New results on some Erdős-Ko-Rado-type problems

by

Gabriel Currier

B.A., Pomona College 2016

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Mathematics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2021

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

New results on some Erdős-Ko-Rado-type problems

---

submitted by Gabriel Currier in partial fulfillment of the requirements for

the degree of Master of Science

in Mathematics

Examining Committee:

Richard Anstee, Professor, Department of Mathematics, UBC

Co-supervisor

József Solymosi, Professor, Department of Mathematics, UBC

Co-supervisor

Abstract

In this thesis, we study two longstanding problems in extremal set theory related to the Erdős-Ko-Rado theorem. In these problems, we have a collection $\mathcal{F}$ of k-subsets of an n-set that contains no copy of some forbidden substructure. The forbidden substructures in question here are known as clusters and simplices and are defined according to intersection and union constraints. The cluster conjecture that we consider was made by Mubayi in 2006 [24] and the simplex conjecture was made in 1974 by Chvátal [5]. We resolve completely the first of these two conjectures, and resolve the second for all but very small values of n.

Lay Summary

Extremal combinatorics deals with the following question: how large can a given structure be, given that it does not contain some specified substructure? Questions of this type are ubiquitous in combinatorics, and have applications all across other areas of mathematics and theoretical science. In this thesis, we consider some extremal problems for finite sets, where the structure in question is a collection of subsets of a set, and our substructure is defined according to the union (defined to be the elements in *any* of the subsets) and the intersection (defined to be the elements in *all* of the subsets) of our collection.

We obtain here some new results, providing the resolution of a conjecture from 2006, and the near-complete resolution of another from 1974.

Preface

Chapter 1 is largely expository. Chapter 2 is mostly adapted from two papers [7] and [6], both of which are sole work of the author. The first of these papers was published in the Journal of Combinatorial Theory Series A, and the second was published in the journal Combinatorica. Both problems are related to problems suggested to the author by Shahriar Shahriari and Ghassan Sarkis.

Table of Contents

Acknowledgements

There are many people to thank here. I would first like to thank my advisors József Solymosi and Richard Anstee for their support and guidance, and in particular to Richard Anstee for many hours spent helping me edit this thesis and the papers on which it is based. I am also significantly indebted to Shahriar Shahriari and Ghassan Sarkis for taking the time a number of years ago to work on some research problems with a bunch of second-year undergraduates. I would most certainly not have made progress on, or even known about these problems were it not for them.

I would like to thank Daniel Di Benedetto, Chi Hoi Yip, and Ethan White for a number of interesting discussions since coming to UBC. I would also like to thank Péter Frankl and the referees of [7] and [6] for helpful comments on the papers on which this thesis is based.

Finally, I would like to thank my parents for their endless support, and for teaching me many things about the world, even if only a little was math.

Chapter 1

Background

1.1 Introduction

The problems studied in this thesis fit into a field that is known commonly as Extremal Set Theory, and revolve around a famous result known as the Erdős-Ko-Rado (EKR) theorem. To give some context for these results, it is worth first discussing EKR and why it is an important theorem. We will jump right in.

Definition 1.1.1. *For a set X and integers n and k, we let $[n] := \{1, \ldots, n\}$, and use 2^X to denote the power set of X. Furthermore, we let $\binom{X}{k}$ denote the set of k-element subsets of X. Our main objects of study will be subsets $\mathcal{F} \subset 2^{[n]}$ or $\mathcal{F} \subset \binom{[n]}{k}$, which we will refer to as families or set systems. In the latter case, where $\mathcal{F} \subset \binom{[n]}{k}$, we will refer to $\mathcal{F}$ as k-*uniform.

There are many "basic" extremal questions that one could ask about a set system $\mathcal{F}$. As some obvious examples, one might choose: how big can $\mathcal{F}$ be before we must have $A \subset B$ for some distinct $A, B \in \mathcal{F}$? Or how big can $\mathcal{F}$ be before we have $A \cap B = \emptyset$ for some $A, B \in \mathcal{F}$? These are both classical problems in extremal set theory. We will focus on the latter, motivating the following definition: a family $\mathcal{F}$ is said to be *pairwise intersecting* if for every $A, B \in \mathcal{F}$ we have $A \cap B \neq \emptyset$. Thus, our question can be rephrased as follows: how big can a pairwise intersecting family $\mathcal{F} \subset 2^{[n]}$ be?

Example 1.1.2. *One obvious example is to take $\mathcal{F} = \{A \in 2^{[n]} : |A| > n/2\}$. Because all sets in $\mathcal{F}$ are large, any $A, B \in \mathcal{F}$ must intersect in at least one* *element. Furthermore, if n is even, then $|\mathcal{F}| = 2^{n-1}$. If n is odd, then it is slightly smaller but comparable. Another example we might consider would be to let $\mathcal{F} = \{A \in 2^{[n]} : 1 \in A\}$. This, again, has size 2^{n-1}, and any two elements of $\mathcal{F}$ must intersect at least in the element 1.*

Looking at these two examples, one might conjecture that the maximum size of a pairwise intersecting subfamily of $2^{[n]}$ is 2^{n-1}. This turns out to be correct, and has a simple and elegant proof.

Theorem 1.1.3. *Let $n \in \mathbb{N}$ and let $\mathcal{F} \subset 2^{[n]}$ be pairwise intersecting. Then $|\mathcal{F}| \leq 2^{n-1}$.*

Proof. For any $A \in 2^{[n]}$, $\mathcal{F}$ can contain only one of A and A^C. Thus, $\mathcal{F}$ can contain at most half of the elements of $2^{[n]}$, so $|\mathcal{F}| \leq \frac{2^n}{2} = 2^{n-1}$. □

This is neat and not too complicated. One might wonder if results of this type are all this straightforward. What if we assume $\mathcal{F}$ to be k-uniform?

Example 1.1.4. *Re-visiting our examples before gives some interesting results. If $k > n/2$, then our first example can still hold in a k-uniform setting. We let $\mathcal{F} = \binom{[n]}{k}$, and observe that $\mathcal{F}$ is pairwise intersecting, and that $|\mathcal{F}| = \binom{[n]}{k}$. This is clearly the best possible. In the case of $k \leq n/2$ it is not so straightforward. However, our second example still applies here, so if we let $\mathcal{F} = \{A \in \binom{[n]}{k} : 1 \in A\}$, we get a pairwise intersecting family with $|\mathcal{F}| = \binom{n-1}{k-1}$.*

Any family $\mathcal{F}$ where every element of $\mathcal{F}$ contains some $x \in [n]$ is known as a *star* centered at x. We observe, as in our example, that the maximum size of a star is $\binom{n-1}{k-1}$, and might again wonder if this is the “right” example to look at. That is, if this is indeed the unique largest pairwise intersecting family. That the answer is yes is given by the well-known Erdős-Ko-Rado (EKR) Theorem [12].

Theorem 1.1.5 (Erdős, Ko, Rado 1961)**.** *Let n, k be integers with $n \leq 2k$, and let $\mathcal{F} \subset \binom{[n]}{k}$ be pairwise intersecting. Then*

$$|\mathcal{F}| \leq \binom{n-1}{k-1}$$

and furthermore, if $n > 2k$, equality implies that $\mathcal{F}$ is a maximum-sized star.

Unlike Theorem 1.1.3, this is non-trivial to prove. It was proven originally for large n by Erdős, Ko and Rado in 1961. The lore behind this problem is that they came up with the proof as early as the 1930s, but interest in this kind of result was so low at the time that they waited almost 30 years to publish their proof. The main results of this thesis resolve some conjectures that originated as generalizations of the Erdős-Ko-Rado theorem. Thus, it seems worthwhile to take some time here to discuss why this is an important theorem.

1.2 Motivation

So, why is EKR an important result? There are many answers to this question, but I will focus on two here. The first is that it has many beautiful proofs that serve to unify a lot of different techniques in extremal set theory. There are many well-known proofs of EKR (we will discuss a few here), each of which highlights a different facet of the problem, and each of which uses a different technique which has proven to be quite useful in tackling other problems. Some of these techniques originated as proofs of EKR, but have evolved to be useful in a variety of other problems.

The second reason is a bit more vague but very important; that is, that set systems are very fundamental and important objects, and that EKR in some sense is the "right" basic question to ask about them. By "right" in this case, we mean a question that is non-trivial and interesting, but has an answer that is within our grasp. Showing this involves an explanation first of why set systems themselves are important, which is not too hard. Showing the latter claim is slightly more difficult, and involves a discussion of the history of graph and hypergraph theory. I will spend the next two sections discussing these motivations.

1.2.1 Proofs of EKR

In this section I will discuss various proofs of EKR, hopefully demonstrating some useful and beautiful techniques and showing how a wide variety of seemingly disparate ideas are brought together for this one theorem. The first and perhaps most famous that I will discuss is often referred to as the Katona cycle method, and originated as a proof of Erdős-Ko-Rado [21]. The idea is, at its core, a simple double counting argument. However, it uses permutations in a very neat and novel way that turns out to be useful in more general contexts as well. As an example, we will use this technique to prove a more general version of EKR later in this thesis (see Theorem 2.3.1) that will feature prominently our main results. It is also, however, a classical proof, so we will go over the original version here.

First proof of Theorem 1.1.5 (Cycle proof). Let $\mathcal{F} \subset \binom{[n]}{k}$ be a pairwise intersecting family, and let $C(n)$ denote the set of all cyclic permutations on n elements. Then, if we have $(a_0, \ldots, a_{n-1}) = \sigma \in C(n)$ and any $\mathcal{G} \subset \binom{[n]}{k}$, we define (with all subscripts henceforth taken mod n)

$$S_\sigma(\mathcal{G}) := \{A \in \mathcal{G} : A = \{a_i, a_{i+1}, \ldots, a_{i+(k-1)}\} \text{ for some } i \in [0, n-1]\}.$$

Observe trivially that $|S_\sigma(\mathcal{G})| \leq n$. Furthermore, for any such $A = \{a_i, \ldots, a_{i+(k-1)}\}$, we say that A has *starting point* i in $\sigma = (a_0, \ldots, a_{n-1})$. We wish to show that

$$|S_\sigma(\mathcal{F})| \leq k$$

for all $\sigma \in C(n)$. To start, we fix $\sigma \in C(n)$, and supposing $S_\sigma(\mathcal{F}) \neq \emptyset$

(otherwise the claim holds trivially) we take $A \in S_\sigma(\mathcal{F})$. Without loss of generality we assume $A = \{a_0, \ldots, a_{k-1}\}$. Then, let $A' \in (S_\sigma(\mathcal{F}) \setminus \{A\})$ and observe that since $A \cap A' \neq \emptyset$, it follows that A' has starting point in either $[n-(k-1), n-1]$ or $[1, k-1]$ Furthermore, because $n \geq 2k$, we cannot have $A_1, A_2 \in S_\sigma(\mathcal{F})$ with starting points $i_1 \in [n-(k-1), n-1]$ and $i_1 + k \in [1, k-1]$ respectively, because this would imply $A_1 \cap A_2 = \emptyset$. Combining these facts gives us $|S_\sigma(\mathcal{F})| \leq k$

Now, we wish to count pairs (σ, A) such that $\sigma \in C(n)$ and $A \in S_\sigma(\mathcal{F})$. First, fixing an element $\sigma \in C(n)$ we know $|S_\sigma(\mathcal{F})| \leq k$. Since $|C(n)| = (n-1)!$ we know then that the number of pairs cannot exceed $(n-1)!k$. Furthermore, each $A \in \mathcal{F}$ will be in $S_\sigma(\mathcal{F})$ for precisely $k!(n-k)!$ different $\sigma \in C(n)$. Combining these gives $(n-1)!k \geq k!(n-k)!|\mathcal{F}|$ or

$$|\mathcal{F}| \leq \frac{(n-1)!}{(k-1)!(n-k)!} = \binom{n-1}{k-1}$$

□

The second proof we will give uses a very common notion in extremal set theory called the *shadow*, and a fundamental theorem known as the Kruskal-Katona theorem.

Definition 1.2.1. *Let n and k be integers, and $\mathcal{F} \subset \binom{[n]}{k}$. Then the down and up-shadows respectively are defined to be*

$$\partial^-_{[n]}(\mathcal{F}) := \{D \in \binom{[n]}{k-1} : D \subset A \text{ for some } A \in \mathcal{F}\}$$

and

$$\partial^+_{[n]}(\mathcal{F}) := \{D \in \binom{[n]}{k+1} : A \subset D \text{ for some } A \in \mathcal{F}\}$$

In plain language, the down shadow consists of all $(k-1)$ sets contained in the elements of $\mathcal{F}$, and the up shadow consists of all $(k+1)$-sets containing elements of $\mathcal{F}$. The Kruskal-Katona theorem tells us the minimum size of a shadow for a family $\mathcal{F}$ of a given size. To state it, we need to extend binomial coefficients to all $x \in \mathbb{R}$ and $k \in \mathbb{Z}^+$ as

$$\binom{x}{k} = \frac{x(x-1)\dots(x-k+1)}{k!}$$

The following is a weaker but very convenient version of the Kruskal-Katona Theorem. It is due to Lovász [23]:

Theorem 1.2.2. *Let n and k be integers with $k < n$, and let $\mathcal{F} \subset \binom{[n]}{k}$. Furthermore, let y be the real number defined by $|\mathcal{F}| = \binom{y}{k}$. Then*

$$|\partial^-_{[n]}(\mathcal{F})| \geq \binom{y}{k-1}.$$

Corollary 1.2.3. *Let $n, k, \mathcal{F}$ be as before, and let y be the real number defined by $|\mathcal{F}| = \binom{y}{n-k}$. Then*

$$|\partial^+_{[n]}(\mathcal{F})| \geq \binom{y}{n-k-1}.$$

We will use this corollary to provide another straightforward proof of EKR, a version of which is due originally to Daykin [8].

Second proof of Theorem 1.1.5 (Shadow proof). We focus first on the case $n = 2k$, and note that for any $A \in \binom{[n]}{k}$, $\mathcal{F}$ can contain either A or A^C, and thus $|\mathcal{F}| \leq \frac{1}{2}\binom{2k}{k} = \binom{2k-1}{k-1}$.

Now, we induct on $n - 2k$, and observe that the case $n - 2k = 0$ has been shown. Suppose $|\mathcal{F}| \geq \binom{n-1}{k-1}$, and note first that while $\mathcal{F} \subset \binom{[n]}{k}$, it can equally well be seen as a collection of k-subsets of $[n+1]$. Using this, we observe that

$$|\mathcal{F}| = |\partial^+_{[n+1]}(\mathcal{F})| - |\partial^+_{[n]}(\mathcal{F})|,$$

which is a straightforward consequence of the fact that

$$\partial^+_{[n+1]}(\mathcal{F}) \setminus \partial^+_{[n]}(\mathcal{F}) = \{A \cup \{n+1\} : A \in \mathcal{F}\}.$$

Now, we provide estimates on $|\partial^+_{[n+1]}(\mathcal{F})|$ and $|\partial^+_{[n]}(\mathcal{F})|$. Observing that $\partial^+_{[n+1]}(\mathcal{F})$ must be pairwise intersecting since $\mathcal{F}$ is, and since $(n+1)-2(k+1) = n-2k-1$ we get by induction that $|\partial^+_{[n+1]}(\mathcal{F})| \leq \binom{n}{k}$. Furthermore, by the Corollary 1.2.3 and the fact that $|\mathcal{F}| \geq \binom{n-1}{k-1} = \binom{n-1}{n-k}$, we have that $|\partial^+_{[n]}(\mathcal{F})| \geq \binom{n-1}{n-k-1} = \binom{n-1}{k}$. Combining these, we get

$$|\mathcal{F}| = |\partial^+_{[n+1]}(\mathcal{F})| - |\partial^+_{[n]}(\mathcal{F})| \leq \binom{n}{k} - \binom{n-1}{k} = \binom{n-1}{k-1}$$

□

The final proof we will give involves another common technique in extremal set theory known as shifting, which was developed for the original proof of Erdős-Ko-Rado [12].

Definition 1.2.4. *Let n and k, with $k < n$, and let $\mathcal{F} \subset \binom{[n]}{k}$ and $A \in \mathcal{F}$. Then, if i, j are integers with $1 \leq i < j \leq n$, we define the shifting operator s_{ij} as follows: $s_{ij}(A) = (A \setminus \{j\}) \cup \{i\}$ if $A \cap \{i,j\} = \{j\}$ and $(A \setminus \{j\}) \cup \{i\} \notin \mathcal{F}$, and $s_{ij}(A) = A$ otherwise. Furthermore, we define $s_{ij}(\mathcal{F}) := \{s_{ij}(A) : A \in \mathcal{F}\}$. If $s_{ij}(\mathcal{F}) = \mathcal{F}$ for all $i < j$, then we say that $\mathcal{F}$ is* shifted.

We make a couple of basic observations: first, it is easily verified that $|s_{ij}(\mathcal{F})| = |\mathcal{F}|$ for all i and j. Furthermore, if $\mathcal{F}$ is pairwise intersecting, then so must be $s_{ij}(\mathcal{F})$. To see this, suppose the contrary, that is, that there is some family $\mathcal{F} \subset \binom{[n]}{k}$ and $i < j$ such that $\mathcal{F}$ is pairwise intersecting but $s_{ij}(\mathcal{F})$ is not. Take $A, B \in s_{ij}(\mathcal{F})$ such that $A \cap B = \emptyset$. Since $\mathcal{F}$ was pairwise intersecting, at least one of A or B must have been shifted, but at most one of them can contain j. Suppose without loss of generality that $j \in A$, and note that since $\mathcal{F}$ was intersecting, we must have $i \in B$. However, by the definition of the shifting operator, this implies $(B \setminus \{i\}) \cup \{j\}$ is in $\mathcal{F}$. However, then $(A \setminus \{j\}) \cup \{i\}$ and $(B \setminus \{i\}) \cup \{j\}$ are both elements of $\mathcal{F}$, and are disjoint. This is a contradiction.

Thus, applying the shifting operator to a family cannot alter the fact that it is pairwise intersecting. Knowing this, it is sufficient to show Theorem 1.1.5 in the case that $\mathcal{F}$ is shifted. For the following proof, we will use the following notation (the first of which will feature prominently also in the proof of our main results later on). For any $D \subset [n]$, we define

$$\triangledown_{\mathcal{F}}(D) := \{B \in \mathcal{F} : D \subset B\}$$

and

$$\overline{\triangledown_{\mathcal{F}}(D)} := \{B \setminus D : B \in \triangledown_{\mathcal{F}}(D)\}.$$

Informally, $\triangledown_{\mathcal{F}}(D)$ is all elements of $\mathcal{F}$ containing D, and $\overline{\triangledown_{\mathcal{F}}(D)}$ is that same family with D taken out from each element. If $D = \{x\}$ for some $x \in [n]$ we write simply $\triangledown_{\mathcal{F}}(x)$ and $\overline{\triangledown_{\mathcal{F}}(x)}$. For the following proof we will use only $x = n$, but later on we will use this notation for arbitrary x.

Third proof of Theorem 1.1.5 (Shifting proof). Suppose that $\mathcal{F} \subset \binom{[n]}{k}$ is shifted. We will show the result by induction on n, and note that in the previous proof we dealt already with the case $n = 2k$, and that the case $k = 1$ is trivial. We assume now that $n > 2k$ and we consider families $\overline{\triangledown_{\mathcal{F}}(n)}$ and $\mathcal{F} \setminus \triangledown_{\mathcal{F}}(n)$. We note that $\mathcal{F} \setminus \triangledown_{\mathcal{F}}(n)$ is a subset of $\mathcal{F}$ and is thus clearly pairwise intersecting, so by induction we have $|\mathcal{F} \setminus \triangledown_{\mathcal{F}}(n)| \leq \binom{n-2}{k-1}$. We claim that $\overline{\triangledown_{\mathcal{F}}(n)}$ must be pairwise intersecting as well. To see this, suppose the contrary, that is, that there exist $D_1, D_2 \in \overline{\triangledown_{\mathcal{F}}(n)}$ such that $D_1 \cap D_2 = \emptyset$, and select $m \in [n-1] \setminus (D_1 \cup D_2)$ (this is possible since $n - 1 - 2(k-1) > 2k - 1 - 2k + 2 \geq 1$). By shiftedness, we must have $D_1 \cup \{m\} \in \mathcal{F}$. However, $D_2 \cup \{n\} \in \mathcal{F}$ as well, and these two elements will be disjoint, contradicting that $\mathcal{F}$ is pairwise intersecting. Thus $\overline{\triangledown_{\mathcal{F}}(n)} \subset \binom{[n-1]}{k-1}$

is pairwise intersecting as well, so by induction $\overline{\nabla_{\mathcal{F}}(n)} \leq \binom{n-2}{k-2}$. Noting that every element of $\mathcal{F}$ either contains n or does not, we have

$$|\mathcal{F}| = |\overline{\nabla_{\mathcal{F}}(n)}| + |\mathcal{F} \setminus \nabla_{\mathcal{F}}(n)| \leq \binom{n-2}{k-2} + \binom{n-2}{k-1} = \binom{n-1}{k-1}$$

thus completing the proof.

□

There are other proofs of Erdős-Ko-Rado that are worth noting, but that we will not go into here. Of particular interest is an algebraic proof that shows shows EKR by bounding the maximum size of an independent set in a graph known as the Kneser graph, which relates to coding theory and has seen a number of other applications. We direct the interested reader to [19] for a good reference on these techniques.

1.2.2 Set systems and hypergraphs

In this section I will try and give another reason why Erdős-Ko-Rado is in some sense fundamental. This involves first establishing a basic fact; that is, that set systems are really just hypergraphs in disguise.

Definition 1.2.5. *A hypergraph $G = (V, E)$ is a tuple consisting of two sets, a vertex set V and an edge set $E \subset 2^V$. G is known as k-uniform if $E \subset \binom{V}{k}$, and if $k = 2$ then G is known simply as an (ordinary) graph.*

You will perhaps notice that the definition of a hypergraph is nearly identical to that of a set system; the only real difference being that in set systems, the vertex set is known instead as the "base set," and the collection of edges is known as a "family." The notation of set systems does provide some advantages, that is, that the base set is often standardized to be $[n]$, making the proofs of many results cleaner. Beyond that, though, they are identical. This makes the motivation of the study of set systems a bit easier.

Hypergraphs have become, especially recently, very important and central objects of study in combinatorics, and have seen very important applications in related fields such as discrete geometry and combinatorial number theory. See, for example [20, 28] and [2].

It is not surprising, then, that we are interested in understanding basic extremal properties of hypergraphs. However, it remains to be answered: why is EKR a good basic extremal question to ask? The answer involves going into a bit of the history of graph theory. The early extremal questions on graphs were concerned, rather naturally, with the most basic structures that researchers could conceive of. One of the earliest examples is the graph K_3, also known as a triangle, which is simply a collection of vertices v_1, v_2, v_3 with all three possible edges between the vertices. One basic extremal question is how many edges must be in a graph before we are guaranteed to find a triangle *somewhere* in our graph.

Example 1.2.6. *Let n be an even integer. Consider the complete bipartite graph with $n/2$ vertices in each class, denoted $K_{n/2,n/2}$. Taking any three vertices in our graph, we observe that at least two must be in the same vertex class, and thus must not share an edge. Thus, $K_{n/2,n/2}$ contains no K_3 as a subgraph.*

This gives us an example of a graph with $n^2/4$ edges that contains no triangle. This seems like a fair number of edges, but one might still wonder: can we do any better? It turns out that the answer is no, as shown by Mantel over 100 years ago.

Theorem 1.2.7 (Mantel, 1907)**.** *Let G be a graph on n vertices, that contains no K_3. Then, G has at most $\lfloor n^2/4 \rfloor$ edges.*

Theorem 1.2.7 was generalized as Turán's theorem [29], one of the most important results in extremal graph theory. Turán's theorem tells us the maximum number of edges in a graph containing no K_r, or complete graph on r vertices. Another important structure in graph theory is the Hamiltonian cycle, which is defined to be a cycle that passes through every vertex of our graph exactly once. Dirac's Theorem [9] gives us a relationship between minimum degree of a graph (denoted $\delta(G)$) and the existence of a Hamiltonian cycle.

Theorem 1.2.8 (Dirac, 1952)**.** *Let G be a graph on n vertices. If $\delta(G) \geq n/2$, then G must contain a Hamiltonian cycle.*

Both Mantel's and Dirac's Theorems have been generalized numerous ways, and have generated numerous interesting questions in graph theory and related fields. While many of these questions concern ordinary graphs, one might wonder what happens when we attempt to generalize these theorems to the hypergraph setting. The answers to these questions in the ordinary graph context has been known for at least 50 years (and over to 100 in the case of Mantel's theorem), and thus one would imagine that we have a good grasp, now, on at least some of the basic hypergraph versions of these questions.

Somewhat surprisingly, despite significant research attention, many of these questions remain largely elusive. This is perhaps most easily demonstrated by examining the hypergraph versions of Turán's Theorem, the most well-studied of which is known as Turán's (3,4) problem. This question asks how many edges we can have in a 3-uniform hypergraph that contains no complete graph on 4 edges, denoted by K_4^3. This seems at first glance to be a very similar question to the one about triangles. It is furthermore quite old, having been proposed originally by Turán in 1961. However, it remains an open problem, even asymptotically. We have now, in the past 20 or so years, near optimal asymptotic results, but these have been obtained only through the use of some reasonably technical and very recent machinery. Even less is known about the analogous question for complete subgraphs on r vertices of ℓ-uniform hypergraphs, for larger r and ℓ.

Hypergraph generalizations of Dirac's theorem have proven somewhat more accessible. Even in that context, though, most near-optimal results have come only in the last 20 years. This is a trend across much of extremal hypergraph theory: fundamental theorems on graphs with reasonably straightforward proofs are often very, very difficult when considered in a hypergraph context. We now come back to our original question: if extremal set theory is in some sense all about hypergraphs, then why is it not focused on these same questions, given that they are so fundamental on ordinary graphs? The answer, quite simply, is that they are too difficult. As discussed here, even when we try to generalize these theorems to edges of size 3 or 4, they become intensely hard. The analogous questions for graphs with arbitrarily large edges are, for the moment, hopelessly out of reach.

The moral of this story is this: extremal problems which are relatively easy on graphs can be hideously difficult when considered in the hypergraph context. Researchers have often then looked away from graphs as a source of extremal questions on hypergraphs, and asked a simpler question: what are the most basic extremal questions that one can ask about hypergraphs of arbitrarily large edge size? One of the most fundamental questions is precisely that which is answered by EKR. That is, how many edges must be in a k-uniform hypergraph before two of them must be disjoint? In the context of an ordinary graph this question is trivial, but as discussed before, for hypergraphs it is not. In fact, it seems in some sense to be the "right" level of difficult; that is, hard enough to be non-trivial and interesting, but easy enough to be provable in a wide variety of ways. The answer to the motivational questions seems a little easier now. Why should we care about EKR? Well, it is one of the most fundamental and basic results known about a fundamental and important object. What's more, as discussed in the previous section, the techniques generated through study of this problem have seen applications in many other problems. It's natural to wonder, if this is such a fundamental theorem, in what ways it can be extended and generalized.

1.3 Generalizations of EKR

One way to view Erdős-Ko-Rado is as a theorem about forbidding intersections of size zero between pairs of sets in a family. From this perspective, there are two very natural generalizations that one might consider. First, we might wonder what happens if we forbid intersections of other sizes, and second, we might ask what happens if our forbidden structure consists of more than two sets. This leads to a number of interesting problems. The first that we will consider, and the one with which there has been probably the most success, is to consider what happens if we forbid intersections that are smaller than a certain threshhold. In other words, what happens if we consider families with no A, B such that $|A \cap B| < t$ for some $t \geq 1$? Such a family is known as *t-intersecting*.

Example 1.3.1. *Let $n, k, t \in \mathbb{N}$, suppose $Y \in \binom{[n]}{t}$ and let $\mathcal{F} = \{A \in \binom{[n]}{k} : Y \subset A\}$. Such a family is known as a t-star, and it is clear that $\mathcal{F}$ is t-intersecting. We can observe furthermore that $|\mathcal{F}| = \binom{n-t}{k-t}$.*

As before, we want to know if $\binom{n-t}{k-t}$ is the maximum possible size of a t-intersecting family. This is one of the few problems of this type where results are known for smaller n. In the original proof of EKR, it was shown that this is in fact the case, assuming $n > n_0(k, t)$. However, the complete result (valid for all $n \geq (k - t + 1)(t + 1)$) was not known until almost over 20 years later, and the situation when $n < (k - t + 1)(t + 1)$ was not known until another 15 years after that, with the landmark result of Ahlswede and Khachatrian [1], which describes the extremal examples for all n for which the problem makes sense.

If we take this problem one step further, we already run into results that have proven inaccessible for small n. Proposed originally in 1974, the Erdős- Sós 'forbidding one intersection' problem conjectures that any family $\mathcal{F}$ containing no A, B with $|A \cap B| = t - 1$ should have size no bigger than $\binom{n-t}{k-t}$, and in 1985, Frankl and Füredi resolved this conjecture for $n \geq n_0(k, t)$

[16]. However, 35 years later, the situation for smaller n is still not known.

We now consider the second question: what if we still want to forbid intersections of size zero, but do so for more than two sets? That is, what if we consider the maximum size of $\mathcal{F}$ containing no A, B, C satisfying some sort of disjointness condition? We will see that this admits a number of possible interpretations, and that the most obvious are not always the most interesting. We could, for example, consider families containing no A, B, C with $A \cap B \cap C = \emptyset$. However, if $\mathcal{F}$ contains two disjoint elements A, B, then taking any remaining $C \in \mathcal{F}$ will give also $A \cap B \cap C = \emptyset$. Thus, this particular generalization is implied by the original version of EKR. Another possibility would be to forbid A, B, C that are pairwise disjoint. This gives us the Erdős matching conjecture [10], which is well understood for large n but for which the situation for small n is not entirely known.

We will consider in this thesis two more generalizations of this type. They are similar in that they both generalize EKR to more than two sets in a way so that the forbidden substructure is disjoint in some sense, but not *too* disjoint, and thus have a different flavor than the above problems. Beyond this, they seem like quite different problems. However, we will show that both conjectures can be proven, even for small n, using variations on the same technique. As is evident at this point, this is one of the rare times that a result of this type is known for a small value of n.

The first has roots in a question asked by Erdős in 1971 [11]. Motivated actually by Mantel's theorem, he wanted know what the maximum size of a set systems containing no "triangle" would be. For set systems, though, the notion of a triangle was not as well defined as it was in ordinary graphs. However, the eventual definition manages to still suit the moniker quite well.

Definition 1.3.2. *Suppose we have $A, B, C \in \binom{[n]}{k}$ such that $A \cap B \cap C = \emptyset$. Suppose furthermore that $A \cap B$, $B \cap C$ and $A \cap C$ are all non-empty. Then we refer to A, B, C as a* triangle.

A maximum-sized star (which is of size $\binom{n-1}{k-1}$) clearly contains no triangle, and it was conjectured that this was in fact the extremal configuration whenever $k \geq 3$.

Motivated by this question, Chvátal took the conjecture a step further. We define a d-simplex to be a collection $A_1, \ldots, A_{d+1} \in \binom{[n]}{k}$ such that $\cap_i A_i = \emptyset$ but $\cap_{i \neq j} A_i \neq \emptyset$ for all $1 \leq j \leq d+1$. It is clear that a 1-simplex is simply two disjoint sets, and that a 2-simplex is a triangle. Chvátal conjectured in [5] that the maximum size of a family $\mathcal{F}$ containing no d-simplex would be $\binom{n-1}{k-1}$ as well, under the assumption that $d+1 \leq k$. In [14] this conjecture was first shown for a wide range of parameters, and in 1987 it was shown to hold for all $n > n_0(k,d)$ by Frankl and Füredi [17]. It was not known to hold for any smaller values of n until the case $d = 2$ (that is, the case of triangles) was resolved completely by Mubayi and Verstraëte in 2005 [27]. Until now, there were no results known for $d \geq 3$ and small n.

The second question we will consider captures a different aspect of the original Erdős-Ko-Rado Theorem. What if our forbidden substructure contains d different sets with empty intersection, but where the sets are not too spread out? In other words, what if we forbid $A_1, \ldots, A_{d+1}$ such that $A_1 \cap \cdots \cap A_{d+1} = \emptyset$ and $|A_1, \ldots, A_{d+1}| \leq s$ for some s? In the 1980s, Katona asked this question in the case that $d = 2$, but the question was later studied for larger d. The following example is due to Frankl and Füredi in [15].

Example 1.3.3. *We start by partitioning $[n]$ into k classes $X_1, \ldots, X_k$ of equal size, and let $\mathcal{F} = X_1 \times \cdots \times X_k$. Then, $\mathcal{F}$ contains no $A_1, \ldots, A_{d+1}$ with $A_1 \cap \cdots \cap A_{d+1} = \emptyset$ and $|A_1 \cup \cdots \cup A_{d+1}| < 2k$ For any $d \geq 2$. Furthermore, we see that $|\mathcal{F}| = \left(\frac{n}{k}\right)^k$, which is in turn larger than $\binom{n-1}{k-1}$.*

Thus, if $s < 2k$, we cannot get a bound of the type $|\mathcal{F}| \leq \binom{n-1}{k-1}$. What about for $s \geq 2k$? Somewhat surprisingly, Frankl and Füredi also showed in [15] that, for $d = 2$ and sufficiently large n, the bound of $\binom{n-1}{k-1}$ holds for *any* $s \geq 2k$, and in [24] Mubayi extended this to all $n \geq 2k$. In the same paper, he denoted any $A_1, \ldots, A_{d+1}$ with $A_1 \cap \cdots \cap A_{d+1} = \emptyset$ and $|A_1 \cup \cdots \cup A_{d+1}| \leq 2k$

a d-cluster, and conjectured that any d-cluster-free family $\mathcal{F}$ would obey the same bound of $|\mathcal{F}| \leq \binom{n-1}{k-1}$. This was, again, shown to be true in the case of $n > n_0(k, d)$ by Mubayi for $d = 3$ in [25] and by Mubayi and Ramadurai for $d \geq 4$ in [27]. However, as with simplices, results for smaller n remained elusive.

The concepts of d-simplices and d-clusters clearly share some similarities, but the conditions $\cap_{i \neq j} A_i \neq \emptyset$ and $|A_1 \ldots, A_{d+1}| \leq 2k$ seem qualitatively very different. Thus, it was somewhat of a surprise when Keevash and Mubayi drew a connection between the two objects in [22]. Specifically, they showed that a set system $\mathcal{F}$ that contained no d-cluster or d-simplex would obey $|\mathcal{F}| \leq \binom{n-1}{k-1}$, under the condition that both k/n and $n/2 - k$ were bounded away from zero. In this same paper, they defined a d-simplex-cluster to be a collection of sets that is both a d-simplex and a d-cluster, and made the following conjecture.

Conjecture 1.3.4 ([22])**.** *Suppose that $3 \leq d+1 \leq k$ and $n \geq \frac{d+1}{d}k$, and let $\mathcal{F}$ be d-simplex-cluster-free. Then,*

$$|\mathcal{F}| \leq \binom{n-1}{k-1}$$

with equality only if $\mathcal{F}$ is a maximum-sized star.

In the upcoming chapter, we will prove some new results on these conjectures about simplices, clusters, and simplex-clusters.

Chapter 2

Main results

2.1 General discussion

In the previous section, we defined d-clusters and simplices. We present these definitions in a unified format here as a reference for the upcoming sections.

Definition 2.1.1. *Suppose $n, k, d \in \mathbb{N}$ be such that $d+1 \leq k$. Suppose furthermore that $A_1, \ldots, A_{d+1} \in \binom{[n]}{k}$ are such that $A_1 \cap \cdots \cap A_{d+1} = \emptyset$. Then, if $\cap_{i \neq j} A_i \neq \emptyset$ we say that $A_1, \ldots, A_{d+1}$ is a d-*simplex*, and if $|A_1 \cup \cdots \cup A_{d+1}| \leq 2k$, we say that it is a d-*cluster*. If both of these hold, we say that $A_1, \ldots, A_{d+1}$ is a d-*simplex-cluster*.*

We will show here a couple of related results, all valid for slightly different values of n, k, d and varying in style and complexity. These results all share theme that they are valid for small n. This is a range that, as discussed before, has historically proven very difficult. We will first prove the d-cluster conjecture.

Theorem 2.1.2. *Let $3 \leq d+1 \leq k$, and $n \geq \frac{d+1}{d}k$. Furthermore, suppose that $\mathcal{F}$ contains no d-cluster. Then,*

$$|\mathcal{F}| \leq \binom{n-1}{k-1}$$

with equality only if $\mathcal{F}$ is a maximum-sized star.

We will give here a proof of Theorem 2.1.2 in the range $n \geq 2k$, which, when combined with a result of Frankl from [13] for $n < 2k$ proves the conjecture for the full range of $n \geq \frac{d+1}{d}k$. Our final result will be the following, which resolves the d-simplex-cluster for almost all values of n.

Theorem 2.1.3. *Suppose that $4 \leq d+1 \leq k$ and $n \geq 2k-d+2$, and that $\mathcal{F} \subseteq \binom{[n]}{k}$ contains no d-simplex-cluster. Then*

$$|\mathcal{F}| \leq \binom{n-1}{k-1},$$

with equality only if $\mathcal{F}$ is a maximum sized star.

Combined with the results of [27] for the case $d = 2$, this resolves Conjecture 1.3.4 for all values of d when $n \geq 2k - d + 2$. This does not cover quite the full range of the conjecture, but it notably covers the range of $n \geq 2k$, which is the range for the orginal EKR theorem.

We take a moment here to discuss the range $n < 2k$. Intersection problems of this type tend to have a slightly different flavor when considered for these values of n. There are sometimes obvious reasons for this - the Erdős-Ko-Rado theorem in its original form, for example, does not make much sense when considered in this context because any two k-sets will automatically intersect. As another example, the problem of clusters is different in this range because the union condition holds automatically, so a d-cluster-free family is simply a family with no $d+1$ sets that have empty intersection. There is, however, some history of results for problems of this type in the range $n < 2k$. The most notable example is perhaps the previously-mentioned Complete Intersection Theorem of Ahlswede and Khachatrian [1]. For simplices, there are two results known to the author. The first is for $d = 2$, when the full range of $n \geq 3k/2$ was shown in [27]. The other example (and the only for general d) is from [13], where the case of $n < k\frac{d}{d-1}$ is resolved.

2.2 Notation

In order to efficiently describe our results we will need to recall some notation discussed briefly in previous sections, as well as a new definition.

Definition 2.2.1. *Suppose $\mathcal{F} \subseteq \binom{[n]}{k}$ and we have $A \in \mathcal{F}$ and $D \subseteq [n]$ and $x \in [n]$. Then, we recall the definitions $\bigtriangledown_{\mathcal{F}}(D), \overline{\bigtriangledown_{\mathcal{F}}(D)}$*

$$\bigtriangledown_{\mathcal{F}}(D) := \{B \in \mathcal{F} : D \subseteq B\}$$

$$\overline{\bigtriangledown_{\mathcal{F}}(D)} := \{B \setminus D : B \in \bigtriangledown_{\mathcal{F}}(D)\}$$

where, as before, if $D = \{x\}$, then we write simply $\bigtriangledown_{\mathcal{F}}(x)$ and $\overline{\bigtriangledown_{\mathcal{F}}(x)}$. We furthermore define $\alpha^i_{\mathcal{F}}(A) \subseteq A$ as

$$\alpha^i_{\mathcal{F}}(A) := \{x \in A : \bigtriangledown_{\mathcal{F}}(A \setminus \{x\}) = i\}.$$

The first two pieces of notation are both related to the common combinatorial notion of link or trace. In some sense it is superfluous to use both $\bigtriangledown_{\mathcal{F}}(D)$ and $\overline{\bigtriangledown_{\mathcal{F}}(D)}$. However, each is useful in different contexts (in particular, $\overline{\bigtriangledown_{\mathcal{F}}(D)}$ will be useful when inducting), and having both at our disposal makes several of our proofs much easier to look at. We note furthermore trivially that $| \bigtriangledown_{\mathcal{F}} (D)| = |\overline{\bigtriangledown_{\mathcal{F}}(D)}|$, which we will use freely when convenient. The third definition can be thought of as a measure of the removability of the elements of $A \in \mathcal{F}$. By this we mean that, if we have $x \in \alpha^i_{\mathcal{F}}(A)$ for some $i \geq 2$, then we can remove x from A without increasing its size very much - that is, we can find $B \in \mathcal{F}$ such that $x \notin B$ but $|A \cup B|$ is small. Furthermore, if $i \geq 3$ we have greater flexibility in choosing B that we will leverage later on. These are useful notions because they provide a way to construct d-clusters and simplices in a way that is both controlled and relatively easy to count.

2.3 An extended version of EKR

One of our primary tools for constructing d-clusters and simplices will be a generalization of EKR that gives a bound on the size of a family $\mathcal{F}$ where some (but not necessarily all) elements of $\mathcal{F}$ intersect all the remaining elements of $\mathcal{F}$. That is, we wish to give an upper bound on the size of an $\mathcal{F}$ that is allowed to have some sub-family that does not have to be pairwise intersecting. The following is due originally to Borg in [3] and a different proof was given by Borg and Leader in [4].

Theorem 2.3.1. *Let $n \geq 2k$, and suppose $\mathcal{F}^* \subseteq \mathcal{F} \subseteq \binom{[n]}{k}$ have the property that for any $A, B \in \mathcal{F}$ such that $A \cap B = \emptyset$, we have $A, B \in \mathcal{F}^*$. Then*

$$|\mathcal{F}| \leq \binom{n-1}{k-1} + \frac{n-k}{n}|\mathcal{F}^*|,$$

where, for $n > 2k$, equality is achieved only if $\mathcal{F}^* = \mathcal{F} = \binom{[n]}{k}$ or if $\mathcal{F}^* = \emptyset$ and $\mathcal{F}$ is a maximum sized star.

Note that this theorem is itself a generalization of EKR; if one sets $\mathcal{F}^* = \emptyset$, then Theorem 2.3.1 says simply that a pairwise intersecting family has size at most $\binom{n-1}{k-1}$. It is, furthermore, intimately related to the notion of cross-intersecting families.

Definition 2.3.2. *Let n, k be integers, and suppose we have $\mathcal{F}, \mathcal{G} \subset \binom{[n]}{k}$. Then, if $A \cap B \neq \emptyset$ for all $A \in \mathcal{F}, B \in \mathcal{G}$, we say that $\mathcal{F}$ and $\mathcal{G}$ are* cross-intersecting.

In fact, the condition required for Theorem 2.3.1 to hold is equivalent to $\mathcal{F}$ and $\mathcal{F} \setminus \mathcal{F}^*$ being cross-intersecting. A more general result of this type, phrased in terms of cross-intersecting families is obtained in Theorem 9 of [18] . The notion of cross-intersecting families is very important in extremal set theory; given the major part that Theorem 2.3.1 plays in our results, the connection between cross-intersection and simplices and clusters is an interesting one and perhaps deserves further attention. We provide here a proof of Theorem 2.3.1 for completeness.

Proof. We will proceed by the Katona cycle method [21]. First, we let $C(n)$ denote the set of all cyclic permutations on n elements. Then, if we have $(a_0, \ldots, a_{n-1}) = \sigma \in C(n)$ and $\mathcal{G} \subset \binom{[n]}{k}$, we define (with all subscripts henceforth taken mod n)

$$S_\sigma(\mathcal{G}) := \{A \in \mathcal{G} : A = \{a_i, a_{i+1}, \ldots, a_{i+(k-1)}\} \text{ for some } i \in [0, n-1]\}.$$

Observe trivially that $|S_\sigma(\mathcal{G})| \leq n$. As before, for any such $A = \{a_i, \ldots, a_{i+(k-1)}\}$, we say that A has *starting point* i in $\sigma = (a_0, \ldots, a_{n-1})$. Now, we wish to prove the following:

(i) $|S_\sigma(\mathcal{F} \setminus \mathcal{F}^*)| \leq k$ for all $\sigma \in C(n)$

(ii) if $S_\sigma(\mathcal{F} \setminus \mathcal{F}^*) \neq \emptyset$, then $|S_\sigma(\mathcal{F}^*)| \leq 2(k - |S_\sigma(\mathcal{F} \setminus \mathcal{F}^*)|)$ for all $\sigma \in C(n)$

Let $\sigma = (a_0, \ldots, a_{n-1})$ as before, suppose $S_\sigma(\mathcal{F} \setminus \mathcal{F}^*) \neq \emptyset$, and take $A \in S_\sigma(\mathcal{F} \setminus \mathcal{F}^*)$. Furthermore, suppose without loss of generality that $A = \{a_0, \ldots, a_{k-1}\}$. Then, let $A' \in (S_\sigma(\mathcal{F}) \setminus \{A\})$ and observe that since $A \cap A' \neq \emptyset$, it follows that A' has starting point in either $[n-(k-1), n-1]$ or $[1, k-1]$. Suppose then that we have $A_1, A_2 \in (S_\sigma(\mathcal{F}) \setminus \{A\})$ with starting points $i_1 \in [n-(k-1), n-1]$ and $(i_1 + k) \in [1, k-1]$ in σ respectively. Since $n \geq 2k$ this implies $A_1 \cap A_2 = \emptyset$ and thus that $A_1, A_2 \in \mathcal{F}^*$. Because only one element of $\mathcal{F}$ may have a given starting point in σ, we can combine these facts to get both (i) and (ii). Now, we define subsets of $C(n)$

$$C_j := \{\sigma \in C(n) : |S_\sigma(\mathcal{F} \setminus \mathcal{F}^*)| = j\},$$

and using (i) we observe that $C_0, C_1, \ldots, C_k$ partition $C(n)$. Using (i) and (ii), and since every $A \in \mathcal{F}$ is in $S_\sigma(\mathcal{F})$ for precisely $k!(n-k)!$ different

$\sigma \in C(n)$ we get

$$|\mathcal{F} \setminus \mathcal{F}^*|k!(n-k)! = \sum_{\sigma \in C(n)} |S_\sigma(\mathcal{F} \setminus \mathcal{F}^*)| = \sum_{1 \le i \le k} |C_i| i$$

and

$$|\mathcal{F}^*|k!(n-k)! = \sum_{\sigma \in C(n)} |S_\sigma(\mathcal{F}^*)| \le n|C_0| + \sum_{1 \le i \le k} |C_i| 2(k-i).$$

Combining these yields

$$\begin{aligned} |\mathcal{F}^*| + \binom{n}{k} |\mathcal{F} \setminus \mathcal{F}^*| &\le \frac{\Big(n|C_0| + \sum_{i=1}^k 2(k-i)|C_i|\Big) + (n/k)\Big(\sum_{i=1}^k i|C_i|\Big)}{k!(n-k)!} \\ &= \frac{n|C_0| + n|C_k| + \sum_{i=1}^{k-1} \frac{in+2k(k-i)}{k}|C_i|}{k!(n-k)!}. \end{aligned} \tag{2.1}$$

A quick calculation gives us that, for all $1 \le i \le k-1$

$$\frac{in + 2k(k-i)}{k} \le \frac{in + n(k-i)}{k} = n, \tag{2.2}$$

with equality only if $n = 2k$. Combining (2.1) and (2.2), since $|C_0| + \cdots + |C_k| = |C(n)| = (n-1)!$, we get

$$\begin{aligned} \left(\frac{k-n}{k}\right)|\mathcal{F}^*| + \binom{n}{k}|\mathcal{F}| &= |\mathcal{F}^*| + \binom{n}{k}|\mathcal{F} \setminus \mathcal{F}^*| \\ &\le \frac{n(|C_0| + \cdots + |C_k|)}{k!(n-k)!} \\ &= \frac{n!}{k!(n-k)!} \\ &= \binom{n}{k}. \end{aligned}$$

Dividing both sides by n/k we get our desired inequality. Now, suppose

$n > 2k$ and we have equality. Note that in this case we do not have equality in (2.2) and so $C(n) = C_0 \cup C_k$. Furthermore, if we take arbitrary $A, A' \in \mathcal{F}$, we can easily construct $\sigma \in C(n)$ such that $A, A' \in S_\sigma(\mathcal{F})$. Since either $\sigma \in C_0$ or $\sigma \in C_k$, this implies that $A, A' \in \mathcal{F}^*$ or $A, A' \in (\mathcal{F} \setminus \mathcal{F}^*)$. Since A and A' were arbitrary, we get that either $\mathcal{F} = \mathcal{F}^*$ or $\mathcal{F} = (\mathcal{F} \setminus \mathcal{F}^*)$. If we assume the former then $|\mathcal{F}| = |\mathcal{F}^*| = \binom{n}{k}$ in which case $\mathcal{F} = \binom{[n]}{k}$. For the latter, we get that $|\mathcal{F}| = |\mathcal{F} \setminus \mathcal{F}^*| = \binom{n-1}{k-1}$ and $\mathcal{F}$ is pairwise intersecting, in which case Theorem 1.1.5 tells us that $\mathcal{F}$ is a star. This completes the proof. □

2.4 Clusters

We will now prove Theorem 2.1.2 by induction on d. Specifically, we will prove something slightly stronger in order to have more powerful inductive tools.

Theorem 2.4.1. *Let $3 \leq d + 1 \leq k \leq n/2$, and suppose $\mathcal{F}^* \subset \mathcal{F} \subset \binom{[n]}{k}$ are such that any d-cluster in $\mathcal{F}$ is contained in $\mathcal{F}^*$. Then*

$$|\mathcal{F}| \leq \binom{n-1}{k-1} + \frac{n-k}{n}|\mathcal{F}^*|$$

with equality only if $\mathcal{F}$ is a maximum-sized star.

We note that the $d = 2$ case of this theorem is equivalent to Theorem 2.3.1, since 1-clusters are simply two disjoint sets. The following result, furthermore, will give us the tools to induct on d. It is a stronger version of a proposition from [25] that was later stated more clearly in [26].

Proposition 2.4.2. *Let $3 \leq d + 1 \leq k$ and $n \geq (d + 1)k/d$. Furthermore, suppose $\mathcal{F}^* \subset \mathcal{F} \subset \binom{[n]}{k}$ has the property that any d-cluster in $\mathcal{F}$ is contained in $\mathcal{F}^*$. Then, if $\{D_1, \ldots, D_d\} \subset \overline{\bigtriangledown_{\mathcal{F}}(x)}$ is a $(d-1)$ cluster, it follows that for each $i \in [d]$, we have $|\overline{\bigtriangledown_{\mathcal{F}}(D_i)}| = 1$ or $D_i \in \overline{\bigtriangledown_{\mathcal{F}^*}(x)}$.*

Proof. Suppose for the sake of contradiction that there exists $i_0 \in [d]$ such that $|\overline{\triangledown_{\mathcal{F}}(D_{i_0})}| \geq 2$ and $D_{i_0} \notin \overline{\triangledown_{\mathcal{F}^*}(x)}$. Then, we know both that $(D_{i_0} \cup \{x\}) \in (\mathcal{F} \setminus \mathcal{F}^*)$ and that $(D_{i_0} \cup \{y\}) \in \mathcal{F}$ for some $y \in [n] \setminus \{x\}$. Then

$$\begin{aligned} |\big(\bigcup_i (D_i \cup \{x\}\big) \cup (D_{i_0} \cup \{y\})| &\leq |D_1 \cup \cdots \cup D_d| + |\{x, y\}| \\ &\leq 2(k-1) + 2 = 2k, \end{aligned}$$

and since, $x, y \notin D_{i_0}$, we see

$$(D_1 \cup \{x\}) \cap \cdots \cap (D_d \cup \{x\}) \cap (D_{i_0} \cup \{y\}) \subset D_1 \cap \cdots \cap D_d = \emptyset.$$

Furthermore, since $x \notin D_{i_0}$, we know that $(D_{i_0} \cup \{y\}) \neq (D_j \cup \{x\})$ for any $j \in [d]$. Thus, $\mathcal{F}$ contains a d-cluster not contained entirely in $\mathcal{F}^*$, which is a contradiction.

□

The following proposition, furthermore, will assist us in a proof of Theorem 2.4.1.

Proposition 2.4.3. *Let $2 \leq k < n$ and $\mathcal{F} \subset \binom{[n]}{k}$. Then, the following hold:*

(i) $\sum_{x \in [n]} |\overline{\triangledown_{\mathcal{F}}(x)}| = k|\mathcal{F}|$,

(ii) $|\{D \in \binom{[n]}{k-1} \ : \ |\overline{\triangledown_{\mathcal{F}}(D)}| = 1\}| \leq \frac{n\binom{n-1}{k-1} - k|\mathcal{F}|}{n-k}$.

Proof. The proof of (i) is straightforward from the definitions, so we will focus on (ii). First, we denote by $\mathcal{F}^C$ the complement of $\mathcal{F}$ in $\binom{[n]}{k}$, and observe that

$$\sum_{x \in [n]} (|\overline{\triangledown_{\mathcal{F}}(x)}| + |\overline{\triangledown_{\mathcal{F}^C}(x)}|) = n\binom{n-1}{k-1}.$$

Using this along with (i), we see

$$\begin{aligned}
\left|\left\{D \in \binom{[n]}{k-1} : |\overline{\triangledown_{\mathcal{F}}(D)}| = 1\right\}\right| &= \left|\left\{D \in \binom{[n]}{k-1} : |\overline{\triangledown_{\mathcal{F}^C}(D)}| = n-k\right\}\right| \\
&\leq \frac{\sum_{x\in[n]} |\overline{\triangledown_{\mathcal{F}^C}(x)}|}{n-k} \\
&= \frac{n\binom{n-1}{k-1} - \sum_{x\in[n]} |\overline{\triangledown_{\mathcal{F}}(x)}|}{n-k} \\
&= \frac{n\binom{n-1}{k-1} - k|\mathcal{F}|}{n-k}.
\end{aligned}$$

□

We now begin with the proof of our main result.

Proof of Theorem 2.4.1. Let $\mathcal{F}^* \subset \mathcal{F}$ be as described. We note as before that the $d = 1$ case appears already in the literature, so we suppose $d \geq 2$. For any $x \in [n]$, we let

$$\overline{\triangledown^*_{\mathcal{F}}(x)} := \{D \in \overline{\triangledown_{\mathcal{F}}(x)} : |\overline{\triangledown_{\mathcal{F}}(D)}| = 1\} \cup \overline{\triangledown_{\mathcal{F}^*}(x)}.$$

Observe first that the sets $\{D \in \overline{\triangledown_{\mathcal{F}}(x)} : |\overline{\triangledown_{\mathcal{F}}(D)}| = 1\}$ over $x \in [n]$ partition $\{D \in \binom{[n]}{k-1} : |\overline{\triangledown_{\mathcal{F}}(D)}| = 1\}$. Using this and Proposition 2.4.3 yields

$$\begin{aligned}
\sum_{x\in[n]} |\overline{\triangledown^*_{\mathcal{F}}(x)}| &\leq \left|\left\{D \in \binom{[n]}{k-1} : |\overline{\triangledown_{\mathcal{F}}(D)}| = 1\right\}\right| + \sum_{x\in[n]} |\overline{\triangledown_{\mathcal{F}^*}(x)}| \\
&\leq \frac{n\binom{n-1}{k-1} - k|\mathcal{F}|}{n-k} + k|\mathcal{F}^*|. \qquad (2.3)
\end{aligned}$$

Furthermore, by Proposition 2.4.2, we know that any $(d-1)$-cluster in $\overline{\triangledown_{\mathcal{F}}(x)}$ is contained in $\overline{\triangledown^*_{\mathcal{F}}(x)}$. Since $\overline{\triangledown_{\mathcal{F}}(x)} \subset \binom{[n]\setminus\{x\}}{k-1}$ and $d \leq (k-1) < (n-1)/2$, we may apply induction on d to get

$$\binom{n-1}{k-1} \geq |\overline{\triangledown^*_{\mathcal{F}}(x)}| + \frac{n-1}{k-1}|\overline{\triangledown_{\mathcal{F}}(x)} \setminus \overline{\triangledown^*_{\mathcal{F}}(x)}| = \frac{n-1}{k-1}|\overline{\triangledown_{\mathcal{F}}(x)}| - \frac{n-k}{k-1}|\overline{\triangledown^*_{\mathcal{F}}(x)}|. \tag{2.4}$$

Then, summing over all $x \in [n]$ and using Proposition 2.4.3 in combination with (2.3), we see

$$\begin{aligned} n\binom{n-1}{k-1} &\geq \frac{n-1}{k-1}\sum_{x\in[n]}|\overline{\triangledown_{\mathcal{F}}(x)}| - \frac{n-k}{k-1}\sum_{x\in[n]}|\overline{\triangledown^*_{\mathcal{F}}(x)}| \\ &\geq \frac{n-1}{k-1}k|\mathcal{F}| - \frac{n-k}{k-1}k|\mathcal{F}^*| - \left(\frac{n-k}{k-1}\right)\frac{n\binom{n-1}{k-1} - k|\mathcal{F}|}{n-k} \\ &= \frac{nk}{k-1}|\mathcal{F}| - \frac{(n-k)k}{k-1}|\mathcal{F}^*| - \frac{n\binom{n-1}{k-1}}{k-1}, \end{aligned}$$

and therefore

$$\frac{nk}{k-1}\binom{n-1}{k-1} \geq \frac{nk}{k-1}|\mathcal{F}| - \frac{(n-k)k}{k-1}|\mathcal{F}^*|.$$

Finally, multiplying both sides by $\frac{k-1}{k^2}$ gives us

$$\binom{n}{k} \geq \frac{n}{k}|\mathcal{F}| - \frac{n-k}{k}|\mathcal{F}^*| = |\mathcal{F}^*| + \frac{n}{k}|\mathcal{F} \setminus \mathcal{F}^*|.$$

Suppose now that we have equality - that is, that $|\mathcal{F}^*| + (\frac{n}{k})|\mathcal{F} \setminus \mathcal{F}^*| = \binom{n}{k}$. This implies two things. First, we get that $|\mathcal{F}| \geq \binom{n-1}{k-1}$ with $|\mathcal{F}| = \binom{n-1}{k-1}$ only if $\mathcal{F}^* = \emptyset$. Additionally, we must have equality in (2.4) for all $x \in [n]$, that is,

$$|\overline{\triangledown^*_{\mathcal{F}}(x)}| + \frac{n-1}{k-1}|\overline{\triangledown_{\mathcal{F}}(x)} \setminus \overline{\triangledown^*_{\mathcal{F}}(x)}| = \binom{n-1}{k-1}.$$

Furthermore, since $(n-1) > 2(k-1)$, we get by induction that either

$\overline{\nabla^*_{\mathcal{F}}(x)} = \overline{\nabla_{\mathcal{F}}(x)} = \binom{[n]\setminus\{x\}}{k-1}$ or $\overline{\nabla_{\mathcal{F}}(x)}$ is a maximum sized star and $\overline{\nabla^*_{\mathcal{F}}(x)} = \emptyset$. Suppose for the sake of contradiction that the latter is true for all $x \in [n]$. Then $|\overline{\nabla_{\mathcal{F}}(x)}| = \binom{n-2}{k-2}$ for all $x \in [n]$, and using proposition 2.4.3 yields

$$|\mathcal{F}| = \frac{\sum_{x\in[n]} |\overline{\nabla_{\mathcal{F}}(x)}|}{k} = \frac{n}{k}\binom{n-2}{k-2} < \frac{n-1}{k-1}\binom{n-2}{k-2} = \binom{n-1}{k-1},$$

which is a contradiction. Thus, there exists $x_0 \in [n]$ such that $\overline{\nabla_{\mathcal{F}}(x_0)} = \binom{[n]\setminus\{x_0\}}{k-1}$, and therefore $\mathcal{F}$ contains a maximum-sized star centered at x_0. If $|\mathcal{F}| = \binom{n-1}{k-1}$ this implies that $\mathcal{F}$ is, itself, a maximum size star centered at x_0 and that $\mathcal{F}^* = \emptyset$. Now, suppose $|\mathcal{F}| > \binom{n-1}{k-1}$ and take $B, B' \in \mathcal{F}$ such that exactly one of B, B' contains x_0 (note that there must exist at least one element of $\mathcal{F}$ not containing x_0). Furthermore, take $Z \subseteq [n]$ such that $|Z| = 2k$ and $(B \cup B') \subset Z$. Then, since $|B \cap B'| \leq k-1$ we know that $|Z \setminus \{x_0\} \setminus (B \cap B')| \geq k$ and thus there exist distinct

$$D_1, \dots, D_{d-1} \in \binom{Z \setminus \{x_0\} \setminus (B \cap B')}{k-1}$$

such that $(D_i \cup \{x_0\}) \neq B, B'$ for all $i \in [d-1]$. Furthermore, since $\mathcal{F}$ contains a maximum-sized star centered at x_0, we get $(D_i \cup \{x_0\}) \in \mathcal{F}$ for all $i \in [d-1]$ and

$$|B \cup B' \cup (D_1 \cup \{x_0\}) \cup \cdots \cup (D_{d-1} \cup \{x_0\})| \leq |Z| = 2k.$$

Additionally, because x_0 is not in one of B or B', we see

$$B \cap B' \cap (D_1 \cup \{x_0\}) \cap \cdots \cap (D_{d-1} \cup \{x_0\}) = \emptyset.$$

Thus, every element of $\mathcal{F}$ is part of a d-cluster. Since all d-clusters in $\mathcal{F}$ are contained in $\mathcal{F}^*$, we get that $\mathcal{F} = \mathcal{F}^*$ and thus that $|\mathcal{F}| = \binom{n}{k}$ and $\mathcal{F} = \binom{[n]}{k}$. This completes the proof. □

2.5 Clusters and Simplices

In this section we will prove Theorem 2.1.3, but we will first need some intermediary results. The first shows us that if we have $A, B \in \mathcal{F}$ such that $A \cap B$ satisfies certain size and removability conditions, then $\mathcal{F}$ must contain a d-simplex-cluster.

Lemma 2.5.1. *Suppose $d+1 \leq k$ and $n \geq 2k-d$, and let $\mathcal{F} \subseteq \binom{[n]}{k}$. Then, if there exist $A, B \in \mathcal{F}$ such that $A \cap B \in \binom{A \setminus \alpha^1_{\mathcal{F}}(A)}{d} \setminus \binom{\alpha^2_{\mathcal{F}}(A)}{d}$, then $\mathcal{F}$ must contain a d-simplex-cluster.*

Proof. Let A, B be as described, with $A \cap B = \{x_1, \ldots, x_d\}$, and suppose without loss of generality that $x_d \in \alpha^i_{\mathcal{F}}(A)$ for some $i \geq 3$. Then, for all $1 \leq j \leq d$, since $x_j \in \alpha^i_{\mathcal{F}}(A)$ for $i \geq 2$, there exists $B_j \in \mathcal{F}$ such that $A \cap B_j = A \setminus \{x_j\}$. Note at this point that $B_1, \ldots, B_d$ may have an element (at most one) of intersection outside of A. However, if this is the case, since $x_d \in \alpha^i_{\mathcal{F}}(A)$ for $i \geq 3$, we can re-choose B_d such that the $B_1, \ldots, B_d$ have empty intersection outside of A. We claim that $B, B_1, \ldots, B_d$ is a d-simplex-cluster. Verifying first the intersection condition gives

$$B \cap B_1 \cap \cdots \cap B_d = (A \cap B) \cap B_1 \cap \cdots \cap B_d = \emptyset,$$

and furthermore

$$|B \cup B_1 \cup \cdots \cup B_d| \leq |A \cup B| + |B_1 \setminus A| + \cdots + |B_d \setminus A| \leq (2k-d) + d = 2k.$$

Finally, we see that $B_1 \cap \cdots \cap B_d = A \setminus B \neq \emptyset$ and that $x_j \in \left(B \cap \left(\bigcap_{j \neq i} B_j\right)\right)$ for all $1 \leq j \leq d$. Thus, $B, B_1, \ldots, B_d$ is a d-simplex-cluster, completing the proof. □

Thus, our task reduces in some sense to proving that $\alpha^1_{\mathcal{F}}(A)$ and $\alpha^2_{\mathcal{F}}(A)$ are small for most $A \in \mathcal{F}$. The following lemma will be used to show this.

Lemma 2.5.2. *Suppose $k < n$ and that $\mathcal{F} \subseteq \binom{[n]}{k}$ is such that $|\mathcal{F}| \geq \binom{n-1}{k-1}$. Then*

$$\sum_{A \in \mathcal{F}} |\alpha^1_{\mathcal{F}}(A)| + \frac{n-k-1}{2(n-k)}|\alpha^2_{\mathcal{F}}(A)| \leq \binom{n-1}{k-1}.$$

Proof. To start, we define $\mathcal{F}^C := \binom{[n]}{k} \setminus \mathcal{F}$, and observe that

$$|\mathcal{F}|k + |\mathcal{F}^C|k = \binom{n}{k}k = n\binom{n-1}{k-1}.$$

Using our assumption that $|\mathcal{F}| \geq \binom{n-1}{k-1}$, we see that

$$\begin{aligned}(n-k)\binom{n-1}{k-1} &\geq k|\mathcal{F}^C| \\ &= \sum_{A \in \mathcal{F}^C} \sum_{1 \leq i \leq (n-k+1)} |\alpha^i_{\mathcal{F}^C}(A)| \\ &\geq \sum_{A \in \mathcal{F}^C} |\alpha^{n-k}_{\mathcal{F}^C}(A)| + |\alpha^{n-k-1}_{\mathcal{F}^C}(A)| \\ &= \sum_{A \in \mathcal{F}} (n-k)|\alpha^1_{\mathcal{F}}(A)| + \frac{n-k-1}{2}|\alpha^2_{\mathcal{F}}(A)|,\end{aligned}$$

which is the desired result. □

Having shown that $\alpha^1_{\mathcal{F}}(A)$ and $\alpha^2_{\mathcal{F}}(A)$ are small for most $A \in \mathcal{F}$, we will want to use this in conjunction with Lemma 2.5.1. However, Lemma 2.5.1 is a statement about d-subsets of A, while Lemma 2.5.2 is about single elements of A. To bridge the gap between these two results, we use the following counting lemma.

Lemma 2.5.3. *Suppose $d + 1 \leq k < n$ and $\mathcal{F} \subseteq \binom{[n]}{k}$. Then, we have*

$$\sum_{A \in \mathcal{F}} \left(\binom{k}{d} - \binom{|A \setminus \alpha^1_{\mathcal{F}}(A)|}{d} + \binom{|\alpha^2_{\mathcal{F}}(A)|}{d} \right) \leq \binom{k-1}{d-1} \sum_{A \in \mathcal{F}} \left(|\alpha^1_{\mathcal{F}}(A)| + \frac{|\alpha^2_{\mathcal{F}}(A)|}{d} \right).$$

Proof. We use here the fact that if $m_1, m_2, \ell \in \mathbb{N}$, then $\binom{m_1}{\ell} - \binom{m_2}{\ell} = \binom{m_1-1}{\ell-1} +$

$\binom{m_1-2}{\ell-1}+\cdots+\binom{m_2}{\ell-1}$, as well as the fact that $|\alpha^1_{\mathcal{F}}(A)|, |\alpha^2_{\mathcal{F}}(A)| \leq k$ for all $A \in \mathcal{F}$. This yields

$$\begin{aligned}\sum_{A\in\mathcal{F}}\left(\binom{k}{d}-\binom{|A\setminus\alpha^1_{\mathcal{F}}(A)|}{d}+\binom{|\alpha^2_{\mathcal{F}}(A)|}{d}\right) &\leq \sum_{A\in\mathcal{F}}\left(\sum_{i=1}^{|\alpha^1_{\mathcal{F}}(A)|}\binom{k-i}{d-1}+\frac{|\alpha^2_{\mathcal{F}}(A)|}{k}\binom{k}{d}\right)\\ &\leq \sum_{A\in\mathcal{F}}\left(|\alpha^1_{\mathcal{F}}(A)|\binom{k-1}{d-1}+\frac{|\alpha^2_{\mathcal{F}}(A)|}{d}\binom{k-1}{d-1}\right),\end{aligned}$$

thus completing the proof. □

We may now proceed with the proof of our main result.

Proof of Theorem 2.1.3. Let $|\mathcal{F}| \geq \binom{n-1}{k-1}$ and suppose $\mathcal{F}$ contains no d-simplex-cluster. Then, for any $D \in \binom{[n]}{d}$, we define the following subset of $\bigtriangledown_{\mathcal{F}}(D)$,

$$\bigtriangledown^*_{\mathcal{F}}(D) := \{A \in \bigtriangledown_{\mathcal{F}}(D) : D \cap \alpha^1_{\mathcal{F}}(A) = \emptyset \text{ and } D \nsubseteq \alpha^2_{\mathcal{F}}(A)\},$$

with $\overline{\bigtriangledown^*_{\mathcal{F}}(D)}$ defined the usual way. Now, suppose we have $A \in \bigtriangledown^*_{\mathcal{F}}(D)$ and $A' \in \bigtriangledown_{\mathcal{F}}(D)$, and observe that by Lemma 2.5.1 and the fact that $\mathcal{F}$ contains no d-simplex-cluster, we have $(A_1 \setminus D) \cap (A_2 \setminus D) \neq \emptyset$. Thus, we may apply Theorem 2.3.1 with $\overline{\bigtriangledown_{\mathcal{F}}(D)}$ as $\mathcal{F}$ and $\overline{\bigtriangledown^*_{\mathcal{F}}(D)}$ as $\mathcal{F}^*$ to get

$$|\overline{\bigtriangledown_{\mathcal{F}}(D)}| \leq \binom{n-d-1}{k-d-1} + \frac{n-k}{n-d}|\overline{\bigtriangledown^*_{\mathcal{F}}(D)}|.$$

Summing over all $D \in \binom{[n]}{d}$, and using Lemmas 2.5.2 and 2.5.3, we obtain

$$
\begin{aligned}
|\mathcal{F}|\binom{k}{d} &= \sum_{D\in\binom{[n]}{d}} |\bigtriangledown_{\mathcal{F}}(D)| \\
&\leq \binom{n-d-1}{k-d-1}\binom{n}{d} + \frac{n-k}{n-d}\sum_{D\in\binom{[n]}{d}} |\bigtriangledown^{*}_{\mathcal{F}}(D)| \qquad (2.5)\\
&= \binom{n-d-1}{k-d-1}\binom{n}{d} + \frac{n-k}{n-d}\sum_{A\in\mathcal{F}}\left(\binom{k}{d} - \binom{|A\setminus\alpha^{1}_{\mathcal{F}}(A)|}{d} + \binom{|\alpha^{2}_{\mathcal{F}}(A)|}{d}\right)\\
&\leq \binom{n-1}{k-1}\binom{k}{d}\frac{n(k-d)}{k(n-d)} + \frac{n-k}{n-d}\binom{k-1}{d-1}\binom{n-1}{k-1} \qquad (2.6)\\
&= \binom{n-1}{k-1}\binom{k}{d}\left(\frac{n(k-d)}{k(n-d)} + \frac{d(n-k)}{k(n-d)}\right)\\
&= \binom{n-1}{k-1}\binom{k}{d},
\end{aligned}
$$

which is our desired inequality. Note that in (2.6) we have also used that $d \geq 3$ and $n \geq 2k-d+2 \geq k+3$. Now, suppose that we have equality, and in particular that we have equality in (2.5). We wish to show that $\mathcal{F}$ is a star. To start, for every $1 \leq \ell \leq k$, we define $\mathcal{G}_\ell \subseteq \binom{[n]}{\ell}$ as follows

$$
\mathcal{G}_\ell := \{D \in \binom{[n]}{\ell} : \bigtriangledown_{\mathcal{F}}(D) = \bigtriangledown_{\binom{[n]}{k}}(D)\}.
$$

The proof will proceed as follows: we will start by showing that $|\mathcal{G}_d| \geq \binom{n-1}{d-1}$ and use this to show that $|\mathcal{G}_{d+1}| \geq \binom{n-1}{d}$. Then, we will show that $\mathcal{G}_{d+1}$ is d-simplex-free, and use this to show that it is a star. This will show by extension that $\mathcal{F}$ is a star.

We start by showing that $|\mathcal{G}_d| \geq \binom{n-1}{d-1}$. To see this, observe that since $n-d > 2(k-d)$, equality in (2.5) implies that, for any $D \in \binom{[n]}{d}$, we have either that $\bigtriangledown_{\mathcal{F}}(D)$ is a maximum-sized $(d+1)$-star or all of $\bigtriangledown_{\binom{[n]}{k}}(D)$. In particular, this implies that $|\bigtriangledown_{\mathcal{F}}(D)| = \binom{n-d}{k-d}$ for all $D \in \mathcal{G}_d$ and $|\bigtriangledown_{\mathcal{F}}(D)| =$

$\binom{n-d-1}{k-d-1}$ for all $D \in \binom{n}{d} \setminus \mathcal{G}_d$. Suppose for the sake of contradiction that $|\mathcal{G}_d| < \binom{n-1}{d-1}$. Then, we get

$$\begin{aligned} |\mathcal{F}|\binom{k}{d} &= |\mathcal{G}_d|\binom{n-d}{k-d} + \left(\binom{n}{d} - |\mathcal{G}_d|\right)\binom{n-d-1}{k-d-1} \\ &< \binom{n-1}{d-1}\binom{n-d}{k-d} + \binom{n-1}{d}\binom{n-d-1}{k-d-1} \\ &= \binom{n-1}{k-1}\binom{k}{d}, \end{aligned}$$

which is a contradiction, so $|\mathcal{G}_d| \geq \binom{n-1}{d-1}$. Now, we will show that $|\mathcal{G}_{d+1}| \geq \binom{n-1}{d}$ by a double-counting argument. Observe that if $D \in \mathcal{G}_d$ and $x \in [n] \setminus D$, then $(D \cup \{x\}) \in \mathcal{G}_{d+1}$. Furthermore, as noted before, for every $D \in \binom{[n]}{d} \setminus \mathcal{G}_d$, we have that $\bigtriangledown_{\mathcal{F}}(D)$ is a maximum-sized $(d+1)$-star. Thus, in this case there exists exactly one $x \in [n] \setminus D$ such that $(D \cup \{x\}) \in \mathcal{G}_{d+1}$. Finally, any element of $\mathcal{G}_{d+1}$ will be counted in this way precisely $d+1$ times, giving us

$$\begin{aligned} |\mathcal{G}_{d+1}| &\geq \frac{|\mathcal{G}_d|(n-d) + \left(\binom{n}{d} - |\mathcal{G}_d|\right)}{d+1} \\ &\geq \frac{\binom{n-1}{d-1}(n-d) + \binom{n}{d} - \binom{n-1}{d-1}}{d+1} \\ &= \binom{n-1}{d}. \end{aligned}$$

We show next that $\mathcal{G}_{d+1}$ must contain no d-simplex. To see this, suppose for the sake of contradiction that $\mathcal{G}_{d+1}$ contains a d-simplex $\{D_1, \ldots, D_{d+1}\}$. We observe first that $\{D_1, \ldots, D_{d+1}\}$ must in fact also be a d-cluster. To see the union condition, note that there must be $\{x_1, \ldots, x_{d+1}\} \subset [n]$ such that $x_j \in D_i$ for all $i \neq j$. By extension we see $|D_i \setminus \{x_1, \ldots, x_{d+1}\}| = 1$, and it follows easily that $|\bigcup_i D_i| \leq 2(d+1)$. Next, we choose (not necessarily distinct) $(k-d-1)$-sets $E_1, \ldots, E_{d+1} \subseteq [n] \setminus (\bigcup_i D_i)$ such that $\bigcap_i E_i = \emptyset$ and $|\bigcup_i E_i| \leq 2(k-d-1)$. Then, because $D_i \in \mathcal{G}_{d+1}$ for all $1 \leq i \leq d+1$, it follows that $(D_i \cup E_i) \in \mathcal{F}$. We claim that $(D_1 \cup E_1), \ldots, (D_{d+1} \cup E_{d+1})$ is a

d-simplex-cluster. To verify this, we check first the union condition

$$\left|\bigcup_i (D_i \cup E_i)\right| = \left|\bigcup_i D_i\right| + \left|\bigcup_i E_i\right| \leq 2(d+1) + 2(k-d-1) = 2k,$$

and the first intersection condition

$$\bigcap_i (D_i \cup E_i) = \left(\bigcap_i D_i\right) \cup \left(\bigcap_i E_i\right) = \emptyset,$$

and finally the second intersection condition

$$\bigcap_{j \neq i} (D_i \cup E_i) = \left(\bigcap_{j \neq i} D_i\right) \cup \left(\bigcap_{j \neq i} E_i\right) \supseteq \left(\bigcap_{j \neq i} D_i\right) \neq \emptyset.$$

However, this contradicts our assumption that $\mathcal{F}$ is d-simplex-cluster-free, so $\mathcal{G}_{d+1}$ must be d-simplex-free. However, the $d+1 = k$ case of Conjecture 1.3.4 was resolved by Chvátal in [5]. Since $|\mathcal{G}_{d+1}| \geq \binom{n-1}{d}$, this implies that $\mathcal{G}_{d+1}$ is a star centered at some $x \in [n]$. We now count the number of elements of $\mathcal{F}$ that contain x by another double counting argument. For every $D \in \mathcal{G}_{d+1}$ there will be $\binom{n-d-1}{k-d-1}$ elements of $\mathcal{F}$ that contain it. Furthermore, for any $A \in \mathcal{F}$ that contains x, it will have $\binom{k-1}{d}$ subsets of size $d+1$ that contain x. From this, we obtain

$$| \bigtriangledown_{\mathcal{F}} (x)| \geq \frac{|\mathcal{G}_{d+1}|\binom{n-d-1}{k-d-1}}{\binom{k-1}{d}} \geq \frac{\binom{n-1}{d}\binom{n-d-1}{k-d-1}}{\binom{k-1}{d}} = \binom{n-1}{k-1}.$$

Thus, $\mathcal{F}$ is a star centered at x. This completes the proof. □

Bibliography

[1] R. Ahlswede and L. Khachatrian, *The complete intersection theorem for systems of finite sets*, European J. Combin. **18** (1997), 125–136.

[2] J. Balogh, R. Morris, and Wojciech Samotij, *The method of hypergraph containers*, Proceedings of the International Congress of Mathematics, Rio de Janeiro, 2018 (2018), 3045–3078.

[3] P. Borg, *A short proof of a cross-intersection theorem of Hilton*, Discrete Math. **309** (2009), 4750–4753.

[4] P. Borg and I. Leader, *Multiple cross-intersecting families of signed sets*, J. Combin. Theory Ser. A. **117** (2010), 583–588.

[5] V. Chvátal, *An extremal set-intersection theorem*, J. London Math. Soc. **9** (1974), 355–359.

[6] G. Currier, *New results on simplex-clusters in set systems*, Combinatorica (2020).

[7] ———, *On the d-cluster generalization of Erdős-Ko-Rado*, J. Combin. Theory Ser. A. **182** (2021).

[8] D. E. Daykin, *Erdős-Ko-Rado from Kruskal-Katona*, J. Combin. Theory Ser. A. **17** (1974), no. 2, 254–255.

[9] G. A. Dirac, *Some theorems on abstract graphs*, Proceedings of the London Mathematical Society **2** (1952), 69–81.

[10] P. Erdős, *A problem on independent r-tuples*, Annales Univesritatis Scientarum Budapestinensis **8** (1965), 93–95.

[11] ———, *Topics in combinatorial analysis*, Proceedings of the Second Louisiana Conference on Combinatorics, Graph Theory and Computing (1971), 2–20.

[12] P. Erdős, H. Ko, and R. Rado, *Intersection theorems for systems of finite sets*, Quart. J. Math Oxford Ser. **12** (1961), 313–320.

[13] P. Frankl, *On Sperner families satisfying an additional condition*, J. Combin. Theory Ser. A. **20** (1976), 1–11.

[14] ———, *On a problem of Chvátal and Erdős on hypergraphs containing no generalized simplex*, J. Combin. Theory Ser. A. **30** (1981), 169–182.

[15] P. Frankl and Z. Füredi, *A new generalization of the Erdős-Ko-Rado theorem*, Combinatorica **3** (1983), 341–349.

[16] ———, *Forbidding just one intersection*, J. Combin. Theory Ser. A. **39** (1985), 160–176.

[17] ———, *Exact solution of some Turán-type problems*, J. Combin. Theory Ser. A. **45** (1987), 226–262.

[18] P. Frankl and A. Kupavskii, *Erdős-Ko-Rado theorem for* $\{0, \pm 1\}$*-vectors*, J. Combin. Theory Ser. A. **155** (2018), 157–179.

[19] C. Godsil and K. Meagher, *Erdős-Ko-Rado: Algebraic approaches*, Cambridge University Press, 2015.

[20] W. T. Gowers, *Hypergraph regularity and and the multidimensional Szemerédi theorem*, Ann. of Math. (2007), 897–946.

[21] G. Katona, *A simple proof of the Erdős-Chao Ko-Rado theorem*, J. Combin. Theory Ser. A. **13** (1972), 183–184.

[22] P. Keevash and D. Mubayi, *Set systems without a simplex or cluster*, Combinatorica **30** (2010), 175–200.

[23] L. Lovász, *Combinatorial problems and exercises*, vol. 2, AMS Chelsea, 2007.

[24] D. Mubayi, *Erdős-Ko-Rado for three sets*, J. Combin. Theory Ser. A. **113** (2006), 547–550.

[25] ——— , *An intersection theorem for four sets*, Adv. in Mathematics **215** (2007), 601–615.

[26] D. Mubayi and R. Ramadurai, *Set systems with union and intersection constraints*, J. Combin. Theory Ser. B. **99** (2009), 639–642.

[27] D. Mubayi and J. Verstraëte, *Proof of a conjecture of Erdős on triangles in set systems*, Combinatorica **25** (2005), 599–614.

[28] V. Rödl and J. Skokan, *Regularity lemma for k-uniform hypergraphs*, Random Structures and Algorithms (2004), 1–42.

[29] P. Turán, *On an extremal problem in graph theory*, Matematikai és Fizikai Lapok **48** (1941), 436–452.