endpoints. Indeed, note that v_1 must be a leaf, and since $n \geq 3$ it follows that no two leaves can be neighbours. Hence the claim is true when $i = 1$. If the claim is true for all $1 \leq j \leq i$ for some i, it follows that $T \setminus \{v_1, \ldots, v_i\}$ is a tree which does not contain any isolated vertices. Note that it also does not contain any edge claimed by Isolator on her first i moves, as for every edge $f_1, \ldots, f_i$ at least one of the endpoints is deleted during the process. Thus v_{i+1} must be a leaf in $T \setminus \{v_1, \ldots, v_i\}$, and note that the unique neighbour of v_{i+1} cannot be a leaf. Indeed, this follows from the fact that $T \setminus \{v_1, \ldots, v_i\}$ contains an edge claimed by Toucher, and hence it contains at least 3 vertices. Thus both claims follow by induction.

Let T' be the tree obtained by deleting the vertices $v_1, \ldots, v_r$, and note that $C \subseteq E(T')$ as none of the vertices $v_1, \ldots, v_r$ is touched. Hence it follows that T' is a reduction of T when we choose $C' = C$, $D' = \emptyset$ and $X' = \emptyset$. Our aim is to reduce T' further by constructing a suitable sequence of reductions $T_0, \ldots, T_t$ for some t with $T_0 = T'$, and for which T_t satisfies $|T_t| - 3|L_t| - 3|C_t| \geq |T| - 5r - 4$. For each i define $g_i = |T_i| - 3|X_i| - 3|C_i|$.

The sequence $T_0, \ldots, T_t$ is obtained as follows. First of all, we take $T_0 = T'$. Given T_{i-1} together with appropriate sets satisfying $X_{i-1} \subseteq L_{i-1} \subseteq X_{i-1} \cup O_{i-1}$ and $D_{i-1} = \emptyset$, we stop the process if we have $X_{i-1} = L_{i-1}$. Otherwise, there exists a leaf $v \in T_{i-1}$ whose unique neighbour w satisfies $vw \in C_{i-1}$. Indeed, since $X_{i-1} \neq L_{i-1}$, it follows that there exists a leaf $v \in L_{i-1} \setminus X_{i-1} \subseteq O_{i-1}$. Let w be chosen so that $N(v) = \{w\}$. Since $v \in O_{i-1}$, we must have $vw \in C_{i-1}$.

Let $v \in L_{i-1}$ and let w be the unique neighbour of v for which we have $vw \in C_{i-1}$. We will prove that in every case we have $g_i \geq g_{i-1} - 1$, and that the property $X_i \subseteq L_i \subseteq X_i \cup O_i$ is preserved.

Case 1. w satisfies $d_{T_{i-1}}(w) = 2$.

Consider T_i obtained by deleting the vertex v and by setting $C_i = C_{i-1} \setminus \{vw\}$, $D_i = \emptyset$ and $X_i = (X_{i-1} \setminus \{v\}) \cup \{w\}$. Note that T_i is certainly a reduction of T_{i-1} as w is a touched vertex in both T_{i-1} and T_i. We certainly have $|C_i| = |C_{i-1}| - 1$ and $|T_i| = |T_{i-1}| - 1$. Since we might also have $v \notin X_{i-1}$, in the worst case we have $|X_i| \leq |X_{i-1}| + 1$. Hence it follows that $g_i \geq g_{i-1} - 1$, and it is easy to see that we also have $X_i \subseteq L_i \subseteq X_i \cup O_i$.

Case 2. w satisfies $d_{T_{i-1}}(w) \geq 3$ and $|\mathcal{E}(w) \cap C_{i-1}| \geq 2$.

Since $|\mathcal{E}(w) \cap C_{i-1}| \geq 2$, it follows that there exists an edge $uw \in C_{i-1}$ with $u \neq v$. Consider T_i obtained by deleting the vertex v and by setting $C_i = C_{i-1} \setminus \{vw\}$, $D_i = \emptyset$ and $X_i = (X_{i-1} \setminus \{v\})$. As before, this is a reduction of T_{i-1}, as w is touched in both T_{i-1} and T_i since $uw \in C_i$. Since $d_{T_{i-1}}(w) \geq 3$, it follows that w is not a leaf in T_i, and hence we have $X_i \subseteq L_i \subseteq X_i \cup O_i$. It is easy to check that we have $g_i \geq g_{i-1} + 2$.

Case 3. w satisfies $d_{T_{i-1}}(w) \geq 3$ and $|\mathcal{E}(w) \cap C_{i-1}| = 1$.

Let $N_{T_{i-1}}(w) \setminus \{v\} = \{v_1, \ldots, v_s\}$. Since $d_{T_{i-1}}(w) \geq 3$, it follows that $s \geq 2$, and since $|\mathcal{E}(w) \cap C_{i-1}| = 1$ it follows that $wv_j \notin C_{i-1}$ for all j. Consider T_i obtained by replacing the edge wv_1 with vv_1 and by setting $C_i = C_{i-1}$, $D_i = \emptyset$ and $X_i = X_{i-1} \setminus \{v\}$. Since $vw \in E_{i-1}$ and T_{i-1} is a tree, it follows that T_i is a connected graph which does not contain a cycle, and hence T_i is also a tree. Again, we obtain that T_i is a reduction of T_{i-1} by taking $f_E(wv_1) = vv_1$. Indeed, this follows from the fact that both v and w are touched vertices.

Note that we have $L_i = L_{i-1} \setminus \{v\}$, as the only vertices whose degrees are affected during the process are v and w, and since $s \geq 2$ it follows that w is not a leaf in T_i. It is easy to see that we have $|T_i| = |T_{i-1}|$, $|X_i| \leq |X_{i-1}|$ and $|C_i| = |C_{i-1}|$. Hence it follows that $g_i \geq g_{i-1}$, and it is also easy to see that we have $X_i \subseteq L_i \subseteq X_i \cup O_i$.

Note that we still need to verify that any sequence of such operations will terminate in a finite time. During every application of Case 1, the number of vertices in T_i decreases by 1, yet the size of T_i remains unaffected in Cases 2 and 3. Thus Case 1 can be applied at most $|T_0|$ times. On the other hand, the number of leaves in T_i decreases by 1 during every application of Cases 2 or 3, so the number of times Cases 2 or 3 can be applied consecutively without applying Case 1 is at most the number of vertices at that particular stage. Hence the total number of applications is at most $|T_0|(|T_0|+1)/2$, which proves that the process must terminate in a finite time.

Let $T_0, \ldots, T_t$ be the sequence of reductions obtained during the process, and let a be the number of times Case 1 is applied. Since $g_i \geq g_{i-1} - 1$ whenever Case 1 is applied, $g_i \geq g_{i-1} + 2$ whenever Case 2 is applied and $g_i \geq g_{i-1}$ whenever Case 3 is applied, it follows that $g_t \geq g_0 - a$.

On the other hand, note that $|C_i| = |C_{i-1}|$ whenever Case 3 is applied, yet $|C_i| = |C_{i-1}| - 1$ whenever Cases 1 or 2 are applied. Since $|C_t| \geq 0$ and $|C_0| = r+1$, it follows that $a \leq r+1$. Thus we must have $g_t \geq g_0 - (r+1)$. Since $X_t = L_t$ and $X_0 = \emptyset$, it follows that

$$|T_t| - 3\,|L_t| - 3\,|C_t| \geq |T_0| - 3\,|C_0| - (r+1)\,.$$

Since $|T_0| = |T| - r$ and $|C_0| = r+1$, it follows that

$$|T_t| - 3\,|L_t| - 3\,|C_t| \geq |T| - 5r - 4.$$

Since T_t is a reduction of T, this completes the proof. □

5.3.2 Delayed version of the game

Let T_t, C_t and L_t be the reduction provided by Lemma 42 and let r be the number of edges claimed by Isolator during the first phase of the game. Then Lemma 41 implies that we have $\alpha(T) \geq r + \alpha(T_t, C_t, L_t)$. Since $|T_t| - 3\,|L_t| - 3\,|C_t| \geq |T| - 5r - 4$, it suffices to prove that for all trees T with n vertices, l leaves and for any set of edges $C \subseteq E$ we have $\alpha(T, C, L) \geq \left\lfloor \frac{n-3l-3|C|+7}{5} \right\rfloor$.

However, for the purposes of the inductive proof it turns out to be convenient to prove a slightly stronger statement. Recall that O is the set of occupied vertices, and since $X = L$ it follows that O is the set of those vertices of degree at least 2 which are an endpoint of an edge in C. Our aim is to prove the following result.

Lemma 43. *Let T be a tree with n vertices and l leaves. Let O be the set of occupied vertices of T and let $C \subseteq E$. Then we have*

$$\alpha(T, C, L) \geq \left\lfloor \frac{n - 3l - 3\,|C| + 7 + \sum_{v \in O} (d(v) - 2)}{5} \right\rfloor. \tag{5.5}$$

In particular, it follows that $\alpha(T, C, L) \geq \left\lfloor \frac{n-3l-3|C|+7}{5} \right\rfloor$.

The proof is an inductive proof first on the size of T and then on the number of leaves in T; however, for simplicity one could view it just as an inductive proof on the size of T. Given a tree T whose leaves are touched and with some edges claimed by Toucher at the start, our aim is to either find suitable edges for Isolator that can help to isolate some vertices, or find suitable substructures of the tree with many touched edges that could be removed without deleting too many vertices that could possibly be isolated. In either case, our aim is to reduce the size of T without reducing the lower bound.

The structures we are in general looking for are two neighbouring vertices of degree 1 or 2, as near such vertices T behaves similarly compared to a path. If no such substructure exists, it follows that the vertices of degree 1 or 2 must be spread out. In particular, there must be vertices of higher degree, which implies that T contains plenty of leaves. The aim of the next Lemma is to make this argument precise. As a consequence, it turns out that if no suitable substructure of T exist, then lower bound of (5.5) cannot be positive, and hence the claim is certainly true.

Since there are several substructures we are considering, the proof splits into many cases and the proofs of some cases are rather long. This is due to the fact that for each substructure, the proof often splits into multiple subcases based on the structure of T on vertices near the substructure. In general, the proofs are fairly easy within each case, and similar ideas are repeatedly used in different cases. In a sense, the most difficult idea is to come up with a suitable lower bound in (5.5) that is strong enough for an inductive argument.

Lemma 44. *Let T be a tree with $n \geq 3$ vertices containing no two adjacent vertices of degree 2 and no leaf adjacent to a vertex of degree 2. Then T contains at least $\frac{n+5}{3}$ leaves.*

Proof. Let d_i be the number of vertices of degree i in T. Since a tree on n vertices contains $n-1$ edges, it is easy to verify that we have

$$\sum_{i=1}^{n} id_i = 2\left(n-1\right) \tag{5.6}$$

and

$$d_1 = 2 + \sum_{i=3}^{n} \left(i-2\right) d_i. \tag{5.7}$$

Let X be the set of vertices in T which have degree 1 or 2, and let Y be the set of vertices in T which have degree at least 3. Note that there are no edges inside X. Indeed, trivially no two leaves can be adjacent in any tree with at least 3 vertices, and by assumption no leaf is adjacent to a vertex of degree 2 and no two vertices of degree 2 are adjacent. Hence it follows that

$$d_1 + 2d_2 = e\left(X, Y\right) \leq \sum_{y \in Y} d\left(y\right) = \sum_{i=3}^{n} id_i. \tag{5.8}$$

Thus (5.6) with (5.8) imply that

$$\sum_{i=3}^{n} id_i \geq n-1.$$

Since we have $3\left(i-2\right) \geq i$ for all $i \geq 3$, it follows that

$$\sum_{i=3}^{n} \left(i-2\right) d_i \geq \frac{1}{3} \sum_{i=3}^{n} id_i \geq \frac{n-1}{3}.$$

Thus (5.7) implies that $d_1 \geq 2 + \frac{n-1}{3} = \frac{n+5}{3}$, which completes the proof. $\square$

Now we are ready to prove Lemma 43.

Proof of Lemma 43. The proof is by induction on n, and for a fixed n we also induct on the number of leaves. Let $C = \{e_1, \ldots, e_k\}$ be the set of edges claimed by Toucher at the start of the game, and for convenience we write $l = |L|$ and $k = |C|$ throughout the proof. We start by checking the base cases. For a fixed n, note that the claim follows if T is a path by Theorem 36. Hence for a given n, the base case for the induction on the number of leaves holds. Thus we may always assume that T is not a path.

Next we prove that the claim holds whenever $n \leq 5$. Since T is not a path, we must have $l \geq 3$. Note that we always have $\sum_{v \in O} (d(v) - 2) \leq \sum_{v \notin L} (d(v) - 2) = l - 2$. Hence for $n \leq 5$ and $l \geq 3$ we have

$$n + 7 - 3l - 3k + \sum_{v \in O} (d(v) - 2) \leq n + 5 - 2l \leq 4,$$

which completes the proof, since we always have $\alpha(T, C, L) \geq 0$. Thus, from now on, we may assume that T has at least 6 vertices and that T is not a path.

We split the proof into cases based on whether T contains suitable substructures. At the end we verify by using Lemma 44 that if T contains none of these substructures, we must have $n - 3l - 3k + 7 + \sum_{v \in O} (d(v) - 2) < 5$, in which case the claim follows trivially.

In the first four cases we consider the situations when C contains two edges that are 'close' to each others or when C contains an edge close to a leaf. In those cases we prove that a suitable part of the tree can be removed already before the start of the game in a way that the resulting tree is a reduction of the original tree, and so that this reduction does not decrease the score. In particular, note that we have $D = I = \emptyset$ in such cases. In the remaining two cases we consider situations when T contains sufficiently many neighbouring vertices of degree 2. In such cases Isolator can increase the score by claiming the edges joining such vertices.

Let us first focus on those cases in which we can simply reduce T before the game starts. Given a reduction T_1 of T with appropriate sets C_1, L_1 and $D_1 = \emptyset$, for convenience we define $d(k) = |C| - |C_1|$, $d(l) = |L| - |L_1|$, $d(n) = |T| - |T_1|$ and

$$d(s) = \sum_{v \in O} (d_T(v) - 2) - \sum_{v \in O_1} (d_{T_1}(v) - 2).$$

For $v \in V(T)$, we define

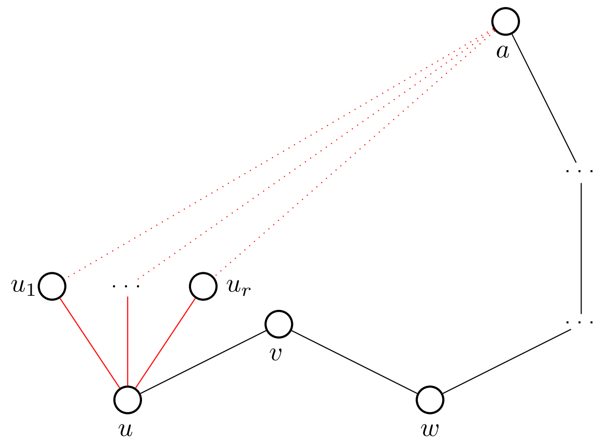
$$d_s(v) = (d_T(v) - 2)\,\mathbb{I}\{v \in O\} - (d_{T_1}(v) - 2)\,\mathbb{I}\{v \in O_1\},$$

where as usual $\mathbb{I}$ denotes the indicator function of an event, and note that we have

$$d(s) = \sum_{v \in V \cup V_1} d_s(v).$$

Finally, we define $D(T, T_1) = d(n) - 3d(l) - 3d(k) + d(s)$. Note that this also depends on the sets C and C_1, but the dependence will not be highlighted in the notation as these sets are clear

Figure 5.1: Construction of T_1. Red rigid edges are edges in T which are replaced with red dotted edges.



from the context.

Let

$$S(T)=\left\lfloor\frac{n-3l-3k+7+\sum_{v\in O}(d(v)-2)}{5}\right\rfloor,$$

and again note that S also depends on C. If $d(n)>0$, the inductive hypothesis implies that we have $\alpha(T_1,C_1,L_1)\geq S(T_1)$. If we also had $D(T,T_1)\leq 0$, it would certainly follow that $S(T_1)\geq S(T)$. Since T_1 is a reduction of T, Lemma 41 implies that $\alpha(T,C,L)\geq\alpha(T_1,C_1,L_1)$, as $D=\emptyset$ implies that $I=\emptyset$. Combining these together, we obtain that $\alpha(T,C,L)\geq S(T)$. Thus we can conclude that when $|T|>|T_1|$, the inductive step follows if we can prove that $D(T,T_1)\leq 0$. We now move on to considering various substructures of T.

Case 1. T contains an unoccupied vertex of degree 2 both of whose neighbours are touched.

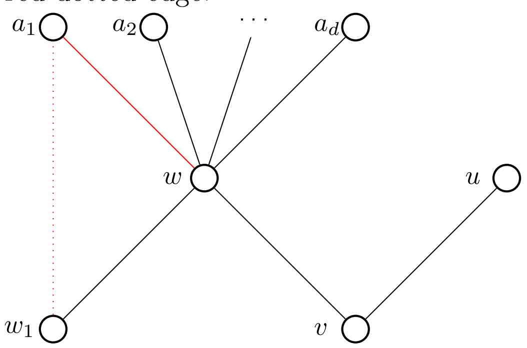
Since $n\geq 3$, it follows that either both of the vertices are occupied, or one of them is occupied and the other is a leaf. We start by reducing the first case to the second case.

Let v be the unoccupied vertex of degree 2, and let u and w be the neighbours of v. By the assumption we know that u and w are occupied. Let $N(u)\setminus\{v\}=\{u_1,\dots,u_r\}$. Since u is occupied and v is unoccupied, there exists j for which we have $uu_j\in C$. Let a be a leaf in T so that every path from a to u must go through w. Consider T_1 obtained by deleting all the edges with u as an endpoint apart from the edge uv and adding the edges $au_1,\dots au_r$, as illustrated in Figure 5.1. We also set C_1 to be the set containing all the edges in C that do not have u as an endpoint, and together with the edges of the form au_i for those i for which we also have $uu_i\in C$.

It is easy to see that T_1 is a reduction of T by taking $f_E(uu_i)=au_i$. Indeed, this follows from the fact that a is touched in T_1 as $au_j\in C$. It is also easy to check that we have $d(k)=0$ and $d(n)=0$. We also have $L_1=(L\setminus\{a\})\cup\{u\}$, which implies that $d(l)=0$. Note that a and u are the only vertices whose degrees are affected during the process. Since we have $d_s(a)=1-r$ and $d_s(u)=r-1$, it follows that $d(s)=0$. In particular, we have $S(T)=S(T_1)$, and thus by Lemma 41 it suffices to only consider the case when one of the neighbours is a leaf.

Hence we may assume that T contains an unoccupied vertex v of degree 2 with neighbours u and w so that u is a leaf and w is occupied. Let w_1 be chosen so that $ww_1\in C$. Since T has at least 6 vertices, we must have $d(w)+d(w_1)\geq 4$. We start by considering the cases

Figure 5.2: Construction of T_1 when $d(w) + d(w_1) \geq 5$ and $c = 0$. Red rigid edge is the edge in T that is replaced with the red dotted edge.



corresponding to $d(w)+d(w_1) = 4$, and it is easy to check that this condition is satisfied exactly when $(d(w), d(w_1)) \in \{(2,2), (3,1)\}$.

If $d(w) = d(w_1) = 2$, consider T_1 obtained by deleting the vertices u, v and w, and by taking $C_1 = C \setminus \{ww_1\}$. Since w_1 is a leaf in T_1, it follows that T_1 is a reduction of T. Since $L_1 = (L \setminus \{v\}) \cup \{w_1\}$, it follows that $d(l) = 0$. It is also easy to check that we have $d(n) = 3$, $d(k) = 1$ and $d(s) = 0$. Thus we have $D(T, T_1) = 0$, and since $|T| > |T_1|$ the claim follows by induction.

If $d(w) = 3$ and $d(w_1) = 1$, let x be chosen so that $N(w) = \{x, v, w_1\}$. Consider T_1 obtained by deleting the vertices u, v and w_1, and by taking $C_1 = C \setminus \{ww_1\}$. Since $N_T(w) = \{x, v, w_1\}$, it follows that w is a leaf in T_1, and hence T_1 is a reduction of T. Note that we have $L_1 = (L \setminus \{u, w_1\}) \cup \{w\}$, and w is the only vertex whose degree is affected during the process. Hence we have $d(n) = 3$, $d(l) = 1$ and $d(k) = 1$, and since $d_s(w) = 1$ we also have $d(s) = 1$. Hence it follows that $D(T, T_1) = -2$, and since $|T| > |T_1|$ the claim follows by induction.

Now suppose that we have $d(w) + d(w_1) \geq 5$. Let $N(w_1) \setminus \{w\} = \{a_1, \ldots, a_c\}$ and $N(w) \setminus \{v, w_1\} = \{a_{c+1}, \ldots, a_d\}$ where one of these sets might be empty. Note that the condition $d(w) + d(w_1) \geq 5$ implies that we have $d \geq 2$. Consider T_1 obtained by deleting the vertices u and v, and by taking $E(T_1)$ to be the set of those edges in T that do not have w or w_1 as their endpoint together with the edges ww_1, w_1a_1 and wa_i for $2 \leq i \leq d$. See Figure 5.2 for illustration when $c = 0$. Finally, we take C_1 to be the set containing the edge ww_1, all the edges in C that do not have w or w_1 as an endpoint, and the unique edge in $\{wa_i, w_1a_i\} \cap E(T_1)$ for those i for which the one of wa_i or w_1a_i that is an edge in T is also contained in C. In particular, it follows that $|C_1| = |C|$.

Note that T_1 is a reduction of T by taking $f(wa_i) = a_1a_i$ for all $2 \leq i \leq c$ since both w and w_1 are touched in T and T_1. It is easy to check that we have $d(n) = 2$, $d(l) \geq 1$ and $d(k) = 0$. It is also easy to see that the only vertices whose degrees are affected are v, w and w_1. We have $d_s(v) = 0$, $d_s(w) = (d - c) - (d - 1) = 1 - c$ and $d_s(w) = \max(c - 1, 0) - 0 \leq c$ as $c \geq 0$. In particular, it follows that $d(s) \leq 1$, and hence we have $D(T, T_1) \leq 0$. Since $|T| > |T_1|$, the claim follows by induction. This completes the proof of Case 1.

Case 2. T contains an edge $e \notin C$ both of whose endpoints are touched.

There are again two possibilities: either both endpoints of e are occupied or one of them is occupied and the other one is a leaf. By using the same argument as in the proof of Case 1, we may assume that one of the endpoints is occupied and the other one is a leaf. Let u and v be the endpoints of e so that u is a leaf and v is occupied, and let w be chosen so that $vw \in C$. Again, we split the proof into cases based on the size of $d(v) + d(w)$, and since T has at least 6 vertices, we have $d(v) + d(w) \geq 4$. Again, the condition $d(v) + d(w) = 4$ implies that we have $(d(v), d(w)) \in \{(2,2), (3,1)\}$.

If $d(v) = d(w) = 2$, consider T_1 obtained by deleting the vertices u and v, and by taking $C_1 = C \setminus \{vw\}$. Since $d_T(w) = 2$, it follows that w is a leaf in T_1, and hence T_1 is a reduction of T. It is easy to check that we have $d(n) = 2$, $d(l) = 0$ and $d(k) =1$. Since w is the only vertex whose degree is affected during the process and $d_s(w) = 0$, it follows that $d(s) = 0$. Hence we have $D(T, T_1) = -1$, and since $|T| > |T_1|$ the claim follows by induction.

If $d(v) = 3$ and $d(w) = 1$, let T_1 be the tree obtained by deleting the vertices u and w, and by taking $C_1 = C \setminus \{vw\}$. Since v is a leaf in T_1, it follows that T_1 is a reduction of T. It is easy to check that we have $d(n) = 2$, $d(l) = 1$ and $d(k) = 1$. Note that v is the only vertex whose degree is affected during the process, and since $d_s(v) = 1$ it follows that $d(s) = 1$. Hence we have $D(T, T_1) = -3$, and since $|T| > |T_1|$ the claim follows by induction.

Finally, suppose that $d(v) + d(w) \geq 5$. Let $N(v) \setminus \{u, w\} = \{a_1, \ldots, a_c\}$ and $N(w) \setminus \{v\} = \{a_{c+1}, \ldots, a_d\}$, where one of these sets might be empty. Note that the condition $d(v)+d(w) \geq 5$ implies that we have $d \geq 2$. Consider T_1 obtained by deleting the vertex u, and similarly as in the proof of Case 1 we take $E(T_1)$ to be the set of those edges in T that do not have v or w as their endpoint together with the edges vw, va_1 and wa_i for $2 \leq i \leq d$. Again, we choose C_1 to be the set containing the edge vw, all the edges in C that do not have v or w as an endpoint, and for those i for which the one of va_i or wa_i that is an edge in T is also in C, the unique edge in $\{va_i, wa_i\} \cap E(T_1)$ is also included in C_1. As before, we have $|C_1| = |C|$. As in the proof of Case 1, since v and w are occupied in both T and T_1, it follows that T_1 is a reduction of T by taking $f_E(va_i) = wa_i$ for all $2 \leq i \leq c$.

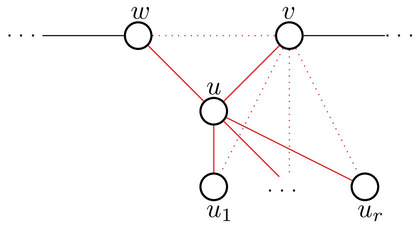
Note that we have $d(n) = 1$, $d(l) \geq 1$ and $d(k) = 0$. Since v and w are the only vertices whose degrees are affected during the process, and since we have $d_s(v) = c$ and $d_s(w) = \max(d - c - 1, 0) - (d - 2) \leq 2 - c$, it follows that $d(s) \leq 2$. Hence we have $D(T, T_1) \leq 0$, and since $|T| > |T_1|$ the claim follows by induction. This completes the proof of Case 2.

Case 3. There exists an edge $e \in C$ whose endpoint is a leaf.

Let u and v be the endpoints of e with u being a leaf. First suppose that we have $d(v) = 2$. Let T_1 be the tree obtained by deleting the vertex u, and by taking $C_1 = C \setminus \{uv\}$. Since v is touched in both T_1 and T, it follows that T_1 is a reduction of T. It is easy to check that we have $d(n) = 1$, $d(l) = 0$, $d(k) = 1$ and $d(s) = 0$. Thus we have $D(T, T_1) = -2$, and since $|T| > |T_1|$ the claim follows by induction.

Now suppose that we have $d(v) \geq 3$, and let $N(v) \setminus \{u\} = \{v_1, \ldots, v_c\}$ where $c \geq 2$. Consider T_1 obtained by replacing the edge vv_1 with uv_1 and by taking C_1 to be the set of all edges in C that are also edges in T_1, and if $vv_1 \in C$ then uv_1 is also included in C_1. It is easy to see that T_1 is a reduction of T by taking $f_E(vv_1) = uv_1$ since u and v are touched vertices in T_1 and T. It is easy to check that we have $d(n) = 0$, $d(l) = 1$ and $d(k) = 0$. Note that u and v are the only vertices whose degrees are affected during the process, and we clearly have $d_s(v) = 1$ and $d_s(u) = 0$. Hence it follows that $d(s) = 1$, and thus we have $D(T, T_1) = -2 \leq 0$. Since

Figure 5.3: Construction of T_1. Red rigid edges are edges in T which are replaced with red dotted edges.



the number of vertices remains the same and the number of leaves decreases by one, the claim follows by induction. This completes the proof of Case 3.

Case 4. There exist distinct edges $e_i, e_j \in C$ which have a common endpoint.

Let u be the common endpoint of e_i and e_j, and let v and w be the other endpoints respectively. By Case 3, we may assume that neither of v or w is a leaf. Let $N(u) \setminus \{v, w\} = \{v_1, \ldots, v_r\}$ with possibly $r = 0$. Consider T_1 obtained by removing the vertex u together with all the edges that have u as an endpoint, and by adding the edges vw and vv_i for all $1 \leq i \leq r$ as illustrated in Figure 5.3. It is easy to check that T_1 is a tree. Let C_1 be the set containing the edge vw, all the edges in C that do not have u as an endpoint and all the edges vv_i for those i for which we also have $uv_i \in C$. Since both u and v are touched, it follows that T_1 is a reduction of T by taking $f_E(uv_i) = vv_i$ for all i.

It is easy to check that we have $d(n) = 1$, $d(l) = 0$ and $d(k) = 1$. Note that the only vertices whose degrees are affected during the process are u, v and w. Since neither of v or w is a leaf, it is easy to check that $d_s(u) = r$, $d_s(v) = (d_T(v) - 2) - (d_T(v) + r - 2) = -r$ and $d_s(w) = 0$. In particular, it follows that $d(s) = 0$, and hence we have $D(T, T_1) = -2$. Since $|T| > |T_1|$, the claim follows by induction, and this completes the proof of Case 4.

From now on, we suppose that T contains no configuration described in Cases 1-4, and hence the edges in C are suitably 'isolated'. We now consider the cases when T contains two adjacent vertices of degree 2 that are both unoccupied. In such case the edges incident with these two vertices could be suitable moves for Isolator. Our aim is to describe a sequence of moves for Isolator that allows her to increase the score in a way that the resulting tree (together with the new moves) has a reduction with sufficiently large score.

From now on, we change our notation slightly: let C' denote the set of edges claimed by Toucher at the start of the delayed game and let O' be the set of occupied vertices at the start of the delayed game. Let D be the set of edges claimed by Isolator during the new moves, and let $\hat{C}$ be the set of edges claimed by Toucher during the new moves, and for convenience we write $\hat{C} = \{f_1, \ldots, f_{|D|}\}$. Finally, we set $C = C' \cup \hat{C}$, and hence C is the set of edges claimed by Toucher at the end of the process, i.e. once the new moves have been played.

In all the cases we are about to consider, the edges in D form a path in T so that all the vertices on this path apart from the endpoints have degree 2 in T, and the endpoints have degree at least 3 or are touched at the end of the process. In particular, it follows that the number of vertices isolated during the process is exactly $|D| - 1$.

Our aim is again to seek for a suitable reduction T_1 of T for these choices of C, D and $X = L$, and as usual we require that $D_1 = \emptyset$ and $X_1 = L_1$. Since there are exactly $|D| - 1$ isolated vertices, Lemma 41 implies that we have $\alpha(T, C, D, L) \geq (|D| - 1) + \alpha(T_1, C_1, D_1, L_1)$. If $|T| > |T_1|$, the inductive hypothesis implies that $\alpha(T_1, C_1, D_1, L_1) \geq S(T_1)$. Note that during the process of claiming new edges we fix a suitable strategy for Isolator, but we allow Toucher to play arbitrary edges on her moves. Hence it follows that $\alpha(T, C', L) \geq \alpha(T, C, D, L)$, as playing the edges in D corresponds to a certain choice of strategy, which may or may not be optimal.

Define $d(n) = |T| - |T_1|$, $d(l) = |L| - |L_1|$, $d(k) = |C| - |C'|$ and let

$$d(s) = \sum_{v \in O'} (d(v) - 2) - \sum_{v \in O} (d(v) - 2).$$

In particular, note that $d(k)$ and $d(s)$ are defined for the initial set-up of the delayed game, and not for the set-up containing the new edges that are played. As before, we define $D(T, T_1) = d(n) - 3d(l) - 3d(k) + d(s)$. Again, if we had $D(T, T_1) \leq 5(|D| - 1)$, it would follow that $S(T_1, C_1) + (|D| - 1) \geq S(T, C')$ and hence it would follow that $\alpha(T, C', L) \geq S(T, C')$, where the dependence on the set of claimed edges C' is highlighted in the notation for clarity. Hence our aim is to prove that we always have $D(T, T_1) \leq 5(|D| - 1)$.

We now focus on the unoccupied vertices of degree 2. First we consider a case when there exists such an unoccupied vertex whose neighbour is a touched vertex, although this case splits into a large number of subcases. Note that if no such unoccupied vertex of degree 2 exists, then by following a path of unoccupied vertices of degree 2 it follows that both of the endpoints of such paths are unoccupied vertices of degree at least 3, and not touched vertices.

Case 5. There exists an unoccupied vertex of degree 2 whose neighbour is touched.

Let v_1 be an unoccupied vertex of degree 2 and let v_0 be the neighbour of v_1 that is touched. Note that by using the same argument as in the proof of Case 1 we may assume that v_0 is a leaf. We start by constructing a sequence of vertices $v_0, v_1, \ldots, v_m$ as follows: given an unoccupied vertex v_i of degree 2 with $v_{i-1}v_i \in E$, let v_{i+1} be chosen so that we have $N(v_i) = \{v_{i-1}, v_{i+1}\}$. Let m denote the index for which the process stops, i.e. m is the least positive integer for which v_m is touched or $d(v_m) \geq 3$. Since v_1 is an unoccupied vertex of degree 2, it follows that $m \geq 2$.

We now split the proof into main cases which mostly depend on the value of m. For convenience, we say that a vertex v is *initially touched* if v is touched in the initial set-up. That is, if v is unoccupied on the initial board but v is an endpoint of one of the edges claimed by Toucher once the game has started, we do not consider v to be initially touched vertex (and we say that v is *initially untouched*).

Case 5.1. $m = 2$.

Note that v_2 cannot be touched by Case 1. Hence by the choice of m we must have $d(v_2) \geq 3$. Suppose that Isolator claims the edge v_1v_2 on her first move. If Toucher claims the edge v_0v_1 on her first move, we stop. Otherwise, Isolator claims the edge v_0v_1 on her second move, and we stop after Toucher's second move.

First consider the case when Isolator managed to claim both of these edges, and consider T_1 obtained by deleting the vertices v_0 and v_1. We set $C_1 = C' \cup \{f_1, f_2\}$, and recall that f_1 and f_2 are the edges claimed by Toucher on her two moves. It is easy to see that T_1 is a reduction of T, and we have $d(n) = 2$, $d(l) = 1$ and $d(k) = -2$. Note that v_2 is the only vertex whose degree changes during the process, and v_2 is initially unoccupied. Hence we have $d(s) \leq 0$, as the additional two moves given to Toucher can only decrease the value of $d(s)$. Since $|D| = 2$, it follows that $D(T, T_1) = 5 \leq 5(|D| - 1)$.

Now suppose that Toucher claimed the edge v_0v_1. Again, consider T_1 obtained by deleting the vertices v_0 and v_1, but in this case we take $C_1 = C'$. It is easy to see that T_1 is a reduction of T, and similarly we have $d(n) = 2$, $d(l) = 1$, $d(k) = 0$ and $d(s) = 0$. Indeed, in this case we have $d(s) = 0$, as Toucher claimed the edge v_0v_1 on her move. Since $|D| = 1$, it follows that $D(T, T_1) = -1 \leq 5(|D| - 1)$.

Case 5.2. $m = 3$ and v_3 is initially touched.

Since T is a tree with least 6 vertices, it follows that v_3 cannot be a leaf, and hence v_3 is occupied. Suppose that Isolator claims the edge v_1v_2 on her first move and one of the edges in $\{v_0v_1, v_2v_3\}$ on her second move. If Toucher has claimed the other one of these edges on one of her first two moves, the process stops after Toucher's second move. Otherwise, Isolator claims the other edge in $\{v_0v_1, v_2v_3\}$, and the process stops after Toucher's third move. The rest of our analysis splits into cases based on the number of neighbours of v_3.

Case 5.2.1. $d(v_3) = 2$.

Let v_4 be chosen so that $d(v_3) = \{v_2, v_4\}$. Since v_3 is occupied, it follows that $v_3v_4 \in C$. We also need to split the proof into cases based on the number of neighbours of v_4, and note that v_4 cannot be a leaf as T contains at least 6 vertices.

Case 5.2.1.1. $d(v_4) = 2$.

Let T_1 be the tree obtained by deleting the vertices v_0, v_1, v_2 and v_3, and let C_1 be the set of those edges in C that are not deleted during the process. Since v_4 is a leaf in T_1 and occupied in T, it follows that T_1 is a reduction of T.

First suppose that Isolator claimed all three edges in $\{v_0v_1, v_1v_2, v_2v_3\}$. Since v_4 is a leaf in T_1 and the edge $v_3v_4 \in C$ is deleted during the process, it is easy to check that we have $d(n) = 4$, $d(l) = 0$, $d(k) = -2$ and $d(s) \leq 0$, as the new edges claimed by Toucher cannot increase the value of $d(s)$. Hence it follows that $D(T, T_1) \leq 10 = 5(|D| - 1)$.

Now suppose that Isolator claimed only two such edges. Hence one of the edges in $\{v_0v_1, v_1v_2, v_2v_3\}$ must be claimed by Toucher, and this edge is deleted together with the edge $v_3v_4 \in C$. Hence it is easy to check that we have $d(n) = 4$, $d(l) = 0$, $d(k) = 0$ and $d(s) \leq 0$, and thus it follows that $D(T, T_1) \leq 4 < 5(|D| - 1)$.

Case 5.2.1.2. $d(v_4) \geq 3$.

Let $N(v_4) \setminus \{v_3\} = \{u_1, \ldots, u_a\}$ where $a \geq 2$. Let T_1 be the tree obtained by deleting the vertices v_0, v_1 and v_2, and by replacing the edge v_4u_1 with v_3u_1. Let C_1 be the set of all edges in C that are also edges in T_1, and if $v_4u_1 \in C$ the edge v_3u_1 is also added to C_1. Hence T_1 is a reduction of T by taking $f(v_4u_1) = v_3u_1$. Note that the only vertices whose degrees are affected during the process are v_3 and v_4, and it is easy to check that we have $d_s(v_3) = 0$ and $d_s(v_4) = (a + 1 - 2) - (a - 2) = 1$. In particular, it follows that $d(s) \leq 1$, and we also have $d(n) = 3$ and $d(l) = 1$.

If Isolator claimed all three edges, it follows that $d(k) = -3$. Hence we have $D(T, T_1) \leq 10 = 5(|D| - 1)$. If Isolator claimed only two such edges, it follows that $d(k) = -1$. Hence we have $D(T, T_1) \leq 4 < 5(|D| - 1)$.

Case 5.2.2. $d(v_3) \geq 3$.

Let T_1 be the tree obtained by deleting the vertices v_0, v_1 and v_2, and note that T_1 is a reduction of T. Since v_3 is the only vertex whose degree is affected during the process and $d_s(v_3) = 1$, it follows that $d(s) \leq 1$. We also have $d(n) = 3$.

If Isolator claimed all three edges, it follows that $d(l) = 1$ and $d(k) = -3$. Hence we have $D(T, T_1) = 10 \leq 5(|D| - 1)$. If Isolator claimed only two such edges, it follows that $d(l) = 1$ and $d(k) = -1$. Hence we have $D(T, T_1) \leq 4 < 5(|D| - 1)$.

Case 5.3. $m = 3$ and v_3 is initially unoccupied.

Since v_3 is initially unoccupied and $m = 3$, it follows that $d(v_3) \geq 3$. Again, suppose that Isolator claims the edge v_1v_2 on her first move and one of the edges in $\{v_0v_1, v_2v_3\}$ on her second move. If Toucher has occupied the other one of these edges on her first two moves, the process stops after the second move of Toucher. Otherwise, Isolator claims the other one of these edges on her third move, and the process tops after the third move of Toucher.

First suppose that Isolator claimed all three edges, and let T_1 be the tree obtained by deleting the vertices v_0, v_1 and v_2, and by taking $C_1 = C$. Then T_1 is a reduction of T, and the only vertex whose degree is affected during the process is v_3. Since $v_3 \notin O'$ it follows that $d(s) \leq 0$, and it is also easy to check that we have $d(n) = 3$, $d(l) = 1$ and $d(k) = -3$. Hence it follows that $D(T, T_1) \leq 9 < 5(|D| - 1)$.

If Isolator claimed only the edges v_1v_2 and v_2v_3, consider the same reduction T_1 as in the previous case, and again we take C_1 to be the set of those edges in C that are also edges in T_1. Again, it is easy to see that T_1 is indeed a reduction of T, as the edge v_2v_3 is occupied by Isolator. It is also easy to check that we have $d(n) = 3$, $d(l) = 1$, $d(k) = -1$ and $d(s) \leq 0$, and hence it follows that $D(T, T_1) \leq 3 < 5(|D| - 1)$.

Finally, suppose that Isolator claimed only the edges v_0v_1 and v_1v_2. Let $N(v_3) \setminus \{v_2\} = \{u_1, \ldots, u_a\}$ where $a \geq 2$. Consider T_1 obtained by deleting the vertices v_0 and v_1, and replacing the edge v_3u_1 with v_2u_1, and by taking C_1 to be the set containing all the edges in C that are also in T_1, and if $v_3u_1 \in C$ then v_2u_1 is also added to C_1. Since both v_2 and v_3 are touched, it follows that T_1 is a reduction of T by taking $f_E(v_3u_1) = v_2u_1$.

Note that v_2 and v_3 are the only vertices whose degrees are affected during the process. Since both of them are initially unoccupied, it follows that $d(s) \leq 0$. It is easy to check that we have $d(n) = 2$, $d(l) = 1$ and $d(k) = -2$. Hence we have $D(T, T_1) \leq 5 = 5(|D| - 1)$.

Case 5.4. $m \geq 4$.

Suppose that Isolator claims the edge v_2v_3 on her first move. Suppose that before a given move of Isolator the set of the edges claimed by Isolator is of the form $\{v_iv_{i+1}, v_{i+1}v_{i+2}, \ldots, v_jv_{j+1}\}$ for some $i \leq 2$ and $2 \leq j \leq m-1$. If $j < m-1$ and if the edge $v_{j+1}v_{j+2}$ is still available, Isolator claims this edge on her next move. Otherwise, if $i \geq 1$ and the edge $v_{i-1}v_i$ is still available, Isolator claims this edge on her next move. If neither of these edges is available, the process stops.

Let $D = \{v_i v_{i+1}, \dots, v_j v_{j+1}\}$ be the set of edges claimed by Isolator at the end of the process. In particular, we have $|D| = j - i + 1$. Hence the number of isolated vertices is $j - i$ as the set of isolated vertices is $\{v_{i+1}, \dots, v_j\}$. Note that we always have $i \leq 2$, $2 \leq j \leq m - 1$ and $j - i \geq 1$, as Toucher cannot claim both of the edges $v_1 v_2$ and $v_3 v_4$ on her first move. We now split the proof into multiple cases, mostly based on the value of j but sometimes also based on whether v_m is touched or $d(v_m) \geq 3$.

Case 5.4.1. $j < m - 2$.

Since $j \neq m - 1$ at the end of the process, it follows that Toucher has claimed the edge $v_{j+1} v_{j+2}$ on one of her moves. Let T_1 be the tree obtained by deleting the vertices $v_0, \dots, v_{j+1}$, and by taking C_1 to be the set of those edges in C that are also edges in T_1. Since $j < m - 2$, it follows that v_{j+2} is a leaf in T_1. Hence T_1 is a reduction of T, and it is easy to see that we have $d(n) = j + 2$, $d(l) = 0$ and $d(s) \leq 0$.

First suppose that we have $i \in \{1, 2\}$. Hence Toucher has claimed at least one of the edges in $\{v_0 v_1, v_1 v_2\}$, and also note that the edge $v_{j+1} v_{j+2}$ claimed by Toucher is deleted. Since Toucher has claimed exactly $j - i + 1$ edges outside C' on her new moves, it follows that $d(k) \geq (i - j - 1) + 2 = 1 + i - j$. Hence we have $D(T, T_1) \leq 4j - 3i - 1$, and by using the facts that $i \leq 2$ and $j - i \geq 1$ we obtain that $5(|D| - 1) - D(T, T_1) \geq j - 2i + 1 \geq 0$.

Now suppose that we have $i = 0$. In this case it follows that $d(k) = -j$, as $v_{j+1} v_{j+2}$ is the only deleted edge claimed by Toucher. Hence we have $D(T, T_1) \leq 4j + 2$, and since $j \geq 2$, it follows that $D(T, T_1) \leq 5j = 5(|D| - 1)$.

Case 5.4.2. $j = m - 2$ and $d(v_m) = 2$.

Since $d(v_m) = 2$, it follows that v_m is initially touched by the definition of m. Again, since $j \neq m - 1$, at the end of the process, Toucher has claimed the edge $v_{m-1} v_m$ on one of her moves. Let T_1 be the tree obtained by deleting the vertices $v_0, \dots, v_{m-1}$, and by taking C_1 to be the set of those edges in C that are also edges in T_1. Since $d(v_m) = 2$, it follows that v_m is a leaf in T_1. Hence T_1 is a reduction of T, and it is easy to check that we have $d(n) = m$, $d(l) = 0$ and $d(s) \leq 0$.

If $i \in \{1, 2\}$, it follows that $d(k) \geq 1 + i - j = 3 + i - m$ by using the same argument as in the proof of Case 5.4.1. Hence we have $D(T, T_1) \leq 4m - 3i - 9$, and thus it follows that $5(|D| - 1) - D(T, T_1) \geq m - 2i - 1$. Since $j \geq i + 1$ and $i \leq 2$, it follows that $m = j + 2 \geq i + 3 \geq 2i + 1$. If $i = 0$, it follows that $d(k) = -j = 2 - m$. Hence we have $D(T, T_1) \leq 4m - 6$, and since $m = j + 2 \geq 4$, it follows that $D(T, T_1) \leq 5(m - 2) = 5(|D| - 1)$.

Case 5.4.3. $j = m - 2$ and $d(v_m) \geq 3$.

Since $j \neq m - 1$, Toucher has claimed the edge $v_{m-1} v_m$ on one of her moves. Let $N(v_m) \setminus \{v_{m-1}\} = \{u_1, \dots, u_a\}$ where $a \geq 2$. Consider T_1 obtained by removing the vertices $v_0, \dots, v_{m-2}$, and by replacing the edge $v_m u_1$ with $v_{m-1} u_1$. Let C_1 be the set of those edges in C that are also edges in T_1, and if $v_m u_1 \in C$ then $v_{m-1} u_1$ is also added to C_1. It is easy to see that T_1 is a reduction of T by taking $f(v_m u_1) = v_{m-1} u_1$, as both v_{m-1} and v_m are touched.

If v_m is initially unoccupied, it is clear that we have $d(s) \leq 0$. Otherwise, we have $d_s(v_m) = (a + 1 - 2) - (a - 2) = 1$ and $d_s(v_{m-1}) = 0$. Hence it follows that $d(s) \leq 1$ in either case. We also have $d(n) = m - 1$ and $d(l) = 1$.

If $i \in \{1,2\}$, Toucher has claimed at least one of the edges v_0v_1 or v_1v_2, and hence we have $d(k) \geq i - j = i + 2 - m$. Hence it follows that $D(T, T_1) \leq 4m - 3i - 9$, and hence we have

$$5(|D| - 1) - D(T, T_1) \geq 5(m - 2 - i) - (4m - 3i - 9) = m - 2i - 1.$$

By using $m = j + 2 \geq i + 3$ and $i \leq 2$, it follows that $D(T, T_1) \leq 5(|D| - 1)$.

If $i = 0$, we have $d(k) = -j - 1 = 1 - m$. Hence it follows that $D(T, T_1) \leq 4m - 6$. Since $m \geq 4$, we have $D(T, T_1) \leq 5(m - 2) = 5(|D| - 1)$.

Case 5.4.4. $j = m - 1$ and $d(v_m) \geq 3$.

Let T_1 be the tree obtained by deleting the vertices $v_0, \dots, v_{m-1}$. Since Isolator has occupied all the edges $v_iv_{i+1}, \dots, v_{m-1}v_m$, it follows that T_1 is a reduction of T regardless of whether v_m is touched or not. Since v_m is the only vertex whose degree is affected during the process and $d_{T_1}(v_m) = d_T(v_m) - 1$, it follows that $d(s) \leq 1$. We also have $d(n) = m$ and $d(l) = 1$.

If $i \in \{1,2\}$, it follows that $d(k) \geq i + 1 - m$. Hence we have $D(T, T_1) \leq 4m - 3i - 5$, and thus it follows that

$$5(|D| - 1) - D(T, T_1) \geq 5(m - i - 1) - (4m - 3i - 5) = m - 2i.$$

Combining this with $m \geq 4$ and $i \leq 2$, we obtain that $D(T, T_1) \leq 5(|D| - 1)$.

If $i = 0$, it follows that $d(k) = -m$. Hence we have $D(T, T_1) \leq 4m - 2$, and since $m \geq 4$, it follows that $D(T, T_1) \leq 5(m - 1) \leq 5(|D| - 1)$.

Case 5.4.5. $j = m - 1$ and $d(v_m) \leq 2$.

Since $d(v_m) \leq 2$, the definition of m implies that v_m is touched. Since T is not a path, we must have $d(v_m) = 2$. Let v_{m+1} be chosen so that $N(v_m) = \{v_{m-1}, v_{m+1}\}$. Since v_m is touched and $v_{m-1}v_m \notin C$, it follows that $v_mv_{m+1} \in C$. We split the proof into subcases based on the degree of v_{m+1}. Note that v_{m+1} cannot be a leaf since T is not a path.

Case 5.4.5.1. $d(v_{m+1}) = 2$.

Let T_1 be the tree obtained by deleting the vertices $v_0, \dots, v_m$. Since v_{m+1} is touched in T and a leaf in T_1, it follows that T_1 is a reduction of T. We clearly have $d(n) = m + 1$, $d(l) = 0$ and $d(s) = 0$.

If $i \in \{1,2\}$, it follows that $d(k) \geq -(j - i + 1) + 2 = i - m + 2$, as at least two edges claimed by Toucher are deleted during the process, namely v_mv_{m+1} and one of v_0v_1 or v_1v_2. Hence we have $D(T, T_1) \leq 4m - 3i - 5$, and thus it follows that

$$5(|D| - 1) - D(T, T_1) \geq 5(m - i - 1) - (4m - 3i - 5) = m - 2i.$$

Again, by using $m \geq 4$ and $i \leq 2$, we obtain that $D(T, T_1) \leq 5(|D| - 1)$.

If $i = 0$, it follows that $d(k) \geq -(j + 1) + 1 = 1 - m$, and thus we have $D(T, T_1) \leq 4m - 2$. Since $m \geq 4$, it follows that $D(T, T_1) \leq 5(m - 1) = 5(|D| - 1)$.

Case 5.4.5.2. $d(v_{m+1}) \geq 3$.

Let $N(v_{m+1}) \setminus \{v_m\} = \{u_1, \dots, u_a\}$ where $a \geq 2$, and let T_1 be the tree obtained by deleting the vertices $v_0, \dots, v_{m-1}$ and by replacing the edge $v_{m+1}u_1$ with v_mu_1. Let C_1 be the set of those edges in C that are also edges in T_1, and if $v_{m+1}u_1 \in C$ then $v_m u_1$ is also added to C_1. Then T_1 is a reduction of T by taking $f_E(v_{m+1}u_1) = v_m u_1$. It is easy to see that we have $d(n) = m$ and $d(l) = 0$. Note that v_m and v_{m+1} are the only vertices whose degrees are affected during the process. Since $d_s(v_{m+1}) \leq 1$ and $d_s(v_m) = 0$, it follows that $d(s) \leq 1$.

If $i \in \{1, 2\}$, it follows that $d(k) \geq -(j - i + 1) + 1 = i + 1 - m$. Hence we have $D(T, T_1) \leq 4m - 3i - 5$. Since $m \geq 4 \geq 2i$, it follows that

$$5(|D| - 1) - D(T, T_1) \geq 5(m - i - 1) - (4m - 3i - 5) = m - 2i \geq 0.$$

If $i = 0$, it follows that $d(k) = -m$, and hence we have $D(T, T_1) \leq 4m - 2$. Since $m \geq 4$, it follows that $D(T, T_1) \leq 5(m - 1) \leq 5(|D| - 1)$, which completes the proof of Case 5.

Suppose that T does not contain any configurations described in Cases 1-5, and let $v_0, \ldots, v_m$ be a maximal path of vertices in T for which v_i is an unoccupied vertex of degree 2 for all $1 \leq i \leq m - 1$, and for which we have $v_i \in N(v_{i-1})$ for all $1 \leq i \leq m$. Since T does not contain any configurations described in Cases 1-5, it follows that v_0 and v_m are also unoccupied, and the maximality assumption implies that we must have $d(v_0) \geq 3$ and $d(v_m) \geq 3$. In our final case we suppose that there exists such a path with $m \geq 3$.

Case 6. There exist $m \geq 3$ and unoccupied vertices $v_0, \ldots, v_m$ satisfying $v_i \in N(v_{i-1})$ for all $1 \leq i \leq m$, $d(v_i) = 2$ for all $1 \leq i \leq m - 1$, $d(v_0) \geq 3$ and $d(v_m) \geq 3$.

Suppose that Isolator claims the edge $v_1 v_2$ on her first move. Suppose that before a given move of Isolator, the set of edges claimed by Isolator is of the form $\{v_i v_{i+1}, v_{i+1} v_{i+2}, \ldots, v_j v_{j+1}\}$ for some $i \in \{0, 1\}$ and $j \leq m - 1$. If $j < m - 1$ and if the edge $v_{j+1} v_{j+2}$ is still available, Isolator claims this edge on her move. Otherwise, if $i = 1$ and the edge $v_0 v_1$ is still available, Isolator claims this edge on her move. If neither of these conditions is satisfied, the process stops.

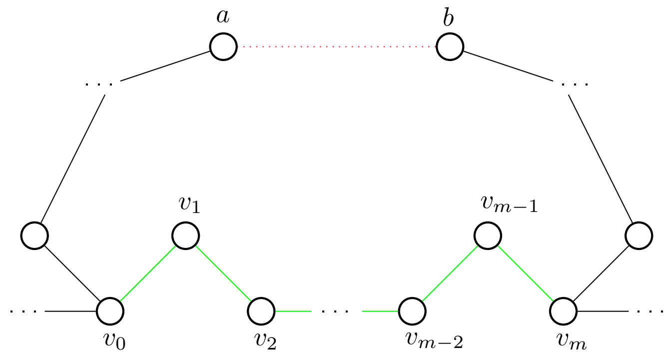
Let $\{v_i v_{i+1}, \ldots, v_j v_{j+1}\}$ be the set of edges claimed by Isolator at the end of such process. Note that we have $|D| = j - i + 1$, $i \in \{0, 1\}$, $1 \leq j \leq m - 1$ and $j - i \geq 1$ as Toucher cannot claim both of the edges $v_0 v_1$ and $v_2 v_3$ on her first move. We split the proof into several cases based on the values of i and j.

Case 6.1. $i = 0$ and $j = m - 1$.

Let S be the graph obtained by deleting the vertices $v_1, \ldots, v_{m-1}$. It is easy to see that S consists of two connected components, both of which are trees. Let a and b be leaves chosen from distinct connected components and let T_1 be the tree obtained by adding the edge ab to the graph S as demonstrated in Figure 5.4, and set $C_1 = C \cup \{ab\}$. Note that the set of leaves in T_1 is exactly $L \setminus \{a, b\}$ since $d_{T_1}(v_0) \geq 3 - 1 = 2$ and $d_{T_1}(v_m) \geq 3 - 1 = 2$. Hence we have $d(l) = 2$. Since $ab \in C_1$, it follows that T_1 is a reduction of T.

Since v_0 and v_m are initially unoccupied, it is easy to see that we have $d(s) \leq 0$. Note that $d(k) = -m - 1$, as Toucher has claimed m new edges, and the edge ab is assigned to Toucher. Finally, it is clear that we have $d(n) = m - 1$. Since $m \geq 3$, it follows that $D(T, T_1) \leq 4m - 4 \leq 5(m - 1) = 5(|D| - 1)$.

Figure 5.4: Construction of T_1. Green edges are edges claimed by Isolator. These edges are deleted during the process, and the red dotted edge is assigned to Toucher.



Case 6.2. $i = 0$ and $j < m - 1$.

Since $j < m-1$, it follows that Toucher has claimed the edge $v_{j+1}v_{j+2}$ on one of her moves. Let S be the graph obtained by deleting the vertices $v_1, \dots, v_{j+1}$, and let a be a leaf in the component of S containing v_0. Consider the tree T_1 obtained by adding the edge av_{j+2} to S, and define C_1 by setting $C_1 = (C \cup \{av_{j+2}\}) \setminus \{v_{j+1}v_{j+2}\}$. Note that T_1 is a reduction of T as both a and v_{j+2} are touched vertices in both T and T_1. Finally, note that we have $d_{T_1}(v_0) \geq 3 - 1 = 2$, and hence we have $d(l) = 1$.

Note that the only vertices whose degrees are affected during the process are a, v_0 and v_{j+2}. Note that $d_s(a) = 0$, and since $v_0, v_{j+2} \notin O'$ it follows that $d(s) \leq 0$. It is easy to check that we also have $d(n) = j + 1$ and $d(k) = -(j+1)$. Thus we have $D(T, T_1) \leq 4j + 1$, and since $j \geq 1$ it follows that $D(T, T_1) \leq 5j = 5(|D| - 1)$.

Case 6.3. $i = 1$ and $j = m - 1$.

Note that the case $(i, j) = (1, m-1)$ is equivalent to the case $(i, j) = (0, m-2)$, which is covered by Case 6.2.

Case 6.4. $i = 1$ and $j < m - 1$.

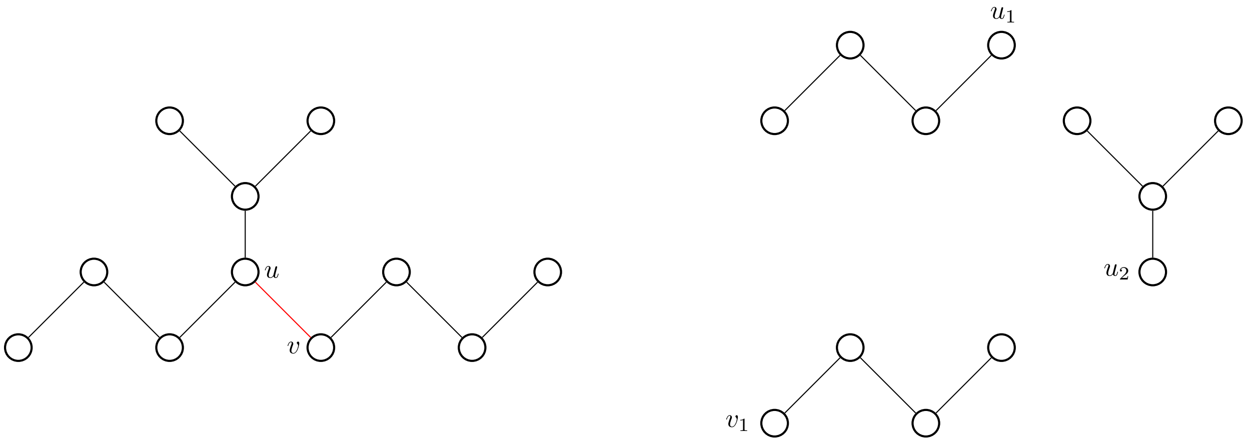
Since $i = 1$ and $j < m - 1$, it follows that Toucher has claimed both of the edges v_0v_1 and $v_{j+1}v_{j+2}$. Let T_1 be the tree obtained by deleting the vertices $v_1, \dots, v_{j+1}$, and by adding the edge v_0v_{j+2}, and set $C_1 = (C \cup \{v_0v_{j+2}\}) \setminus \{v_0v_1, v_{j+1}v_{j+2}\}$. Note that T_1 is a reduction of T as both v_0 and v_{j+2} are touched before and after the reduction.

Note that the degree of any vertex that is not deleted is not affected during the process. Since $v_0 \in O' \setminus O_1$ and $d_{T_1}(v_0) = d_T(v_0) \geq 3$, it follows that $d_s(v_0) \leq -1$. Hence we must have $d(s) \leq -1$. Since the edges v_0v_1 and $v_{j+1}v_{j+2}$ claimed by Toucher are deleted during the process and the edge v_0v_{j+2} is given to Toucher, it follows that $d(k) = 1 - j$, and it is easy to see that we have $d(n) = j+1$ and $d(l) = 0$. Hence it follows that $D(T, T_1) \leq 4j - 3$. Since $|D| - 1 = j - 1$ and $j \geq i+1 \geq 2$, it follows that we have $D(T, T_1) \leq 5(|D| - 1)$, which completes the proof of Case 6.

Note that introducing the term $\sum_{v \in O}(d(v) - 2)$ in order to strengthen the inductive hypothesis was crucial in the proof of Case 6.4.

Let $C = \{e_1, \dots, e_k\}$ be the set of edges claimed by Toucher at the start of the game. Our aim is to prove that if T together with this particular collection C does not contain any of the

Figure 5.5: Illustration of one stage of the process. The red edge $uv \in C_i$ is deleted, and since $d_i = 3$ the tree splits into 3 components.



configurations described in Cases 1-6, we must have $S(T) \leq 0$. For each $1 \leq i \leq k$ let $e_i = a_i b_i$, and let $d_i = d_T(a_i) + d_T(b_i) - 2$.

We say that a graph T is a *forest* if every connected component of T is a tree. We define a sequence of forests $T_0, \ldots, T_k$ and collections of edges $C_0, \ldots, C_k$ as follows. First of all, we set $T_0 = T$ and $C_0 = C = \{e_1, \ldots, e_k\}$, and at every stage we have $C_i = \{e_{i+1}, \ldots, e_k\}$.

Given T_i and C_i, let X be the connected component of T_i containing the edge e_i, and note that X is a tree since T_i is a forest. Let Y be the forest consisting of d_i trees obtained by removing the vertices a_i and b_i and the edge $a_i b_i$, and by adding one new vertex to each connected component S_j of Y joined by an edge to the vertex of S_j that is a neighbour of a_i or b_i. Note that such a vertex always exists in each connected component, and such vertex is also unique since X is a tree. Finally, we set T_{i+1} to be the union of Y and all the components of T_i apart from X. One stage of the process is illustrated in Figure 5.5.

Note that by Claims 2, 3 and 4 it follows that all $a_1, \ldots, a_k, b_1, \ldots, b_k$ are distinct vertices, none of them is a leaf in T, and there are no other pairs of neighbours among these vertices apart from the pairs $a_i b_i$ (before such an edge is deleted). By Claim 1 it follows that every connected component of T_k contains at least 4 vertices, and also none of the a_i or b_i is a leaf in any T_j (before they are deleted).

Note that during the i^{th} step of the process, the number of connected components increases by $d_i - 1$, as one connected component splits into d_i connected components. Hence the number of connected components in T_k is

$$D = 1 + \sum_{i=1}^{k} (d_i - 1) = 1 - k + \sum_{i=1}^{k} d_i. \tag{5.9}$$

Let $n_1, \ldots, n_D$ be the number of vertices in each connected component and let $l_1, \ldots, l_D$ be the number of leaves in each connected component. Note that during the i^{th} stage of the process, the number of vertices increases by $d_i - 2$, as we delete the vertices a_i and b_i, and add d_i new vertices that are leaves. Hence we have

$$\sum_{i=1}^{D} n_i = n + \sum_{i=1}^{k} (d_i - 2) = n - 2k + \sum_{i=1}^{k} d_i. \tag{5.10}$$

Since none of the vertices $a_1, \ldots, a_k, b_1, \ldots, b_k$ is a leaf at any stage of the process before they are deleted, it follows that the number of leaves increases by d_i on the i^{th} stage. Hence we have

$$\sum_{i=1}^{D} l_i = l + \sum_{i=1}^{k} d_i. \tag{5.11}$$

Let S be a connected component in T_k. Note that if S contains a vertex of degree 2 whose neighbour is a leaf, we can backtrack the process and find a vertex of degree 2 in T whose neighbour is a touched vertex, which contradicts Case 5. Hence we may assume that no vertex of degree 2 in S has a leaf as a neighbour.

If S contains two vertices of degree 2 that are neighbours, it follows that there exists a path of vertices $v_0, \ldots, v_t$ in S for some $t \geq 3$ with $d(v_0) \geq 3$, $d(v_{t+1}) \geq 3$ and $d(v_i) = 2$ for all $1 \leq i \leq t$. Since none of these vertices is a leaf in S, it follows that these vertices also formed a path satisfying the same condition in T, and all of these vertices were unoccupied in T. This contradicts Case 6.

Hence in every connected component there is no vertex of degree 2 whose neighbour is a leaf or another vertex of degree 2. Since each connected component is a tree with at least 4 vertices, Lemma 44 implies that we have $3l_i \geq n_i + 5$ for all i. Adding these inequalities for all $i \in \{1, \ldots, D\}$, and by using (5.9), (5.10) and (5.11) we obtain that

$$3\left(l + \sum_{i=1}^{k} d_i\right) \geq n - 2k + \sum_{i=1}^{k} d_i + 5 - 5k + 5\sum_{i=1}^{k} d_i.$$

This can be rearranged to

$$3l + 3k \geq n + 5 + 3\sum_{i=1}^{k} d_i - 4k. \tag{5.12}$$

Note that $O(T) = \{a_1, \ldots, a_k, b_1, \ldots, b_k\}$, and hence it follows that

$$\sum_{v \in O(T)} (d(v) - 2) = \sum_{i=1}^{k} (d(a_i) + d(b_i) - 4) = \sum_{i=1}^{k} (d_i - 2) = -2k + \sum_{i=1}^{k} d_i.$$

Hence (5.12) can be written as

$$3l + 3k \geq n + 5 + \sum_{v \in O(T)} (d(v) - 2) + 2\sum_{i=1}^{k} d_i - 2k. \tag{5.13}$$

Since none of a_i or b_i is a leaf, it follows that $d_i = d(a_i) + d(b_i) - 2 \geq 2$. Hence we have

$2\sum_{i=1}^{k} d_i - 2k \geq 2k \geq 0$. In particular, (5.13) implies that we have

$$n + 7 - 3k - 3l + \sum_{v \in O(T)} (d(v) - 2) \leq 2,$$

and thus we must have $S(T) \leq \lfloor \frac{2}{5} \rfloor = 0$. Hence the claim follows trivially, as we always have $\alpha(T, C, L) \geq 0$. Since we always have $\sum_{v \in O(T)} (d(v) - 2) \geq 0$, the second part of the Lemma follows immediately. □

We are now ready to prove Theorem 37.

Proof of Theorem 37. Let T be a tree with n vertices. Suppose that during the first phase of the game Isolator follows the strategy specified in Lemma 42, and let r be the number of edges claimed by her during the first phase of the game. Let T', C' and $X' = L'$ be given by Lemma 42. Since $|I| = r$, it follows that $\alpha(T) \geq r + \alpha(T', C', L')$. Since the second phase is equivalent to the delayed game $F(T', C', L')$, Lemma 43 implies that we have

$$\alpha\left(T', C', L'\right) \geq \left\lfloor \frac{|T'| - 3|C'| - 3|L'| + 7}{5} \right\rfloor.$$

Since Lemma 42 guarantees that we have

$$\left|T'\right| - 3\left|C'\right| - 3\left|L'\right| \geq n - 5r - 4,$$

it follows that

$$\alpha(T) \geq r + \left\lfloor \frac{n - 5r - 4 + 7}{5} \right\rfloor = \left\lfloor \frac{n+3}{5} \right\rfloor,$$

which completes the proof of Theorem 37. □

There are many questions that are open concerning the value of $u(G)$ for general G. Dowden, Kang, Mikalački and Stojaković [17] gave bounds for $u(G)$ that depended on the degree sequence of the graph G. In particular, they concluded that if the minimum degree of G is at least 4, we have $u(G) = 0$. They also proved that there exists a 3-regular graph satisfying $u(G) > 0$, and they proved that for all 3-regular graphs we have $u(G) \leq \frac{n}{8}$. It would be interesting to know what the largest possible proportion of isolated vertices in a connected 3-regular graph is.

Chapter 6

Intervals in the Hales-Jewett theorem

6.1 Introduction

In order to state the Hales-Jewett theorem we need some notation. Given positive integers k and n, note that the cube $[k]^n$ can be viewed as the set of all words in symbols $\{0, \ldots, k-1\}$ of length n. A set $L \subset [k]^n$ is called a *combinatorial line* if there exist a non-empty set $S \subseteq \{1, \ldots, n\}$ and integers $a_i \in [k]$ for all $i \notin S$ such that

$$L = \{(x_1, \ldots, x_n) : x_i = a_i \text{ for all } i \notin S \text{ and } x_i = x_j \text{ for all } i, j \in S\}.$$

The set S is called the *active coordinate set* of L.

Theorem 45. *(Hales-Jewett, [22]). For any k and r there exists N so that whenever $[k]^n$ is r-coloured for $n \geq N$, there exists a monochromatic combinatorial line.*

As noted by Conlon and Kamčev [15], by following Shelah's proof of the Hales-Jewett theorem [42] it can be shown that for sufficiently large n one can always find a monochromatic combinatorial line whose active coordinate set S is a union of at most $HJ(k-1, r)$ intervals, where $HJ(k-1, r)$ is the smallest integer n for which the Hales-Jewett theorem holds for $k-1$ and r.

In the case $k = 3$, since $HJ(2, r) = r$, this says that one can always find a monochromatic line whose active coordinate set is a union of at most r intervals. Conlon and Kamčev proved in [15] that this bound is tight for r odd: in other words, they showed that for each odd r there exists an r-colouring of $[3]^n$ for any n for which every monochromatic line has active coordinate set made up of at least r intervals. They conjectured that this would also be the case for r even. In particular, for $r = 2$, they conjectured that for all n there exists a 2-colouring of $[3]^n$ for which there exists no monochromatic combinatorial line whose active coordinate set is an interval.

In this chapter we prove that, perhaps surprisingly, their conjecture is false when $r = 2$. This can be stated in the following form.

Theorem 46. *For all sufficiently large n, whenever $[3]^n$ is 2-coloured there exists a monochromatic combinatorial line whose active coordinate set S is an interval.*

6.2 Proof of Theorem 46

The idea of the proof is as follows. By applying Ramsey's theorem, we will pass to a subspace on which the colour of a word depends only on its 'pattern' of intervals, and not on its 'breakpoints' which are the places where the word changes from one letter to another. Once this is done, we can consider some particular small patterns.

For a word w, define the *pattern* $\overline{w}$ to be the word obtained from w by contracting every interval on which w is constant to a single letter. We define the particular words $s_1 = 021$, $s_2 = 0121$, $s_3 = 0201$, $s_4 = 02121$ and $s_5 = 02021$. Set t_i to be the length of the word s_i. Define recursively $n_0 = 4$, and for $i \geq 1$ let $n_i = R^{(t_i-1)}(n_{i-1})$ where $R^{(t)}(s) = R^{(t)}(s,s)$ is the t-set Ramsey-number. Finally, set $N = n_5 + 1$.

Suppose that $n \geq N$ and let c be a 2-colouring of $[3]^n$. For a word w, define the set of *breakpoints* $T(w)$ by setting $T(w) = \{a_1, \ldots, a_m\}$ if $w_{a_{i-1}+1} = \cdots = w_{a_i}$ and $w_{a_i} \neq w_{a_i+1}$ for all $1 \leq i \leq m+1$, with the convention $a_0 = 0$, $a_{m+1} = n$. For example, $w = 0011222000$ has breakpoints $T(w) = \{2, 4, 7\}$.

Let s be a word of length t and let $T_1 = \{a_1, \ldots, a_{t-1}\} \subseteq \{1, \ldots, n-1\}$ be a set of size $t-1$. We say that $w \in [3]^n$ has *breakpoints in* T_1 *with pattern* s if $T(w) = T_1$ and $\overline{w} = s$. For example, $w = 0011222000$ has breakpoints in $T(w) = \{2, 4, 7\}$ with pattern $s = 0120$. Note that if $\overline{w} = s$, then there exists a unique set T_1 of size $|s| - 1$ for which w has breakpoints in T_1 with pattern s.

Set $Y_5 = \{1, \ldots, n-1\}$. Given a set Y_i satisfying $|Y_i| \geq n_i$ and a certain specific pattern p, we will recursively define a set Y_{i-1} satisfying $|Y_{i-1}| \geq n_{i-1}$ so that the words w which have breakpoints in Y with pattern p for some $Y \subseteq Y_{i-1}$ all have the same colour.

Recall that t_i is the length of the word s_i defined at the start of the theorem. For all $A \in \{1, \ldots, n-1\}^{(t_i-1)}$ define s_i^A to be the unique word which has breakpoints in A with pattern s_i. Let c_i denote the 2-colouring of the set $Y_i^{(t_i-1)}$ given by $c_i(A) = c\left(s_i^A\right)$. By Ramsey's theorem and the choice of n_i, there exists $Y_{i-1} \subseteq Y_i$ with $|Y_{i-1}| \geq n_{i-1}$ for which $Y_{i-1}^{(t_i-1)}$ is monochromatic, say with colour d_i. In particular, if w is a word with $\overline{w} = s_i$ and $T(w) \subseteq Y_{i-1}$, then we have $c(w) = d_i$. Thus we obtain sets $Y_0 \subseteq Y_1 \subseteq \cdots \subseteq Y_5$ with $|Y_0| \geq 4$ and colours $d_1 \ldots, d_5$ so that c_i restricted to $Y_i^{(t_{i+1}-1)}$ is a constant d_{i+1}.

Note that it is impossible to choose colours $d_1, \ldots, d_5$ without at least one of the following sets

$$
\begin{aligned}
N_1 &= \{d_1, d_2\} \\
N_2 &= \{d_1, d_3\} \\
N_3 &= \{d_2, d_4\} \\
N_4 &= \{d_3, d_5\} \\
N_5 &= \{d_1, d_4, d_5\}
\end{aligned}
$$

having just one element (i.e. all colours being equal). Indeed, if the first four sets contain both colours, we must have $d_2 = d_3$ and $d_1 = d_4 = d_5$, and the second condition implies that the last set contains only one colour.

Let $a_1 < a_2 < a_3 < a_4$ be elements in Y_0. We use the shorthand $w = [b_1 b_2 b_3 b_4 b_5]$ for the word which has $w_i = b_j$ for all $a_{j-1} < i \leq a_j$, with the convention $a_0 = 0$ and $a_5 = n$. Note that we allow $b_i = b_{i+1}$. Hence we have $T(w) \subseteq \{a_1, \ldots, a_4\} \subseteq X_0$ and $\overline{w} = \overline{b_1 b_2 b_3 b_4 b_5}$.

Set $w_1 = [02221]$, $w_2 = [01121]$, $w_3 = [02001]$, $w_4 = [02121]$ and $w_5 = [02021]$. It is easy to verify that for all i we have $\overline{w_i} = s_i$ and $T(w_i) \subseteq Y_0 \subseteq Y_{i-1}$. Furthermore, set $v_1 = [00021]$, $v_2 = [00121]$, $v_3 = [02011]$ and $u_1 = [02111]$. As before, it is easy to verify that we have $\overline{v_i} = s_i$, $\overline{u_1} = s_1$, and by construction we have $T(v_i) \subseteq Y_0 \subseteq Y_{i-1}$ and $T(u_1) \subseteq Y_0$. Thus by the choice of the sets Y_i, it follows that we have $c(w_i) = d_i$, $c(v_i) = d_i$ and $c(u_1) = d_1$.

It is straightforward to verify that

- v_1, w_2, w_1 forms a combinatorial line l_1 with $S_1 = \{a_1 + 1, \ldots, a_3\}$
- w_3, u_1, w_1 forms a combinatorial line l_2 with $S_2 = \{a_2 + 1, \ldots, a_4\}$
- v_2, w_2, w_4 forms a combinatorial line l_3 with $S_3 = \{a_1 + 1, \ldots, a_2\}$
- w_3, v_3, w_5 forms a combinatorial line l_4 with $S_4 = \{a_3 + 1, \ldots, a_4\}$
- w_5, w_4, w_1 forms a combinatorial line l_5 with $S_5 = \{a_2 + 1, \ldots, a_3\}$

It is clear that the colours used to colour the points of the line l_i are exactly the colours in the set N_i. As observed earlier, one of the sets N_i contains only one colour, which implies that the associated line l_i is monochromatic. Since the active coordinate set of each line l_i is an interval, this completes the proof. $\square$

Chapter 7

Induced saturation of P_6

7.1 Introduction

A graph G is said to be *H-saturated* if G does not contain a copy of H, but adding any edge from G^c to G creates a copy of H. It is clear that for any non-empty H there exists such G with a given number of vertices.

The notion of saturation can be generalised to induced subgraphs in the following way. A graph G is said to be *H-induced-saturated* if G does not contain an induced copy of H, but removing any edge from G creates an induced copy of H and adding any edge of G^c to G creates an induced copy of H.

It is not clear whether for a given H there exists a graph G which is H-induced-saturated. Martin and Smith [34] studied a similar problem from a quantitative perspective, and they proved that for $H = P_4$ there is no such G satisfying the property, where P_n denotes the path on n vertices. For convenience, we say that H is *induced saturated* if there exists some G which is H-induced-saturated. Behrens, Erbes, Santana, Yager and Yeager [5] proved that graphs belonging to a few simple families are induced saturated, and they also proved some quantitative results.

It is natural to ask what happens when $H = P_n$ for other values of n. The cases $H = P_2$ and $H = P_3$ are trivial, as one can take G to be an empty graph or a clique respectively. Axenovich and Csikós [2] gave examples of families of graphs that are induced saturated, and they also gave an easier proof of the fact that P_4 is not induced saturated. However, their examples of induced saturated families did not include P_n for any $n \geq 5$, and they asked whether the graphs $H = P_n$ are induced saturated for $n \geq 5$. The aim of this chapter is to provide an example which shows that P_6 is induced saturated.

7.2 The construction

Theorem 47. *There exists a graph with 16 vertices which is P_6-induced-saturated.*

Proof. Let $\mathbb{F} = \mathbb{F}_2(\alpha) / \left(\alpha^4 + \alpha + 1\right)$ be the finite field of order 16, and note that α is a generator of the multiplicative group $\mathbb{F}^\times$. Let S be the set of non-zero cubes in $\mathbb{F}$, i.e.

$$S = \left\{1,\, \alpha^3,\, \alpha^2 + \alpha^3,\, \alpha + \alpha^3,\, 1 + \alpha + \alpha^2 + \alpha^3\right\}.$$

Define a graph G whose vertex set is $\mathbb{F}$ and whose edges are given by $xy \in E(G)$ if and only $x - y \in S$. This graph is an example of a Cayley graph, and it is also known as the Clebsch graph. For later purposes, it will be convenient to use the algebraic way of defining the graph.

First of all, note that the map $x \to \alpha^{3i}x + \beta$ is an automorphism of G for any $i \in \{0, \dots, 4\}$ and $\beta \in \mathbb{F}$. Indeed, let $\theta(x) = \alpha^{3i}x + \beta$ for some i and β. Then for all $x, y \in \mathbb{F}$ we have $\theta(x) - \theta(y) = \alpha^{3i}(x - y)$. Thus $\theta(x) - \theta(y)$ is a non-zero cube in $\mathbb{F}$ if and only if $x - y$ is, and thus $\theta(x)\theta(y) \in E(G)$ if and only if $xy \in E(G)$.

Given any edge $xy \in E(G)$, let i be chosen so that $x - y = \alpha^{3i}$. Then θ defined by $\theta(z) = \alpha^{3i}z + y$ is an automorphism of G satisfying $\theta(0) = y$ and $\theta(1) = x$. In particular, the edge 01 is mapped to the edge xy under this automorphism, and hence it follows that for any two edges $e, f \in E(G)$ there exists an automorphism θ satisfying $\theta(e) = f$, i.e. the group of automorphisms acts transitively on the edges of G.

For convenience, we call the edges of G^c as non-edges. Define the particular non-edges $f_1 = 0\alpha^{10}$ and $f_2 = 0\alpha^{14}$. Given a non-edge $e = xy$, since $E(G)$ consists of pairs xy satisfying $x - y = \alpha^{3i}$ for some i, it follows that there exists i for which one of the conditions $x - y = \alpha^{3i+1}$ or $x - y = \alpha^{3i+2}$ is satisfied. In the first case, note that θ given by $\theta(z) = \alpha^{3i-9}z + y$ maps f_1 to e, and in the second case note that θ given by $\theta(z) = \alpha^{3i-12}z + y$ maps f_2 to e. By taking inverses, it follows that for an arbitrary non-edge e there exists an automorphism θ mapping e to either f_1 or f_2.

We will now check that G satisfies all the required properties.

Claim 1. G does not contain an induced copy of P_6.

Proof of Claim 1. Suppose that G contains an induced copy of P_6. Since the group of automorphisms acts transitively on $E(G)$, we may assume that 0 is one of the endpoints of the induced path, and 1 is the only neighbour of 0 on this induced path. Thus $T = N(0)^c \cap N(1)^c$ contains an induced path on three vertices.

It is easy to verify that

$$T = \left\{\alpha,\, 1+\alpha,\, \alpha^2,\, 1+\alpha^2,\, \alpha+\alpha^2,\, 1+\alpha+\alpha^2\right\}.$$

Hence $G[T]$ is a union of three disjoint edges corresponding to the pairs $\{\alpha, 1+\alpha\}$, $\left\{\alpha^2, 1+\alpha^2\right\}$ and $\left\{\alpha+\alpha^2,\, 1+\alpha+\alpha^2\right\}$. Thus $G[T]$ does not contain an induced path on three vertices, so G cannot contain an induced copy of P_6. □

Claim 2. Adding any non-edge to G creates an induced copy of P_6.

Proof of Claim 2. By the earlier observations, it suffices to consider the cases when the non-edge is $0\alpha^{10}$ or $0\alpha^{14}$. Consider the particular elements $x_1 = 0$, $x_2 = \alpha+\alpha^3$, $x_3 = \alpha$, $x_4 = \alpha+\alpha^2+\alpha^3$, $x_5 = \alpha^2$, $x_6 = \alpha^{10} = 1+\alpha+\alpha^2$ and $x_7 = \alpha^{14} = 1+\alpha^3$, and let $R = \{x_1, \dots, x_7\}$. It is easy to verify that $G[R]$ is a union of induced P_5 whose vertices are x_1, x_2, x_3, x_4, x_5 in this order, together with two isolated vertices x_6 and x_7. Hence adding either of the non-edges $0\alpha^{10}$ or $0\alpha^{14}$ creates an induced copy of P_6 in G. □

Claim 3. Removal of any edge from G creates an induced copy of P_6.

Proof of Claim 3. By our earlier observation we may assume that the edge removed is 01. It is easy to check that $v_1 = \alpha + \alpha^2 + \alpha^3$, $v_2 = 1$, $v_3 = 1 + \alpha^3$, $v_4 = \alpha^3$, $v_5 = 0$ and $v_6 = \alpha^2 + \alpha^3$ forms an induced copy of P_6 in this case. $\square$

From the claims above it follows that G is P_6-induced-saturated graph. $\square$

We now mention some other recent work on induced saturation of paths. Cho, Choi and Park [11] found the first infinite family of induced saturated paths. They generalised our construction of G and proved that P_{3n} is induced-saturated for every $n \geq 2$. In addition, they gave examples of some other graphs G that are P_n-induced-saturated for some specific small values of n. As an example, they noted that the Petersen graph is also P_6-induced-saturated.

Inspired by this observation, Dvořák [18] proved that for all $n \geq 6$ there exists a graph G_n that is P_n-induced saturated, where the graphs G_n are natural generalisations of the Petersen graph obtained as follows. For each n, the vertex set of G_n is $\{v_1, \dots, v_{n-1}, w_1, \dots, w_{n-1}\}$, with the edges $v_i w_i$, $v_i v_{i+1}$, $v_1 v_n$ and $w_j w_k$ for any i and for any j and k satisfying $j - k \not\equiv \pm 1 \pmod{n-1}$. Finally, the case $n = 5$ was settled by Spiegel [43], and independently by Bonamy, Groenland, Johnston, Morrison and Scott [10] who proved that P_5 is induced saturated. As a consequence of these results, it follows that P_n is induced saturated whenever $n \neq 4$.

Chapter 8

Bibliography

[1] G. Aubrun, S.J. Szarek. Alice and Bob Meet Banach - The Interface of Asymptotic Geometric Analysis and Quantum Information Theory. Mathematical Surveys and Monographs, American Math. Soc, 2017.

[2] M. Axenovich, M. Csikós. Induced Saturation of Graphs. *Discrete Math.* **342** (4): 1195-1212, 2019.

[3] J. Balogh, W. Samotij. On the Chvátal-Erdős triangle game. *Electron. J. Combin.* **18** (1): Paper 72, 2011.

[4] J. Beck. Combinatorial Games: Tic-Tac-Toe Theory. Cambridge University Press, 2008.

[5] S. Behrens, C. Erbes, M. Santana, D. Yager, and E. Yeager. Graphs with induced-saturation number zero. *Electron. J. Combin.* **23** (1): Paper 1.54, 2016.

[6] Bollobás. Combinatorics. Cambridge University Press, 1986.

[7] Bollobás, I. Leader. Exact Face-isoperimetric inequalities. *European J. Combin.* **11** (4): 335-340, 1990.

[8] Bollobás, I. Leader. Compressions and isoperimetric inequalities. *J. Combin. Theory Ser. A*, **56** (1): 47-62, 1991.

[9] Bollobás, I. Leader. Personal communication, 2017.

[10] M. Bonamy, C. Groenland, T. Johnston, N. Morrison, A. Scott. Induced Saturation for P_5. https://tomjohnston.co.uk/blog/2020-05-22-induced-saturation-for-paths.html, 2020.

[11] E. Cho, I. Choi, B. Park. On induced saturation for paths. *European J. Combin.* **91**, 2021.

[12] V. Chvátal, P. Erdős. Biased positional games. *Ann. Discrete Math.* **2**: 221–229, 1978.

[13] G. Clements. Another Generalization of the Kruskal-Katona Theorem. *J. Combin. Theory Ser. A* **68** (1): 239-245, 1994.

[14] G. Clements, B. Lindström. A generalization of a combinatorial theorem of Macaulay. *J. Combin. Theory* **7:** 230-238, 1969.

[15] D. Conlon and N. Kamčev. Intervals in the Hales-Jewett theorem. *arXiv preprint, arXiv:1801.08919*, 2017.

[16] T. Danh, D. Daykin. Ordering integer vectors for coordinate deletion. *J. Lond. Math. Soc.* **55** (3): 417-427, 1997.

[17] C. Dowden, M. Kang, M. Mikalački, M. Stojaković. The Toucher-Isolator game. *Electron. J. Combin.* **26** (4): Paper 4.6, 2019.

[18] V. Dvořák. P_n-induced saturated graphs exist for all $n \geq 6$. *Electron. J. Combin.* **27** (4): Paper 4.43, 2020.

[19] P. Frankl, Z. Füredi, G. Kalai. Shadows of colored complexes. *Math. Scand.* **63**: 169-178, 1988.

[20] Z. Füredi, J. R. Griggs. Families of finite sets with minimum shadows. *Combinatorica* **6** (4): 355-363, 1986.

[21] H. Gebauer. On the Clique-Game. *European J. Combin* **33** (1): 8–19, 2012.

[22] A. Hales, R. Jewett. Regularity and positional games. *Trans. Amer. Math. Soc.* **106**: 222–229, 1963.

[23] L. Harper. Optimal numberings and isoperimetric problems on graphs. *J. Combin. Theory* **1** (3): 385-393, 1966.

[24] G. Katona. The Hamming-sphere has minimum boundary. *Studia Sci. Math. Hungar.* **10**: 131-140, 1975.

[25] G. Katona. A theorem of finite sets. Theory of Graphs, Akademiai Kiado, 187–207, 1968.

[26] P. Keevash, E. Long. Stability for vertex isoperimetry in the cube. *J. Combin. Theory Ser. B* **145**: 113-144, 2020.

[27] D. Kleitman. Extremal hypergraph problems. Surveys in Combinatorics, Cambridge University Press, 44-65, 1979.

[28] J. Kruskal. The number of simplices in a complex. Mathematical Optimization Techniques, University of California Press, 251–278, 1963.

[29] I. Leader, E. Räty. A Note on Intervals in the Hales-Jewett Theorem. *Electron. J. Combin.* **25** (3): Paper 3.15, 2018.

[30] U. Leck. Nonexistence of a Kruskal-Katona type theorem for subword orders. *Combinatorica* **24**: 305-312, 2004.

[31] D. Lubell. A short proof of Sperner's lemma. *J. Combin. Theory*, **1** (2): 299, 1966.

[32] E. London. A new proof of the colored Kruskal–Katona theorem. Discrete Math. **126** (1-3): 217-223, 1994.

[33] L. Loomis, H. Whitney. An inequality related to the isoperimetric inequality. *Bull. Amer. Math. Soc.* **55**: 961–962, 1949.

[34] R. Martin, J. Smith. Induced saturation number. *Discrete Math.* **312** (21): 3096–3106, 2012.

[35] L. Meshalkin. Generalization of Sperner's theorem on the number of subsets of a finite set. *Theory Probab. Appl.*, **8** (2): 203-204, 1963.

[36] E. Räty. A grid generalisation of the Kruskal-Katona theorem. *arXiv preprint, arXiv:1908.02253,* 2019.

[37] E. Räty. An achievement game on a cycle. *Integers* **20**: Paper G7, 2020.

[38] E. Räty. Coordinate deletion of zeroes. *Electron. J. Combin.* **26** (3): Paper 3.50, 2019.

[39] E. Räty. Induced Saturation of P_6, *Discrete Math.* **343** (1), 2020.

[40] E. Räty. The Toucher-Isolator Game on Trees. *arXiv preprint, arXiv:2001.10498,* 2020.

[41] E. Räty. Uniqueness in Harper's vertex-isoperimetric theorem. *Discrete Math.* **343** (4), 2020.

[42] S. Shelah. Primitive recursive bounds for van der Waerden numbers. *J. Amer. Math. Soc.* **1** (3): 683–697, 1988.

[43] C. Spiegel. Personal communication, 2019.

[44] K. Yamamoto. Logarithmic order of free distributive lattice. *J. Math. Soc. Japan*, **6** (3-4): 343-353, 1954.