Turán Problems for Hypergraphs

By

Nathan W Lemons

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

in the

DEPARTMENT OF MATHEMATICS AND ITS APPLICATIONS

of the

CENTRAL EUROPEAN UNIVERSITY

BUDAPEST, HUNGARY

Approved:

Committee in Charge, 2008

Contents

Abstract

We consider two different Turán-type problems for hypergraphs. The first concerns uniform and non-uniform hypergraphs avoiding cycles of a given length. Here we use the loosest definition of a cycle (due to Berge). We are able to bound the number of edges of l-uniform hypergraphs containing no cycle of length $2k+1$ by $O(n^{(k+1)/k})$ if $l \geq 3$ and n is the number of vertices of the hypergraph. We give the same bound to l-uniform hypergraphs avoiding a cycle of length $2k$. These orders of magnitudes are shown to be sharp when $k = 2, 3$ or 5. We also consider the problem for non-uniform hypergraphs. Here we are able to bound the total size of the hypergraph ($\sum_{h \in \mathcal{E}(\mathcal{H})} |h|$) by $O(n^{1+1/k})$ for hypergraphs $\mathcal{H}$ if either $\mathcal{H}$ contains no cycle of length $2k$ or $\mathcal{H}$ contains no cycle of length $2k+1$.

The second problem is a perturbation of the famous Erdős-Ko-Rado Theorem. We find the largest possible unbalance of k-uniform hypergraphs whose edges have pairwise non-trivial intersections. The unbalance of such a hypergraph is defined as the size (number of edges) of the hypergraphs minus the size of the largest degree in the hypergraph.

To introduce these two problems, we give a short introduction to the history of Turán type problems paying close attention to those problems and results which motivate and or relate to our results.

Acknowledgments

The author would like to thank Ervin Győri for many helpful discussions.

1 Introduction

The first Turán type result (Mantel's Theorem) appeared almost exactly 100 years ago. However, it was not until the second half of the 20th century, following the publication of Turán's famous paper, that Turán type problems became a subfield of their own within the larger field of extremal combinatorics. At this time, this subfield has grown so large that we are forced to restrict our attention to a small segment of the possible problems, results and techniques.

Turán problems are questions of the following sort. Let $\mathcal{F}$ be a family of graphs. How many edges can a graph have if the graph contains no member of $\mathcal{F}$ as a subgraph? (We will also ask this question in the context of hypergraphs.) As we shall see, these easily stated problems are often quite difficult to solve; the theory is quite deep and the methods needed to solve these problems can be varied and complex. In what follows, we present a short history of the most notable Turán-type results including some of the important proofs. Of course in a short paper such as this, we are forced to be selective; the material represented here reflects the author's bias and is by no means a complete exposition on the subject of Turán type problems in combinatorics. In deciding which proofs to include (that are not our own) we tried to choose proofs that were elegant, simple, and above all important either in technique or result to the general field. We have tried to give an idea of the many different approaches and techniques that have been used to solve Turán problems. However, here also we have had to leave out important and nice results. Before we jump into the actual mathematics, we start with a few definitions.

A *graph* G consists of a finite vertex set $V(G)$ and a collection of edges $E(G) \subseteq \binom{V(G)}{2}$ which are subsets of $V(G)$ of size two. If we allow a graph to have multiple copies of an edge, we call it a *multigraph*. We often refer to the number of edges in the graph G as $e(G)$. As usual, the complete graph on n vertices is referred to as K_n while $K_{n_1,n_2,\ldots,n_r}$ refers to the complete r-partite graph with parts of size $n_1, n_2, \ldots, n_r$. The *Turán Graph*, $T_{n,r}$, is the complete r-partite graph on n vertices in which the partite sets are as close in size as possible to each other: if n_i is the size of the i^{th} partite set, then we require $|n_i - (n/r)| < 1$ for all i. The length of a cycle or a path refers to the number of edges in it. Unless otherwise specified, C_k and P_k will refer to a cycle, respectively path, of length k. Often we refer to cycles by their geometric equivalents; a triangle refers to a 3-cycle, a pentagon to a 5-cycle.

A *hypergraph* $\mathcal{H}$ is a generalization of a graph where an edge (or *hyperedge*) can be a subset of any size of the vertex set. If all the edges are of size 2, then the hypergraph is just a graph. Note that our notation distinguishes graphs from hypergraphs; graphs are represented with uppercase letters, hypergraphs with script letters. In general, we call a hypergraph k-*uniform* if all of its edges are of size k. We also sometimes call such a hypergraph a k-*graph*. If the edges of the hypergraph have varying sizes, the hypergraph is called *nonuniform*.

For a graph G, or a collection of graphs $\mathcal{G}$, $\mathrm{ex}(n, G)$ (respectively $\mathrm{ex}(n, \mathcal{G})$) denotes the maximal number of edges a graph on n vertices can have without containing the graph G (or a member of $\mathcal{G}$) as a subgraph. The study of Turán type problems is concerned with determining the values of $\mathrm{ex}(n, G)$ for various graphs G. One can ask similar questions for hypergraphs; given a fixed k-uniform hypergraph, $\mathcal{H}$, we let $\mathrm{ex}_k(n, \mathcal{H})$ denote the maximum number of edges in a k-uniform hypergraph can have without containing $\mathcal{H}$ as a sub-hypergraph. If it is clear from the context that we are talking about k-uniform hypergraphs, we may suppress the k. For a forbidden k-uniform hypergraph, $\mathcal{H}$, the Turán density is defined as

$$\pi(\mathcal{H}) = \lim_{n\to\infty} \frac{\mathrm{ex}_k(n, \mathcal{H})}{\binom{n}{k}} \tag{1.1}$$

The first Turán type result was proved by Mantel[81] in 1907. Mantel showed that if a graph on n vertices has more than $n^2/4$ edges, then it must contain a triangle. This result did not generate much attention until after Turán proved a much more general theorem. Turán's seminal paper can easily be credited for popularizing and indeed starting this field of study; this is why it bears his name. In this paper, Turán found the largest graphs which do not contain a K_k for any fixed k. (These are the Turán graphs.) Interestingly, Turán first approached this problem from a Ramsey theoretical viewpoint. Ramsey's theorem states that for a given k, there is a function $n_0(k)$ such that $\forall n > n_0(k)$, any graph on n vertices which does not contain

k independent vertices must contain a K_k as a subgraph. Turán then asked, if a graph on n vertices contains enough edges, can one insure the existence of a K_k as a subgraph? Of course, Turán not only answered this question in the affirmative, but also found the unique maximal graphs containing no K_k. Here we give Turán's elegant proof[97] which was published in 1941. (See also [98])

1.1 Turán's Theorem

Theorem 1.1 (Turán)**.** *Let G be a graph on n vertices containing no K_k with at least $e(T_{n,k-1})$ edges. Then $G = T_{n,k-1}$.*

Proof. By induction on n, the number of vertices in G. Without loss of generality, suppose G contains a subgraph $H = K_{k-1}$. Write $e(G) = e + f + g$ where e is the number of edges in H, f the number of edges in $G \backslash V(H)$, and g the number of edges joining H to $G \backslash V(H)$. By the induction hypothesis, we know $f \leq e(T_{n-k+1,k-1})$. Also, each vertex not in H can have at most $k-2$ neighbors in H. Thus $g \leq (n-k+1)(k-2)$. As $e = \binom{k-1}{2}$, we conclude $e(G) \leq e(T_{n,k-1})$. Then by assumption, $e(G) = e(T_{n,k-1})$ and we must show that $G = T_{n,k-1}$.

As equality holds, we must have each vertex not in H connected to exactly $k-2$ neighbors of H. Then we can partition the vertices of G (including the vertices of H) into $k-1$ classes depending on which $k-2$ vertices in H they neighbor. Clearly, these classes must by independent; thus G is complete

$(k-1)$-partite. It is easy to check that the largest complete $(k-1)$-partite graph on n vertices is the Turán graph $T_{n,k-1}$. □

The simplicity and elegance of Turán's result inspired further analysis. One can ask similar questions about any graphs - not just complete graphs. Interestingly, it turns out that Turán's result lies much deeper than one might guess at first. In fact, the Turán graphs, $T_{n,p}$, are very similar to the extremal graphs for many other classes of graphs. Here the theory of extremal graphs bifurcates; there are two distinct classes of graphs which one can forbid. If the forbidden graph has chromatic number at least three, then its extremal graph is very close indeed in size and structure to the appropriate Turán graph. In this case, there are some very nice structural theorems that more or less solve the problem. On the other hand, the situation is quite different if the forbidden graph has chromatic number two. In this case, while there are some known bounds, many of the extremal graphs are completely unknown. Following the terminology of Simonovits[93], we call the first type of Turán problem "non-degenerate" and the second, "degenerate."

The difference between these two classes of problems can hardly be overstated. In non-degenerate problems, the extremal graphs all have $\Theta(n^2)$ edges. However, in the degenerate problems, the extremal graphs have $o(n^2)$ edges. Many of the known structural theorems known give quantitative results with an error term of $o(n^2)$, which when applied to degenerate problems, renders such theorems trivial. Thus there is very little general theory concerning the structures of the extremal graphs for degenerate problems. Indeed, most of the results are partial results and some of the most basic problems remain unanswered. However, this is not the case at all for non-degenerate problems. As the theory of non-degenerate problems builds directly upon Turán's results, we examine them first.

1.2 Non-Degenerate Problems

After Turán's Theorem, the second major step the theory of non-degenerate problems is a theorem of Erdős and Stone[38], published in 1946.

Theorem 1.2 (Erdős-Stone). *Let $K_{n,p+1}$ be a regular $(p+1)$-partite graph. Then*

$$ex(n, K_{n,p+1}) = \left(1 - \frac{1}{p}\right)\binom{n}{2} + O(n)$$

Twenty years later, Erdős and Simonovits realized that this had very deep implications for the generalized Turán problem. For a family of graphs $\mathcal{F}$, we define its *subchromatic number* to be one less than the minimum chromatic number over all the graphs in the family. Erdős and Simonovits[39] showed that the Erdős-Stone theorem easily implies the following:

Theorem 1.3 (Erdős-Stone-Simonovits). *If $\mathcal{F}$ is a family of forbidden subgraphs whose subchromatic number is p, then*

$$ex(n, \mathcal{F}) = \left(1 - \frac{1}{p}\right)\binom{n}{2} + o(n^2) \tag{1.2}$$

We note that the Erdős-Stone-Simonovits Theorem is trivial in the degenerate case. For families with subchromatic number one, that is families containing a graph with chromatic number 2, the first term in Equation 1.2 disappears and the theorem only tells us that the extremal graphs have $o(n^2)$ edges. We do not know what the correct order of magnitude is. However, in the non-degenerate case, when the subchromatic number of the family is at least 2, the Erdős-Stone-Simonovits Theorem gives not only the correct order of magnitude for the extremal family size, but also the correct coefficient. In fact, even more is known about the extremal graphs. They are close structurally to the Turán graphs. The following two theorems (developed by Erdős and Simonovits - see [25], [26], and [91]) show this.

Theorem 1.4 (Structure Theorem)**.** *Let $\mathcal{F}$ be a family of forbidden subgraphs with subchromatic number p. If $\{G_n\}$ are the extremal graphs on n vertices for this family, then G_n can be obtained from $T_{n,p}$ by adding or deleting at most $o(n^2)$ edges.*

As one might expect at this point, not only are all the extremal graphs very similar to the Turán graphs, but also, any graph close in size to the extremal graph must also be similar to the Turán graph. This is expressed in the following stronger result:

Theorem 1.5 (Stability Theorem)**.** *Let $\mathcal{F}$ be a family of forbidden subgraphs with subchromatic number p. For $\epsilon > 0$, there is a δ and $n_0(\epsilon)$ such that if G is a graph on n vertices containing no member of $\mathcal{F}$, $n > n_0(\epsilon)$, and*

$e(G) > ex(n, \mathcal{F}) - \delta n^2$, then G can be obtained from $T_{n,p}$ by adding or deleting at most ϵn^2 edges.

The Stability Theorem says that the Turán graphs are not only the unique extremal graphs avoiding a K_k (for given k), but also any other graphs avoiding a K_k with approximately the same number of edges as $T_{n,k-1}$ will have the same structure as $T_{n,k-1}$ with just a few edges changed. In this sense, the structure of the Turán graphs is 'stable' - graphs with close to the same number of edges must have basically the same structure. In some ways, Turán's problem is now completely solved in the non-degenerate case; we know the extremal graphs and even the structure of graphs that are almost extremal. However, this is in no way the end of the theory; there are still more structural questions one can ask about the extremal graphs.

One way to further study the structure of the extremal graphs is to take a common characteristic of the extremal graph and forbid it as well. If the resulting extremal graphs are quite different from the originals, we know the new forbidden trait is an important characteristic of the original extremal graphs. Such problems are called perturbation problems. For instance, two of the most basic characteristics of the Turán graphs $T_{n,k}$ are that they have large independent sets and are k-colorable. We consider both perturbation problems resulting from limiting each of these characteristics from the extremal graph.

To perturb the size of the independent sets in our extremal graphs, we can ask: what is the largest graph on n vertices with no more than $f(n)$

independent vertices which does not contain a K_k? Note that if we perturb the graphs in this way too much, no graph will satisfy our requirements. That is, Ramsey's theorem says that graphs with no K_k must contain large independent sets (for large enough n.) In other words, large independent sets are an integral part of the extremal graphs avoiding a K_k. Thus for our question to be interesting, we should make sure $f(n)$ does not grow too slowly. In this way, we can perhaps quantify the importance of having large independent sets. Let $ex(n, K_k, f)$ denote the size of the largest graph on n vertices having no more than $f(n)$ independent vertices and containing no K_k. Erdős and Sós[36] proved that if f is linear in n, but smaller than n/k (the number of independent vertices in $T_{n,k}$) then $ex(n, K_{k+1}, f)$ is $\Omega(n^2)$ smaller than $ex(n, K_{k+1})$. For $f(n) = o(n)$ they proved the following:

Theorem 1.6 (Erdős-Sós[36]). *There is a constant, $c > 0$, such that if $g(n) = c\sqrt{n} log\ n$ and $g(n) \leq f(n) = o(n)$, then*

$$ex(n, K_{k+1}) \leq ex(n, K_{2k+1}, g) \leq ex(n, K_{2k+1}, f) \leq ex(n, K_{k+1}) + o(n^2)$$

Note that this theorem gives the exact bound (including the constant factor) for a wide range of f. It is interesting that the results are different when the order of the forbidden complete graph is even - see for instance [32]. Next we consider chromatic perturbation; can we have large graphs with chromatic number $(k + 1)$ avoiding a family $\mathcal{F}$ of forbidden subgraphs whose subchromatic number is k? The following theorem by Simonovits[92] answers that question in the negative; if a graph G containing no member of $\mathcal{F}$ has almost the same number of edges as the Turán graph $T_{n,k}$, then it must also have chromatic number k.

Theorem 1.7 (Simonovits)**.** *Let $\mathcal{F}$ be a family of graphs with subchromatic number k whose extremal graph is the Turán graph $T_{n,k}$ for all $n > n_0$. Then there is a constant K such that if a graph G on n vertices contains no member of $\mathcal{F}$ and has chromatic number greater than k, then*

$$e(G) < ex(n, \mathcal{F}) - (n/k) + K$$

We consider another perturbation problem related to chromatic perturbation. Erdős[28] asked the following question: How many edges must one delete to make a triangle free graph bipartite? He conjectured that the blown up pentagon was the extremal graph; that at most $n^2/25$ edges needed to be deleted. (The blown up pentagon is a five-partite graph with parts $V_1, \ldots V_5$ such that all the vertices are connected between parts V_i and V_j iff $|i - j| \in \{1, 4\}$ modulo 5; otherwise there are no edges between V_i and V_j.) In the same paper, Erdős also asked how many pentagons a triangle-free graph can contain. Again he conjectured that the answer was $(n/5)^5$ with the extremal graph being a blown up pentagon. The best known upper bound is approximately $1.03(n/5)^5$ which was proved by Győri[62]. Motivated by these questions, Bollobás and Győri asked the converse question: how many triangles can a pentagon-free graph contain? They proved[10] that such a graph could contain at most $\Omega(n^{3/2})$ triangles. They also showed that 3-uniform hypergraphs with no C_5 had at most $\Omega(n^{3/2})$ edges. As we will see in the next section, this is the order of magnitude of the extremal graph containing no 4-cycle. These problems are closely related. In fact, Bollobás and Győri's constructions are built upon maximal C_4-free graphs. We will return to this question later in Section 1.4.

Having considered some structural properties of extremal graphs, we look at the inverse problem: which graphs can be extremal graphs? More specifically, which sequences of graphs $\{G_n\}$ on n vertices can occur as extremal graphs? Simonovits[91], [92] provides a nice characterization of the forbidden families for which the Turán graphs are the extremal graphs:

Theorem 1.8. *A family $\mathcal{F}$ of forbidden graphs with subchromatic number k, has the Turán graph $T_{n,k}$ as an extremal graph (for large n) iff there exists an $F \in \mathcal{F}$ and an edge $e \in E(F)$ such that $\chi(F - e) = k$. In addition, if the Turán graphs are extremal for infinitely many values of n, then for all n large enough, the Turán graph is the only extremal graph.*

1.3 Degenerate Problems

Having provided a brief overview of the non-degenerate case, we now turn our attention to the degenerate problems. Here, in contrast to the non-degenerate problems, there is very little general theory and indeed many of the most basic questions are unsolved. There are some important partial results, but a break through is needed. It is a testament to the difficult nature of the degenerate problems that the extremal graphs for two of the most basic families of bipartite graphs remain unsolved: the even cycles, and the complete bipartite graphs. On the one hand, we exclude extremely sparse bipartite graphs, on the other, very dense bipartite graphs. In both cases, upper bounds are know which are generally believed to be correct (up to constant factors) but the constructions (or proof of their existence) are missing.

The simplest two-chromatic graph is a path. The extremal numbers for graphs not containing a path of a certain length were first found by Erdős and Gallai[31] in 1959.

Theorem 1.9 (Erdős-Gallai). *$ex(n, P_k) = \frac{1}{2}(k-1)n$, equality holds if k is a divisor of n.*

The extremal graphs in this case, are the disjoint union of complete graphs on k vertices. This result was improved on later by Faudree and Schelp[40]. In Section 2.2.2 we will make use of the following corollary of the Erdős-Gallai Theorem.

Lemma 1.10 (Győri-Lemons[66]). *Let G be a graph on vertex set $\mathcal{V}$. Let $\mathcal{P}$ be a proper coloring of $\mathcal{V}$. Suppose that there is no path on $2l$ vertices with endpoints in different color classes. Then $e(G) \leq 2(l-1)|\mathcal{V}|$.*

Note that if G is given a discrete coloring (every vertex its own color), then Lemma 1.10 is almost the same as the Erdős-Gallai Theorem; however it is off by a factor of two. It is unclear what constant factor in Lemma 1.10 should be; perhaps it is best possible. Interestingly, the situation is very different for paths of even length. If G is a two-colored bipartite graph on n vertices, then it can have as many as $n^2/4$ edges without containing any paths of even length with endpoints in different color classes! (Clearly in a bipartite graph, all even length paths start and end in the same part.) Thus it is unclear what the corresponding statement for even length paths should be. Certainly the proof of Lemma 1.10 relies heavily on the fact that the path is of odd length.

Proof of Lemma 1.10. First note that we can find a bipartition of $\mathcal{V}$ such that each color class sits completely within one partition class and so that at least half the edges are between the two parts, i.e., we can find a bipartite subgraph of G by deleting at most half the edges such that each color class is completely within one part. This is clear because if we assign color classes randomly to either part A or to part B with probability $1/2$, then the expected number of edges between A and B is $(1/2)e(G)$. Thus there must be at least one bipartition of the color classes with at least half the edges of G between the two classes. Let $G^1 = (A, B)$ be such a bipartite subgraph of G. Now any path on $2l$ vertices in G^1 has endpoints both in A and in B; the endpoints are in different color classes. So we can have no such path. The classical theorem of Erdős and Gallai[31] says that $e(G^1) \leq (l-1)|\mathcal{V}|$. We conclude $e(G) \leq 2e(G^1) \leq 2(l-1)|\mathcal{V}|$ as desired. □

Returning to the Erdős-Gallai Theorem, we note that the same bound holds for the extremal graph containing no star $K_{1,k}$. This led Erdős and Sós to conjecture that for every tree on $k+1$ vertices, the extremal graph with no such tree has at most $\frac{1}{2}(k-1)n$ edges. This was proved for large enough trees by Ajtai, Komlós, and Szermerédi[2], [3].

Theorem 1.11. *There is a constant c such that for all $k > c$, and for any tree T on $k+1$ vertices,*

$$ex(n, T) \leq \frac{1}{2}(k-1)n$$

The same bound was proved by Sidorenko[90] for all trees on k vertices with a vertex with at least $(k-2)/2$ neighboring leaves.

The relative completeness of the extremal problem for trees is rare. Very few of the non-degenerate problems have been solved. The rest of this section is devoted to the two most famous classes of bipartite graphs for which we do not know the extremal graphs: the even cycles and the complete bipartite graphs.

As far back as 1938, Erdős[23] found the exact order of magnitude for a graph with no cycle of length four:

$$ex(n, C_4) = \Theta(n^{3/2})$$

This result was partially generalized by Kővári, T. Sós, and Turán[75] who gave an upper bound on the size of complete bipartite graphs. (Note that $C_4 = K_{2,2}$.)

Theorem 1.12. *Suppose $r \leq s$. Then $ex(n, K_{r,s}) \leq \frac{1}{2}(s-1)^{1/2}n^{2-(1/r)} + O(n)$.*

The proof is a nice example of the power of double counting, a common technique in combinatorics, so we present it here.

Proof. Let G be a graph on n vertices with no $K_{r,s}$. Count the stars $K_{1,r}$ in G. Every set of r vertices is in such a star for at most $(s-1)$ different vertices. Thus there are at most $(s-1)\binom{n}{r}$ such stars. However, if $d_1, d_2, \ldots, d_n$ are the degrees of the vertices in G, then the total number of such stars is $\sum_{i=1}^{n}\binom{d_i}{r}$. Applying Jensen's Inequality we get

$$n\binom{2e(G)/n}{r} \leq \sum_{i=1}^{n}\binom{d_i}{r} \leq (s-1)\binom{n}{r}$$

□

Erdős, Rényi, and Sós[35] and (independently) Brown[12] showed that for $r = s = 2$ and infinitely many values of n, Theorem 1.12 is best possible. Namely, for infinitely many n,

$$ex(n, C_4) = \frac{1}{2}n^{3/2} + O(n^) \tag{1.3}$$

This was generalized by Füredi[57], who showed:

Theorem 1.13. *For $t \geq 1$, $ex(n, K_{2,t}) = \frac{1}{2}\sqrt{t-1}n^{3/2} + O(n)$.*

Füredi's result is quite impressive as there are very few asymptotics known for degenerate problems. Indeed the correct orders of magnitude of the extremal graphs are mostly unknown for degenerate problems. Füredi's constructions are closely related to the construction given in Theorem 1.19.

Theorem 1.12 is also known to give the correct order of magnitude for $r = s = 3$ (see Brown[12][13]), but in general, the correct order of magnitude is unknown. The best lower bound was proved by Erdős and Spencer[37] using a probabilistic technique:

Theorem 1.14 (Erdős-Spencer).

$$ex(n, K_{t,t}) \geq \frac{1}{2}n^{2-2/(t+1)}$$

However, the above examples suggest that the upper bound is indeed the correct bound. In the case of $K_{3,3}$, while the correct constant factor is not known, the order of magnitude of the extremal graphs is known. Due to a nice construction of Brown[12] [13], we have

Theorem 1.15 (Brown). *For infinitely many values of n, $ex(n, K_{3,3}) \geq \frac{1}{2}n^{5/3} + o(n^{5/3})$.*

We now direct attention to a second famously unsolved problem. Bondy and Simonovits[11] proved in 1974 the famous even cycles theorem.

Theorem 1.16 (Bondy-Simonovits).

$$ex(n, C_{2k}) < 100kn^{1+1/k}$$

In fact they showed something a bit stronger:

Theorem 1.17 (Even Cycles Theorem). *If G is a graph on n vertices with at least $100kn^{1+1/k}$ edges, then G contains a C_{2t} for all $t \in [k, kn^{1/k}]$.*

Again, the first general lower bounds were done by Erdős[24] using probabilistic methods.

Theorem 1.18. $ex(n, C_{2k}) \geq \Theta(n^{1+1/(2k)})$.

Margulis[82], Imrich[69], and Lubotzy, Phillips, and Sarnak[79] have used results in number theory and eigenvalue methods in graph theory to construct *Ramanujan graphs* which have large chromatic number and girth. These graphs show:

$$\mathrm{ex}(n, C_{2k}) \geq \Theta(n^{1+3/(4k+21)}) \tag{1.4}$$

The only cases where orders of magnitude are known are when $k = 2, 3, 5$. These constructions were done using finite geometry (see Benson[6] and Erdős-Rényi-Sós[35]) and then simplified by Wenger[99]. New constructions have also been given by Lazebnik, Ustimenko, and Woldar[77].

Here we give the Erdős-Rényi-Sós-Brown construction for extremal graphs avoiding a four-cycle. First note that a careful inspection of Theorem 1.12 reveals that to obtain an extremal graph, almost all the vertices should have degree approximately $\sqrt{n}$ and any pair of vertices should have exactly one neighbor in common. Luckily, structures like this exist 'in nature,' namely the lines of a projective plane have this very property.

Theorem 1.19. *For infinitely many values of* n*,*

$$ex(n, C_4) \geq \frac{1}{2}n^{3/2} + O(n^{4/3})$$

.

Proof. Fix a prime p and let the vertices of G be the p^2 pairs (x, y) of residues (modulo p). Then connect the vertex (x, y) to the vertex (a, b) iff $ax+by = 1$. (If the resulting graph contains loops, delete them; their are at most p^2 of them.) As there are p solutions to the equation $ax + by = 1$, even after deleting loops, there are at least $\frac{1}{2}p^2(p-1)$ edges in G. It is easy to check that G has no four-cycle as such an event would imply the existence of two lines in the projective plane intersecting in two different places. □

The construction[12], [13] giving the best lower bound for graphs containing no $K_{3,3}$ also arises from geometry. Due to the importance of these constructions, we will reproduce this as well. The idea is to make the infinite graph whose vertices are the points in Euclidean 3-space where two points are connected iff they lie at distance one from each other. Now in this graph there is clearly no $K_{3,3}$; the three points of one class cannot be collinear as no point is equidistant from three collinear points. Thus the three must lie on a circle. But in 3-space, at most two points can be equidistant from three points on a circle.

Proof of Theorem 1.15. Let p be a prime of the form $k-1$. Let G be the graph whose vertices are the triples (x, y, z) of residue classes (modulo p). Two vertices (x, y, z) and (a, b, c) are adjacent iff

$$(a-x)^2 + (b-y)^2 + (c-z)^2 = 1 \tag{1.5}$$

Now a theorem of Lesbegue implies that for fixed (a, b, c), Equation 1.5 has $p(p-1)$ solutions. Thus G has $\frac{1}{2}n^{5/3} + o(n^{5/3})$ edges and G has no $K_{3,3}$. □

We remark that the proof that G has no $K_{3,3}$, while outlined above, is not complete - the language of analytical geometry is needed to formalize the proof. It is interesting that for $p = 4k + 1$, the sphere defined by $(a - x)^2 + (b - y)^2 + (c - z)^2 = 1$ may have three collinear points which is why we must require $p = 4k - 1$.

Another famously open problem (probably due to Erdős) is:

Conjecture 1.20. $ex(n, \{C_4, C_6, \dots, C_{2k}\}) = \Omega(ex(n, C_{2k}))$

Of course, for the first few cases, this is known to be true. Indeed, Győri[64] has proved that if G is any C_6-free graph, then one can delete at most half the edges of G to make it C_4-free as well.

We close this section with a conjecture by Simonovits and Erdős[93]. It is somewhat of an inverse problem: which values are realized as the correct order of magnitude of some extremal graph? They conjecture that all rationals between 1 and 2 are realizable:

Conjecture 1.21. *Let $1 \leq r \leq 2$ be a rational. Then there exists a graph G such that $ex(n, G) = \Theta(n^r)$.*

1.4 Hypergraphs

We now move on to a natural generalization and extension of extremal graph theory; extremal hypergraph theory. Almost all of the questions we posed above for graphs can also be asked for hypergraphs. Indeed, for each specific type of extremal graph problem, there are usually several possible different analog hypergraph problems. In general the hypergraph generalizations of graph problems are much more complex and difficult than the original problems. Thus, despite a plenitude of problems, it is even a challenge to find problems which are solvable. For instance, a natural analog of Turán's Theorem to hypergraphs is completely unsolved - even the simplest cases of this analog question are unknown.

Let $K_k(l)$ be the k-uniform hypergraph on l vertices with every possible edge. This is a natural extension of complete graphs to hypergraphs; note that $K_2(l)$ is just the complete graph on l vertices. Recall that the Turán density $\pi(\mathcal{H})$ of a k-uniform hypergraph $\mathcal{H}$ is $\lim_{n\to\infty} \mathrm{ex}_k(n, \mathcal{H})/\binom{n}{k}$ where $\mathrm{ex}_k(n, \mathcal{H})$ is the size of the largest k-uniform hypergraph on n vertices not containing $\mathcal{H}$ as a subgraph. Letting $\pi_k(l) = \pi(K_k(l))$, it is easy to check that this limit exists for all k and l, but still, no value of π is known for $l > k \geq 3$. Turán himself, in his seminal paper[97] asked how large a 3-uniform hypergraph could be if it does not contain a $K_3(4)$. He gave a construction showing $\pi(K_3(4)) \geq \frac{5}{9}$ and conjectured it was best possible, but to this day, the question remains open.

That is not due to lack of exposure: the problem has been famously open ever since Turán's original paper appeared with his conjecture and there are some notable partial results. In the case mentioned above with $k = 3$ and $l = 4$, other non-isomorphic constructions have been found, all of which have the same size as Turán's original construction. Brown[13] found six, Todorov[96] eight, and Kostochka[73] $2^{(n/3)-2}$ different non-isomorphic families. If the conjecture is true, these different constructions give a clue as to why the problem is so hard; there is no common structure to the extremal hypergraphs. The best upper bound for $\pi(K_3(4))$ is $(3+\sqrt{17})/12 < 0.593$, due to Chung and Lu[21]. This is an improvement on work by Giraud[58], who gave $(-1+\sqrt{21})/6$ as an upper bound. Even earlier, de Caen[18] proved $\pi(K_3(4)) \leq 0.6213$ an improvement on the result of Katona, Nemetz, and Simonovits[72] who showed $\pi(K_3(4)) \leq 9/14$.

Due to the seemingly intractable nature of the Turán problem for complete k-graphs, Katona[70] suggested that a generalization of Mantel's theorem to hypergraphs should be studied. The problem is the following: Find the maximum number of edges in a k-uniform hypergraph on n vertices such that the symmetric difference of any two edges is not contained in a third edge. We write $\mathcal{D}_k$ for the family of k-uniform hypergraphs composed of three edges $\{A, B, C\}$ such that $A \vartriangle B = (A\backslash B) \cup (B\backslash A) \subseteq C$. Bollobás[9] conjectured the following:

Conjecture 1.22. $ex_k(n, \mathcal{D}_k) = \left\lfloor \frac{n}{k} \right\rfloor \left\lfloor \frac{n+1}{k} \right\rfloor \cdots \left\lfloor \frac{n+k-1}{k} \right\rfloor$

Furthermore, he conjectured that the complete, equipartite k-uniform hypergraph is the only extremal hypergraph. Frankl and Füredi[45] proved Conjecture 1.22 for $n \leq 2k$, and Bollobás[9] solved the case $k = 3$. In 1987, Sidorenko[89] proved the case $k = 4$. Here we present Sidorenko's proof for the cases $k = 3, 4$. First we note the following simpler problem posed by de Caen[17]. What is the largest k-uniform family which does not contain 3 members A, B, and C such that $|A \cap B| = k-1$ and $A \triangle B \subseteq C$. Letting $\mathcal{A}_i = \{\{1,2,\dots,k\},\{1,2,\dots,k-1,k+1\},\{i,i+1,\dots,i+k-1\}\}$ and $\mathcal{S}_k = \{\mathcal{A}_2,\mathcal{A}_3,\dots,\mathcal{A}_k\}$ we can rewrite de Caen's problem as: find $\mathrm{ex}_k(n,\mathcal{S}_k)$.

Theorem 1.23 (Sidorenko)**.** *For $k = 3$ and 4, $ex(n,\mathcal{S}_k) = \lfloor \frac{n}{k} \rfloor \lfloor \frac{n+1}{k} \rfloor \cdots \lfloor \frac{n+k-1}{k} \rfloor$*

We remark that this clearly implies Conjecture 1.22 for $k = 3, 4$ as

$$\left\lfloor \frac{n}{k} \right\rfloor \left\lfloor \frac{n+1}{k} \right\rfloor \cdots \left\lfloor \frac{n+k-1}{k} \right\rfloor \leq \mathrm{ex}_k(n,\mathcal{D}_k) \leq \mathrm{ex}_k(n,\mathcal{S}_k)$$

Sidorenko's proof utilizes the Lagrange function of a hypergraph (first used by Motzkin and Strauss[85] to give a new proof of Turán's theorem for graphs, and by Frankl and Rödl[54][55] in this case of hypergraphs.) For $\mathcal{H}$ a k-uniform hypergraph on n vertices, we associate the polynomial

$$f(\mathcal{H}, x_1, x_2, \dots, x_n) = \sum_{H \in \mathcal{H}} \prod_{i \in H} x_i$$

Then the Lagrange function is defined as

$$\lambda(\mathcal{H}) = \max\{f(\mathcal{H}, \mathbf{x}) : \sum_i x_i = 1, x_i \geq 0\}$$

Here we use $\mathbf{x}$ to denote the vector $(x_1, x_2, \dots, x_n)$. The *support* of $\mathbf{x}$ is Supp($\mathbf{x}$)=$\{i : x_i > 0\}$. We say that $\mathcal{H}$ is a 2-*cover* if every pair of vertices is contained in an edge of $\mathcal{H}$. We call $\mathcal{H}(j) = \{H \backslash \{j\} : j \in H \in \mathcal{H}\}$ the *link* of vertex j. Also, for $S \subseteq \mathcal{V}(\mathcal{H})$, we let $\mathcal{H}|_S$ represent the hypergraph defined on the vertices S with edge set $\{h \in \mathcal{E}(\mathcal{H}) : h \subseteq S\}$. Frankl and Rödl[54] show that if $\mathbf{x}$ is chosen to achieve the maximum ($f(\mathcal{H}, \mathbf{x}) = \lambda(\mathcal{H})$) and have the smallest support, then the induced hypergraph $\mathcal{H}|_{\text{Supp}(\mathbf{x})}$ is a 2-cover. In this case they also show,

$$f(\mathcal{H}(j), \mathbf{x}_{-j}) = k\lambda(\mathcal{H}) \tag{1.6}$$

where $\mathbf{x}_{-j} = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$. We are now ready to present Sidorenko's proof.

Proof of Theorem 1.23. Let $\mathcal{H}$ be an $\mathcal{S}_k$-free family. Let $\mathbf{x}$ be chosen so that $f(\mathcal{H}, \mathbf{x})$ is maximal and such that $\text{Supp}(\mathbf{x})$ is minimal. For $i, j \in \text{Supp}(\mathbf{x})$, we must have the links $\mathcal{H}(i)$ and $\mathcal{H}(j)$ disjoint. Otherwise there are two edges $H_1, H_2 \in \mathcal{H}$ such that $|H_1 \cap H_2| = k - 1$ and $H_1 \triangle H_2 = \{i, j\}$. But $i, j \in \text{Supp}(\mathbf{x})$ which implies the existence of an edge $H_3 \in \mathcal{H}|_{\text{Supp}(\mathbf{x})}$ containing both i and j. But then $\mathcal{H}$ is not $\mathcal{S}_k$-free, so this cannot happen. Thus the polynomials $f(\mathcal{H}(j), \mathbf{x}_{-j})$ have no common terms and therefore

$$\sum_{j \in \text{Supp}(\mathbf{x})} f(\mathcal{H}(j), \mathbf{x}_{-j}) \leq \sigma_{k-1}(\text{Supp}(\mathbf{x}))$$

where $\sigma_{k-1}(\text{Supp}(\mathbf{x}))$ is the elementary symmetric polynomial of rank $k-1$ with variables from $\text{Supp}(\mathbf{x})$. Letting $m = |\text{Supp}(\mathbf{x})|$, and as $x_i \geq 0$, we can bound the right hand side by $\sigma_{k-1}(\frac{1}{m}, \dots, \frac{1}{m})$. Using Equation 1.6, we get

$$km\lambda(\mathcal{H}) \leq \frac{\binom{m}{k-1}}{m^{k-1}} \tag{1.7}$$

For $k = 3$, the right hand side is clearly at most km/k^k for $m \geq k$. The same is true for $k = 4$ and $m \geq k$ if $m \neq 5$. However, m cannot be 5 as then either $||\mathcal{H}|_{\text{Supp}(\mathbf{x})}|| = 1$ in which case it is not a 2-cover, or $||\mathcal{H}|_{\text{Supp}(\mathbf{x})}|| \geq 2$

and $\forall H, H' \in \mathcal{H}|_{\text{Supp}(\mathbf{x})},\ |H \cap H'| = 3 = k-1$ implying that $\mathcal{H}$ is not $\mathcal{S}_k$-free, a contradiction. $\square$

Conjecture 1.22 remains open for $k \geq 5$. When $k = 3$, there are two non-isomorphic forbidden systems: $\mathcal{H}_1 = \{123, 124, 234\}$ and $\mathcal{H}_2 = \{123, 124, 345\}$. Given Bollobás' result, the next step has been to find the Turán density of the respective extremal systems. Frankl and Füredi[44] proved $\pi(\mathcal{H}_2) = 2/9$:

Theorem 1.24 (Frankl-Füredi)**.** *$ex_3(n, \mathcal{H}_2) = \left\lfloor \frac{n}{3} \right\rfloor \left\lfloor \frac{n+1}{3} \right\rfloor \left\lfloor \frac{n+2}{3} \right\rfloor$ for all $n \geq$ 3000.*

Frankl and Füredi[47] also gave a construction for an extremal 3-graph containing no $\mathcal{H}_1$ which has Turán density 2/7. de Caen[16] proved that 1/3 is an upper bound for $\pi(\mathcal{H}_1)$. This was improved upon by Mubayi[83] to $\frac{1}{3} - (4.5305 \times 10^{-6})$ and by Talbot[95] to $.32975 < \frac{1}{3} - \frac{1}{280}$. This is currently the best upper bound for $\pi(\mathcal{H}_1)$.

There are some other hypergraphs besides $\mathcal{H}_2$ for which the Turán density is known. The most notable is the Fano plane. In 1976, V. T. Sós[94] conjectured that the Turán density of the Fano plane was $\frac{3}{4}$. The conjecture remained open for almost thirty years until de Caen and Fürei[19] proved its veracity. Their simple proof inspired more work in the Turán densities of various simple 3-uniform hypergraphs. (See for instance [84].)

The complimentary question to the Turán density of hypergraphs is the determination of the Turán numbers $T(n, k, l)$. This is the size of the minimum k-uniform hypergraph such that every l-set is contained in at least one edge. Note that $T(n,k,l) = \binom{n}{k} - \mathrm{ex}_k(n, K_k(l))$. This formulation of the problem has perhaps been studied even more than the original; in this form the question is very similar to packing questions. Another related question is the determination of $f_k(n,v,t) = \mathrm{ex}_k(n, \mathcal{L}_k(v,t))$ where $\mathcal{L}_k(v,t)$ is the collection of all k-uniform hypergraphs on v vertices with t edges. Ruzsa and Szemerédi[88] showed that $f_k(n,6,3)$ has a non-polynomial order of magnitude. Specifically, they showed that $f_k(n,6,3) = o(n^2)$ and for all $\epsilon > 0$, lim $f_k(n,6,3)/n^{2-\epsilon} = \infty$. This is remarkable as it seems likely that in the case of graphs, all the extremal graphs have polynomial orders of magnitude. Frankl and Füredi [48] found a single hypergraph whose extremal hypergraph also has a non-polynomial order of magnitude. These results show how much more complicated Turán's problem is for hypergraphs as compared to graphs. (Though as we have seen, the degenerate problem in extremal graph theory is already quite hard!)

2 Hypergraphs Avoiding Cycles of Given Length

Having looked at extensions of complete graphs in the setting of hypergraphs, we consider the problem of finding extremal hypergraphs avoiding cycles of a given length. It is not surprising that there are many possible definitions of a cycle in a hypergraph. In fact, we have already discussed one possible generalization of a triangle in hypergraphs due to Katona: namely a triple of edges such that the symmetric difference of two of them is contained within the third. Here we note another important generalization (following the approach of Berge[7][8]). A k-cycle in a hypergraph $\mathcal{H}$ is a sequence of k distinct vertices $v_0, \ldots, v_{k-1}$ and distinct edges $H_0, \ldots, H_{k-1}$ such that for all $0 \leq i \leq k-1$, $v_i, v_{i+1} \subset e_i$ (here we consider the subscripts modulo k). Hypergraphs avoiding such cycles are the focus of this Section. Of course there are even other types of cycles in hypergraphs that have been studied, but we do not have space here to cover all such problems.

2.1 The Berge Approach

We now consider a hybrid approach to hypergraph Turán problems. In contrast to the problems outlined in Section 1.4, here we consider extremal hypergraph problems that have a distinctly graph theoretic flavor. This point of departure is closely associated with Berge[7][8], who grafted graph theory problems and methods into the theory of hypergraphs. The difference between this approach and previously described approaches lies in how one generalizes the basic structures of graphs for hypergraphs. For instance, as we have already seen, there are several possible generalizations of complete graphs in a hypergraph setting. Berge's approach to the question of such generalization problems is the following. Suppose we want to find an appropriate generalization of a graph G to a hypergraph. The information contained in the graph G can be expressed in an adjacency matrix, or even in an incidence matrix - a graph is an incidence structure of points and lines where each line is incident with exactly two points. Now if $M(G)$ is the incidence matrix of the graph G, the Berge hypergraph generalization of G is any hypergraph H with incidence matrix $M(H)$ where $M(G)_{i,j} = 1 \Rightarrow M(H)_{i,j} = 1$. In other words, the same incidences occur in H (or possibly more) as in G.

Note that the definition above of a k-cycle in a hypergraph follows this approach exactly; all that matters is that the correct incidents occur.

We will consider the Turán problem on hypergraphs for such cycles. However as we have seen, the realm of hypergraphs is quite large and it is important to keep in mind some differences between different hypergraph settings. For instance, one can ask how large a r-uniform hypergraph can be containing no cycle of length k or how large a non-uniform hypergraph can be containing no cycle of length k. These questions can be quite different from each other, namely, in the non-uniform case it is often very hard to prove upper bounds with good constant coefficients. In the uniform setting it is somewhat easier to prove tighter theorems with sharp results. Another possible variation is to allow multiple edges. In some cases this is impractical; one could have as many copies of a give 3-edge as one wants and still avoid a cycle of length k for any fixed $k \geq 4$. However, if the size of the edges is larger than the size of the forbidden cycle, it makes sense to find the largest multi-hypergraphs avoiding the given cycle; in this case we do.

One of the most striking aspects of this Berge-approach to Turán problems in hypergraphs is how strongly connected the problems are to the analogous graph theoretic questions. However, there are some notable differences. When considering forbidden cycles in the theory of graphs, the parity of the cycle is of utmost importance; there is a huge difference between excluding a $2k$ cycle and excluding a $2k+1$ cycle from a graph (forbidding the first is a degenerate problem, while forbidding the second is a non-degenerate problem.) However, when one considers this problem for hypergraphs with edges of size at least 3, this difference between odd and even cycles is erased. This is true for both uniform and non-uniform hypergraphs.

The following two theorems nicely demonstrate both the similarities to the analogous graph theory problems and the differences. The first was proved by Kostochka and Verstraëte[74].

Theorem 2.1. *Let $\mathcal{H}$ be a hypergraph on $n \geq 3$ vertices whose edges are all of size at least $k \geq 2$. If $\mathcal{H}$ contains no even cycles then* $\sum_{h \in \mathcal{E}(\mathcal{H})}(|h|-1) \leq \lfloor \frac{k}{k-1}(n-1) \rfloor - 1$

This theorem clearly shows the connections between graphs and hypergraphs as it gives very similar bounds for both graphs and hypergraphs. On the other hand, in the case of odd cycles, Gyárfás, Jacobson, Kézdy, and Lehel[60] proved the following.

Theorem 2.2. *Let $\mathcal{H}$ be a hypergraph on n vertices whose edges are all of size at least $k \geq 3$. If $\mathcal{H}$ contains no odd cycle then $\sum_{h \in \mathcal{E}(\mathcal{H})}(|h|-1) \leq 2n-2$.*

Of course, it is important here that no 2-edges are allowed here; in a graph one can have $n^2/4$ edges without having an odd cycle. As we shall see, this is not an unusual result; hypergraphs with forbidden odd or even cycles behave similarly to each other as long as all the edges are of size at least 3.

2.2 Uniform Hypergraphs

We will now consider extremal hypergraphs with just one forbidden cycle. The first such result was found by Győri[64] who, motivated by a problem in number theory, found the maximal number of edges in an unbalanced C_6-free bipartite graph. Later he realized that this question can be reformulated in the language of hypergraphs. (It is clear that a C_6-free bipartite graph can be thought of as the incidence graph of a C_3-free (multi-)hypergraph.) Győri[65] was able to improve his previous bounds by looking at the hypergraph case. Indeed, he proved

Theorem 2.3 (Győri)**.** *Let $\mathcal{H}$ be a (multi-)hypergraph on n vertices containing no triangles. Then for large enough n,*

$$\sum_{H \in \mathcal{H}} (|H|-2) \leq n^2/8$$

His result is best possible. Here we prove a special case of Theorem 2.3 (due to Győri[65] and independently, Frankl, Füredi and Simonyi[[50]]).

Theorem 2.4. *Let $\mathcal{H}$ be a 3-uniform hypergraph on n vertices. Then*

$$e(\mathcal{H}) = \sum_{H \in \mathcal{H}} (|H| - 2) \leq n^2/8$$

Proof. First note that for any three distinct edges, H_1, H_2, H_3 in $\mathcal{H}$ with $|H_1 \cap H_2| = |H_1 \cap H_3| = 2$, we must have $H_1 \cap H_2 = H_1 \cap H_3$. Otherwise the three edges constitute a triangle. Thus for a given hyperedge $H = \{x, y, z\}$ we may assume without loss of generality that for all $H^{'} \in \mathcal{H}$ distinct from H, we have that $\{x, y\}, \{x, z\} \not\subset H^{'}$. Then we can replace H with two graph edges, namely x, y and x, z. Doing this for every edge, we produce a graph G on the vertices of $\mathcal{H}$. Note that in this graph, there are no multiple edges. It is also easy to see that G is triangle-free. But then Mantel's theorem says that $e(G) \leq n^2/4$, so we conclude that $e(\mathcal{H}) \leq n^2/8$ as desired. □

Motivated by this result and the conjecture of Erdős mentioned in Section 1.2, Győri and Bollobás[10] asked the following two questions.

1) How many triangles can a C_5-free graph have?

2) How many edges can a 3-uniform C_5-free hypergraph have?

As mentioned in Section 1.2, they proved[10] that the answer to both these questions is $\Theta(n^{3/2}) = \Theta(ex_2(n, C_4))$.

Theorem 2.5 ([10])**.** *Let $\mathcal{H}$ be a 3-uniform hypergraph on n vertices containing no 5-cycle. Then* $e(\mathcal{H}) \leq \sqrt{2}n^{3/2} + 4.5n$

Theorem 2.6 ([10]). *If G is a graph on n vertices containing no $\mathcal{C}_5$ then the number of triangles in G is at most* $(\sqrt{2}/4+1)n^{3/2}+o(n^{3/2})$.

It is not an accident that both these answers are of the same order of magnitude as that of C_4-free graphs; the constructions giving the lower bounds are built upon C_4-free graphs. We reproduce those here:

Construction 2.7. Let G_0 be a C_4-free bipartite graph with classes of size $n/3$ and approximately $(n/3)^{3/2}$ edges. Double each vertex in one of the classes and add an edge joining the old and new copy. Call this new graph G. Clearly G has n vertices and $2(n/3)^{3/2}+n/3$ edges. The number of triangles in G is equal to the number of edges in G_0. Also, it is easy to check that G has no 5-cycles.

Construction 2.8. Let G_0 be a C_4-free bipartite graph with classes of size $n/3$. Again, double each vertex in one of the classes to turn each edge into a triple. Call the resulting hypergraph $\mathcal{H}$. Now $\mathcal{H}$ has n vertices, and $e(\mathcal{H})=e(G)=(n/3)^{3/2}+o(n^{3/2})$ edges. It is easy to check that here as well, there is no 5-cycle in $\mathcal{H}$.

To emphasize the close relation between even and odd cycles (in hypergraphs) we note that Lazebnik and Verstraëte[76] showed that the extremal r-uniform hypergraphs of girth 5 have also order of magnitude $n^{3/2}$ (here as always n is the number of vertices.)

Theorem 2.9 (Lazebnik-Verstraëte). *Let $\mathcal{H}$ be a 3-uniform hypergraph on n vertices with girth at least 5. Then* $e(\mathcal{H})\le \frac{1}{6}n\sqrt{n-\frac{3}{4}}+\frac{1}{12}n$

Theorem 2.9 is sharp when $k = 3$ for infinitely many values of n while Theorem 2.5 is sharp up to a constant factor. Thus extremal 3-uniform hypergraphs with girth 5 have the same order of magnitude as extremal 3-uniform hypergraphs containing no cycle of length 5.

Of course, one can ask questions (1) and (2) above for other odd cycles as well. The general cases for question (1) were solved by Győri and Li[67] who proved that

Theorem 2.10 (Győri-Li). *If G is a C_{2k+1}-free graph, then the number of triangles in G is at most*

$$\frac{(2k-2)(16k-8)}{3} ex(n, C_{2k})$$

Remark 2.11. This theorem is essentially sharp (i.e. up to the constant factor) if, as conjectured the functions $\mathrm{ex}(n, C_4, c_6, \ldots, C_{2k})$ and $\mathrm{ex}(C_{2k})$ are essentially the same.

On the other hand, the generalization of question (2) was solved by Győri and Lemons[66] who proved

Theorem 2.12 (Győri-Lemons). *Let $\mathcal{H}$ be a 3-uniform hypergraph containing no cycle of length $2k+1$. Then $e(G) < 4k^4 n^{(k+1)/k} + 15k^4 n + 10k^2 n$.*

This will be proved in Section 2.2.2 and in Section 2.2.1 we will prove the following related result:

Theorem 2.13 (Győri-Lemons). *Let $\mathcal{H}$ be a 3-uniform hypergraph containing no cycle of length $2k$. Then $e(\mathcal{H}) = O(n^{1+1/k})$.*

In fact we will show that the same order of magnitude is an upper bound in Theorems and for all l-uniform hypergraphs if $l \geq 3$. We note that both theorems give the correct orders of magnitude for the extremal hypergraphs if the following famous conjecture is true for graphs:

$$\text{ex}(n, C_4, C_6, \ldots, C_{2k}) = \Theta(n^{(k+1)/k})$$

(This is known to be true for $k = 2, 3, 5$.) This again shows how closely connected the graph and hypergraph Turán problems are for forbidden cycles. If we take an extremal bipartite graph with girth $2k+2$, then from this graph, we can easily build a 3-uniform hypergraph with no cycles of length $2k$ or $2k+1$ as the following construction shows.

Construction 2.14. Let G be a an extremal bipartite graph with girth $2k+2$ and each part containing n vertices. Take $2k-2$ copies of each of the vertices in one part; in this way we now have $2kn$ vertices. Now from each edge $e = \{x, y\}$ of G we produce $\binom{2k-1}{2}$ 3-edges. Suppose x is the vertex of e which we copied $2k-2$ times. Then we take all the $\binom{2k-1}{2}$ pairs made up from x and its copies. To each of these pairs we add the vertex y; in this way we obtain a 3-uniform hypergraph $\mathcal{H}$. (This will produce a $\mathcal{C}_{2k+1}$-free hypergraph. If we want to produce a $\mathcal{C}_{2k}$-free hypergraph, we simply take one less copy of the each of the copied vertices.)

If C is a cycle in $\mathcal{H}$, then we call C' in G the trace of C, if C' is obtained by identifying each duplicated pair of $\mathcal{H}$. Note that C' will have length no more than that of C. It is clear that $\mathcal{H}$ can have no cycle of length $2k+1$ as such a cycle would trace out a cycle of length l, $1 < l \leq 2k$ in the graph G, a contradiction.

Construction 2.14 can be modified by taking all $(m-1)$-sets from within the copied vertices (instead of the pairs) together with the unduplicated vertex to obtain $2k+1$-free (or $2k$-free) m-uniform hypergraphs from bipartite graphs of girth $2k+2$ (as long as $2k > m$.) Thus our results can be stated in the following manner:

Theorem 2.15.

$$O(n^{1+1/k}) \geq ex_m(n, C_{2k}) \geq \Omega(ex_2(n, C_4, C_6, \ldots, C_{2k}))$$

$$O(n^{1+1/k}) \geq ex_m(n, C_{2k+1}) \geq \Omega(ex_2(n, C_4, C_6, \ldots, C_{2k}))$$

2.2.1 Forbidden Even Cycles

While our results are similar for even and odd cycles, the proof techniques are quite different. The difference stems from the difference in graph theory between graphs avoiding a given even or a given odd cycle. That is, our bound concerning l-uniform hypergraphs containing no $\mathcal{C}_{2k}$ is built up from the Bondy-Simonovits Theorem. The following lemma is our main tool in this approach.

Lemma 2.16 (Reduction Lemma)**.** *Let $2 \leq l < m$ and let $\mathcal{H}$ be a m-uniform hypergraph with no cycle of length k. Define the l-uniform hypergraph $\mathcal{H}_1$ in the following way. Order the edges of $\mathcal{H}$ and going through the edges one by one, pick an l-set from each to be in $\mathcal{H}_1$, but at each point, pick this l-set* *to have as small multiplicity as possible. Obviously the hypergraph $\mathcal{H}_1$ will also have no k-cycle and each of its hyperedges will have multiplicity no more than $\sum_{j=1}^{k-1}\binom{j-l}{m-l}+k+1$.*

Remark 2.17. Note that if $k \leq m$ then the lemma trivially holds; if a hyperedge of $\mathcal{H}_1$ has multiplicity more than $k+1$ then there is an hyperedge h of $\mathcal{H}$ such that every l-subset of h is contained within more than k edges and we can clearly find a k-cycle within h, a contradiction.

The first statement of Theorem 2.15 immediately follows as a corollary of this result when $l=2$ and the Bondy-Simonovits Theorem[11]. However, the constant factor produced in this way is far from sharp. (And of course Theorem 2.2) is a special case of Theorem 2.15. Note that we can not use the same technique to prove the second statement of Theorem 2.15 as extremal graphs containing no odd cycle can have $\Theta(n^2)$ edges. Instead, we will prove that 3-uniform hypergraphs with no $\mathcal{C}_{2k+1}$ have at most $O(n^{1+1/k})$ edges. We will then use this Reduction Lemma to extend the result to any l-uniform hypergraph containing no $\mathcal{C}_{2k+1}$ (for $l \geq 3$.) Due to these consequences, Lemma 2.16 is quite an important result; yet at the same time it has a simple proof.

Proof of Lemma 2.16. We suppose that the cycle is even; the proof for odd cycles is exactly the same and is thus omitted. Suppose the lemma is not true. Let $\mathcal{H}$ be an m-uniform hypergraph containing no $\mathcal{C}_{2k}$ and let $\mathcal{H}_1$ be a l-uniform hypergraph defined as in the statement of the lemma and suppose $\mathcal{H}_1$ has an edge, e_1, of multiplicity greater than $M=\sum_{j=1}^{2k-1}\binom{j-l}{m-l}+2k+1$.

We will find a sequence of hyperedges $h_1, \ldots, h_{2k}$ in $\mathcal{H}$ forming a $2k$-cycle. To distinguish between the hypergraphs $\mathcal{H}$ and $\mathcal{H}_1$, we will refer to the edges of the former as hyperedges and those of the later simply as edges. Similarly we will use the variables h_i for hyperedges in $\mathcal{H}$ and the variables e_i for edges in $\mathcal{H}_1$. For $h \in \mathcal{E}(\mathcal{H})$, we will refer to the l-set chosen in the construction of $\mathcal{H}_1$ to represent h as $e(h)$.

Consider the last hyperedge which contributed to the multiplicity of e_1 in the construction of $\mathcal{H}_1$. Call this hyperedge h_1. (Thus for instance, $e(h_1) = e_1$.) By definition, each l-set contained within h_1 must be an edge in $\mathcal{H}_1$ and indeed, must have multiplicity at least M. Otherwise when we picked $e(h_1)$ from h_1 to be in the graph, we would have chosen one of the other less-represented l-sets in h_1. Let the vertices $x_0, x_1 \in e_1$ and pick $e_2 \subset h_1$ with $x_1 \in e_2 \neq e_1$. Let $y_1 \in e_2 \backslash \{x_1\}$.

As e_2 has multiplicity at least M in $\mathcal{H}_1$, we can find $h_2 \in \mathcal{E}(\mathcal{H})$ with $e(h_2)$ a multiple of e_2 such that every l-set contained in h_2 is an edge in $\mathcal{H}_1$ with multiplicity at least $M-1$. Note that h_1 does not contribute to the multiplicity of e_2 as $e(h_1)$ is a multiple of e_1.

Claim 2.18. *Let $i \leq 2k-1$ be odd. Suppose we can find distinct vertices $x_1, x_2 \ldots, x_{(i-1)/2}$ and $y_1, y_2, \ldots, y_{(i-1)/2}$ and distinct edges $e_1, e_2, \ldots, e_{i-1} \in \mathcal{E}(\mathcal{H}_1)$ and $h_1, h_2, \ldots, h_{i-1} \in \mathcal{E}(\mathcal{H})$ such that the following conditions hold:*

- $\forall j \leq \frac{i-1}{2}, x_j \in (e_{2j-1} \cap e_{2j})$ *and* $y_j \in (e_{2j} \cap e_{2j+1})$
- $\forall j < i-1, e_j \cup e_{j+1} \subseteq h_j$

- $h_{i-1} \not\subseteq (\cup_{j \le (i-1)/2}\{x_j, y_j\})$
- *each l-set in h_{i-1} has multiplicity at least* $M - \sum_{j=1}^{i-1} \binom{j-l}{m-l} - (i-2)$

Then we can find an edge of $\mathcal{H}_1$, $e_i \subset h_{i-1}$ and a vertex $x_{(i+1)/2} \in e_i$ such that $x_{(i+1)/2}$ is distinct from the vertices $\{x_j, y_j\}_{j=1}^{(i-1)/2}$. Moreover, there exists a hyperedge h_i of $\mathcal{H}$, containing e_i such that h_i is not contained in $\{x_j, y_j\}_{j=1}^{(i-1)/2} \cup \{x_{(i+1)/2}\}$ and such that each l-set of h_i occurs with multiplicity at least $M - \sum_{j=1}^{i} \binom{j-l}{m-l} - (i-1)$.

Claim 2.19. *Let $i \le 2k$ be even. Suppose we can find distinct vertices $x_1, x_2 \dots, x_{i/2}$ and $y_1, y_2, \dots, y_{(i/2)-1}$ and distinct edges $e_1, e_2, \dots, e_{i-1} \in \mathcal{E}(\mathcal{H}_1)$ and $h_1, h_2, \dots, h_{i-1} \in \mathcal{E}(\mathcal{H})$ such that the following conditions hold:*

- $\forall j \le \frac{i}{2}, x_j \in (e_{2j-1} \cap e_{2j})$ *and* $\forall j \le \frac{i}{2} - 1, y_j \in (e_{2j} \cap e_{2j+1})$
- $\forall j < i-1, e_j \cup e_{j+1} \subseteq h_j$
- $h_{i-1} \not\subseteq (\cup_{j \le (i/2)-1}\{x_j, y_j\} \cup \{x_{i/2}\})$
- *each l-set in h_{i-1} has multiplicity at least* $M - \sum_{j=1}^{i-1} \binom{j-l}{m-l} - (i-2)$

Then we can find an edge of $\mathcal{H}_1$, $e_i \subset h_{i-1}$ and a vertex $y_{i/2} \in e_i$ such that $y_{i/2}$ is distinct from the vertices $\{x_j, y_j\}_{j=1}^{(i/2)-1}$ and $x_{i/2}$. Moreover, there exists a hyperedge h_i of $\mathcal{H}$, containing e_i such that h_i is not contained in $\{x_j, y_j\}_{j=1}^{i/2}$ and such that each l-set of h_i occurs with multiplicity at least $M - \sum_{j=1}^{i} \binom{j-l}{m-l} - (i-1)$.

Remark 2.20. The proofs of these statements are almost exactly the same. The only difference lies in the parity of i; in one case an $x_{(i-1)/2}$ is produced, in the other a $y_{i/2}$ is produced. To avoid unnecessary repetition, we will only prove Claim 2.18.

Proof of Claim 2.18. We find $x_{(i+1)/2}$, e_i and h_i in the following way. By the third condition, we can pick a new edge $e_i \subset h_{i-1}$ such that there exists a vertex

$$x_{\frac{i+1}{2}} \in e_i \backslash \Big(\bigcup_{j \leq \frac{i-1}{2}} \{x_j, y_j\} \Big)$$

By the fourth condition, e_i has multiplicity at least $M - \sum_{j=1}^{i-1} \binom{j-l}{m-l} - (i-2)$ in $\mathcal{H}_1$. Let $\mathcal{E}_i^* = \{h \in \mathcal{E}(\mathcal{H}) : e(h)$ is a multiple of $e_i\}$. We want to pick $h \in \mathcal{E}_i^*$ such that h is not contained within the x_j and y_j. As there are i such points, there are at most $\binom{i-l}{m-l}$ such forbidden edges in $\mathcal{E}_i^*$. Thus, letting $\mathcal{E}_i = \{h \in \mathcal{E}_i^* : h \not\subseteq \cup\{x_j, y_j\}\}$, we have

$$|\mathcal{E}_i| \geq |\mathcal{E}_i^*| - \binom{i-l}{m-l} - (i-2) \geq M - \sum_{j=1}^{i} \binom{j-l}{m-l} - (i-2)$$

Then picking h_i to be the last hyperedge (in our original ordering of the edges) of $\mathcal{E}_i$, we must have that every l-set in h_i has multiplicity in $\mathcal{H}_1$ at least $M - \sum_{j=1}^{i} \binom{j-l}{m-l} - (i-1)$. In this way we have picked x_i, e_i, and h_i such that the above five conditions still hold. $\square$

Thus by induction on i and using Claims 2.18 and 2.19, we can find $2k$ vertices $x_1, \ldots, x_k$ and $y_1 \ldots, y_k$ such that there are $2k$ edges e_i in $\mathcal{H}_1$ and $2k$ hyperedges in $\mathcal{H}$ such that for $1 \leq j \leq k$ we have: $x_j \in e_{2j-1} \cap e_{2j}$ and

$y_j \in e_{2j} \cap e_{2j+1}$. We also have for $2 \leq j \leq 2k$ that $e_j \subset h_{j-1} \cap h_j$. Thus we conclude that for $1 \leq j \leq k-1$, we have $x_j, x_{j+1} \in h_{2j}$ and $y_j, y_{j+1} \in h_{2j+1}$. Also we clearly have $x_k, y_k \in e_{2k} \subset h_{2k}$ and $x_1, y_1 \in e_1 \subset h_1$. Thus we have produced the following $2k$-cycle in $\mathcal{H}$: $x_1, \ldots, x_k, y_k, \ldots, y_1$ (in the odd case: $x_1, \ldots, x_k, x_{k+1}, y_1, \ldots, y_k$); we have arrived at our contradiction. □

We state the important special case when $m = 3$ and $l = 2$ of Lemma 2.16 as a corollary as we will use it several times in the proof of Theorem 2.2.

Corollary 2.21. *Let $\mathcal{H}$ be a 3-uniform hypergraph with no cycle of length k. Define the 2-graph G in the following way. Order the edges of $\mathcal{H}$ and going through the edges one by one, pick a 2-set from each to be in G, but at each point, pick this 2-set to have as small multiplicity as possible. Then the graph G will also have no k-cycle and each of its edges will have multiplicity no more than $\binom{2k+1}{2} + 1$.*

2.2.2 Forbidden Odd Cycles

We now turn our attention to bounding the size of a 3-uniform hypergraph containing no $\mathcal{C}_{2k+1}$; that is we will prove Theorem 2.2. This, together with the Reduction Lemma will imply the verity of the second part of Theorem 2.15. The general proof technique which we describe here will be used again to bound the size of non-uniform hypergraphs avoiding a given cycle (in Section 2.3).

Suppose $\mathcal{H}$ is a hypergraph containing no cycle of length $2k+1$. The idea is to pick a vertex $x \in \mathcal{V}(\mathcal{H})$ and prove that each off the first k successive neighborhoods of x must be approximately the average degree in $\mathcal{H}$ times the size of the previous neighborhood. If it is also true that $\mathcal{H}$ contains no cycles of length less then $2k$, then each of these neighborhoods will be disjoint. As the k^{th} neighborhood can be no larger then $O(n)$, we can argue that the average degree in $\mathcal{H}$ should be no more than $O(n^{1/k})$.

However, we can not assume that these successive neighborhoods are distinct; thus while the we utilize the same type of argument, the technicalities are a bit more complicated. Namely, we will define disjoint subsets of $\mathcal{V}(\mathcal{H})$ which will act as successive neighborhoods of the fixed vertex x and which will still be approximately as large as the real neighborhoods.

Definition 2.22. An *approximate k-neighborhood structure* of the hypergraph $\mathcal{H}$ around the vertex $x \in \mathcal{V}(\mathcal{H})$ consists of $k+1$ pairwise disjoint sets $S_0, \ldots, S_k \subset \mathcal{V}(\mathcal{H})$, a subset $\mathcal{E}_S \subset \mathcal{E}(\mathcal{H})$ and maps $\pi : S_i \to S_{i-1}$ and $\phi : \cup_{i>0} S_i \to \mathcal{E}_S$ such that $S_0 = \{x\}$, ϕ is a bijection and for each $v \in S_i$, $\{v, \pi(v)\} \subseteq \phi(v)$. An approximate k-neighborhood structure is *greedily defined* if for each $i > 0$, there is no edge $e \in \mathcal{E}(\mathcal{H}) \backslash \phi(\cup_{j=1}^{i} S_j)$ with $e \cap S_{i-1} \neq \emptyset$ and $e \nsubseteq (S_{i-1} \cup S_i)$.

It is clear that greedily defined approximate neighborhood structures always exist and can easily be found as the following algorithm shows. Pick $x \in \mathcal{V}(\mathcal{H})$ and set $S_0 = \{x\}$ and $\mathcal{E}_S = \emptyset$. Suppose we have defined already S_{i-1}. Then S_i is defined in the following way.

(*) If there is an edge $h \in \mathcal{E}(\mathcal{H})\backslash\mathcal{E}_S$ with $h \cap S_{i-1} \neq \emptyset$ and $h \not\subset (S_{i-1} \cup S_i)$, then pick $u \in h\backslash(\cup_{j\leq i})$ and $v \in h \cap S_{i-1}$. Add u to S_i and h to $\mathcal{E}_S$ and set $\phi(u) = h$ and $\pi(u) = v$.

Repeat (*) until there are no more such edges. Then we have defined S_i. Before we proceed with the proof of Theorem 2.2, we state two important lemmas which we will use. The first is a type of complimentary result to the generalization of the Erdős-Gallai Theorem (see Theorem 1.10.) The second is a result about approximate neighborhood structures. We state and prove both lemmas before the proof of the theorem.

Lemma 2.23. *Let $\mathcal{H}$ be a $\mathcal{C}_{2k+1}$-free 3-uniform hypergraph with approximate neighborhood structure $(S_i, \mathcal{E}_S, \pi, \phi)$ around a fixed vertex x. Then if $i \leq k$ and G is a graph on S_i such that $\exists \mathcal{E}^* \subset \mathcal{E}(\mathcal{H})\backslash\mathcal{E}_S$ and a bijection $\psi : E(G) \to \mathcal{E}^*$ such that $\forall e \in E(G), e \subset \psi(e)$. Then ignoring multiplicities (in G), $e(G) < 2|S_i|\binom{k}{2}$.*

Proof. We partition the edges of G into sets $E_1, E_2, \ldots, E_i$, where $e = \{v_1, v_2\} \in E_j$ if j is the smallest integer such that $\pi^j(v_1) = \pi^j(v_2)$ (here and later, π^j is the j^{th} power of the map π.) Note that for all $v \in S_i$, $\pi^i(v) = x$ so there always exists such an integer. Now if $u, v \in \mathcal{V}(G)$ are in the same connected component of $G|_{\mathcal{E}_j}$ then clearly $\pi^j(u) = \pi^j(v)$. We claim there can be no path in $G|_{E_j}$ of length $2(k-j)+1$ with endpoints u and v where $\pi^{j-1}(u) \neq \pi^{j-1}(v)$. Otherwise, there exists a path in $\mathcal{H}|_{\mathcal{E}(\mathcal{H})\backslash\mathcal{E}_S}$ of length $2(k-j)+1$ between u and v. But, since j is the smallest integer with

$\pi^j(u) = \pi^j(v)$, there is a path of length $2j$ in $\mathcal{H}|_{\mathcal{E}_S}$ between u and v. These two paths in $\mathcal{H}$ are vertex disjoint (except for the common vertices u and v) since the first is contained within $\cup_{l<i} S_l$ and the second within S_i. Thus the two paths form a $2k+1$ cycle in $\mathcal{H}$, a contradiction. Applying Lemma 1.10, we find that, ignoring multiplicities, $e(G|_{E_j}) \leq 2(k-j)|S_i|$. As this is true for any j, $1 \leq j \leq k$, we conclude:

$$\begin{aligned} e(G) &\leq 2|S_i| \sum_{j=1}^{i} (k-j) \\ &= 2|S_i| [\binom{k}{2} - \binom{k-i}{2}] \\ &< 2|S_i| \binom{k}{2} \end{aligned} \tag{2.1}$$

□

Lemma 2.24. *Let G be a bipartite graph with parts A and B. Fix i and k with $0 < i < k$. Suppose there are partitions $\mathcal{R}_j, 0 \leq j \leq i$ of A with the following properties:*

(1) For each $j < i$, $\mathcal{R}_{j+1}$ is a refinement of $\mathcal{R}_j$. $\mathcal{R}_0$ is the trivial partition (there is only one partition class) and $\mathcal{R}_i$ is the discrete partition.

(2) For $0 \leq l \leq i-1$, there are no paths of length $2(k-i+l)$ with endpoints $u, v \in A$ such that u, v are in the same partition class of $\mathcal{R}_l$ but are in different classes of $\mathcal{R}_{l+1}$.

Then the average degree in G is less than $2k^2$.

Proof. Suppose not. Then we can delete vertices of G of degree less than k^2 and their incident edges without decreasing the average degree of G. Thus we may assume that the minimum degree of G is at least k^2. Consider the partition $\mathcal{R}_1$ of A. We say a vertex $v \in B$ is of Type 1 if no more than k of its neighbors lie outside some class, say R_1^m of $\mathcal{R}_1$. In this case, we delete the edges incident with v not going into R_1^m. (We do this for all Type 1 vertices.) We call all other vertices in B Type 2 vertices. Now we have deleted at most $k|B|$ edges from G. Thus, $e(G) \geq k^2(|A|+|B|) - k|B| > k(k-1)(|A|+|B|)$. Thus the average degree in G is now no less than $2k(k-1)$, so again, we may delete all vertices with degree less than $k(k-1)$ without lowering the average degree. After these modifications, the minimal degree of G will be no less than $k(k-1)$.

Claim 2.25. *$\forall u, v \in B$, if there is a path of length $2(k-i+1)$ from u to v in G, then u and v are Type 1 vertices both of whose neighbors lie in the same partition class of $\mathcal{R}_1$.*

Proof. Consider a path P_0 of length $2(k-i-1)$ with endpoints in B. Let $P_0 = b_1, a_1, b_2, \ldots, a_{k-i-2}, b_{k-i-1}$. Suppose at least one of the endpoints (say b_1) of P_0 is of Type 2. Then we may extend P_0 to a path of length $2(k-i)$ with endpoints in different partition classes of $\mathcal{R}_1$. This is clear: first we pick a neighbor of b_{k-i-1} not already in P_0. We can do this because $d(b_{k-i-1}) \geq k(k-1)$. This new vertex, a_{k-i-1}, is in some partition class of $\mathcal{R}_1$: say R_1^m. But then since b_1 is Type 2 and has degree at least $k(k-1)$, we can find a neighbor of b_1 that is neither in P_0 nor in R_m^1. But this contradicts the property (2) of G above.

If both b_1 and b_{k-i-1} are of Type 1 then by the same reasoning, their neighbors (outside of P_0) must lie in the same partition class of $\mathcal{R}_1$. □

Let $P_0 = b_1, a_1, b_2, \ldots, a_{k-i-2}, b_{k-i-1}$ be a path of length 2(k-i+1) with endpoints in B. Then by Claim 2.25, all the neighbors of b_1 and b_{k-i-1} lie in the same partition class of $\mathcal{R}_1$, say R_1^m.

Claim 2.26. *In any extension of P_0 from b_{k-i+1}, all the vertices of B in the extension are of Type 1 and have neighbors in R_1^m.*

Proof. We show that any extension of P_0 by two points has this property and that b_2 is also of Type 1 with all of its neighbors in R_1^m. The rest follows by induction.

Extend P_0 from b_{k-i+1} by two more points; a_{k-i+1} and then b_{k-i+2}. Call this path P_1. (We can do this because the minimal degree of G is at least $k(k-1)$.) Suppose we can find a neighbor of b_{k-i+2} that is neither in P_1 nor in R_1^m. But then, there is a path of length $2(k-i+1)$ from this vertex to a_1 contradicting property (2) of G as $a_1 \in R_1^m$. Since the degree of b_{k-i+2} is at least $k(k-1)$, we must assume that b_{k-i+2} is a Type 1 vertex with all of its neighbors in R_1^m. Now there is a path of length $2(k-i+1)$ between b_2 and b_{k-i+2}. Thus by Claim 2.25, both b_2 and b_{k-i+2} are Type 1 vertices with neighbors in the same partition class of $\mathcal{R}_1$. We know that the neighbors of b_{k-i+2} lie in R_1^m, therefore the neighbors of b_2 do as well. □

Claim 2.27. *For $1 \leq l \leq i-1$, we can extend P_0 to P_l, a path of length 2(k-i+l-1) such that all the neighbors of the endpoints b_1 and b_{k-i+l} lie within one* *partition class of $\mathcal{R}_l$. Call this partition class R_l^m. Furthermore, by deleting appropriate edges in G, we can insure that in any further extension of P_l from $b_{k-i+l-1}$, all the vertices of the extension in B will only have neighbors in R_l^m and that the minimum degree of G is at least $k(k-l)$.*

Proof. We use induction on l. The first case is proved in Claims 2.25 and 2.26 above. Suppose the statement holds for $j-1$, j fixed. We reclassify the vertices of B. Now, a vertex $v \in B$ is of Type 1 if no more than k of its neighbors lie outside some class R_j^m of $\mathcal{R}_j$. In this case we delete all the edges incident with v not going into R_j^m. All other vertices of B are Type 2 vertices. Again we deleted no more than $k|B|$ edges from G. Thus we have $e(G) \geq k(k-j+1)(|A|+|B|) - k|B| > k(k-j)(|A|+|B|)$. Thus the average degree in G is now no less than $2k(k-j)$ and we can delete vertices of G of degree less than $k(k-j)$ without decreasing the average degree of G. In this way we guarantee that the minimum degree of G is at least $k(k-j)$. Now suppose that one of the endpoints of P_{j-1} is a Type 2 vertex; say b_1. Then we can extend P_{j-1} to a path of length $2(k-i+j)$ with endpoints in different partition classes of $\mathcal{R}_j$, contradicting property (2) of G above. We do this by first picking a neighbor of $b_{k-i+j-1}$ not in P_{j-1}. Let this point be in the partition class R_j^m of $\mathcal{R}_j$. Then since b_1 is a Type 2 vertex and has degree at least $k(k-j) \geq 2k$ ($j \leq i-1 < k-1$) we can pick a neighbor of b_1 that is neither in P_{j-1} nor in the partition class R_j^m. We conclude that both b_1 and $b_{k-i+j-1}$ are of Type 1 and only have neighbors within one class of $\mathcal{R}_j$.

Call this class R_j^m. Note that we can find extensions of P_{j-1} as the minimum degree in G is at least $k(k-j) > 2k$. It is clear that in any extension of P_{j-1}, the new vertices of the path in B must also be of Type 1 and only have neighbors in R_j^m. The argument is exactly the same as in the proof of Claim 2. □

Finally, we can find a path of length $2(k-2)$ the neighbors of whose endpoints lie in the same partition class of $\mathcal{R}_{i-1}$. Since the minimum degree in B is at least $k^2 - k(k-2)$, we can extend our path into A from both ends. But then we get a path of length 2(k-1) contradicting property (2) of G. □

We are ready to prove the main theorem in this section.

Proof of Theorem 2.2. Suppose not. Let $\mathcal{H}$ be a $\mathcal{C}_{2k+1}$-free 3-uniform hypergraph with at least $4k^4n^{(k+1)/k} + 15k^4n + 10k^2n$ edges on n vertices. Since the average degree in $\mathcal{H}$ is at least $12k^4n^{1/k} + 45k^4 + 30k^2$, if we delete all vertices of $\mathcal{H}$ of degree less than $4k^4n^{1/k} + 15k^4 + 10k^2$ together with their adjacent edges, the average degree in $\mathcal{H}$ will not decrease. Thus we may suppose without loss of generality that the minimum degree, δ, in $\mathcal{H}$ is at least $4k^4n^{1/k} + 15k^4 + 10k^2$. Pick x in $\mathcal{V}(\mathcal{H})$ and let $(S_i, \mathcal{E}_S, \pi, \phi)$ be a greedily defined approximate neighborhood structure centered around x.

The proof of the theorem follows directly from the following proposition.

Proposition 2.28. *The sets S_i, for $1 \leq i \leq k$ have the following properties:*

(A) There are no more than $k^4(|S_i|+|S_{i-1}|)$ edges of $\mathcal{E}(\mathcal{H})\backslash\mathcal{E}_S$ intersecting S_i in one point and S_{i-1} in two points.

(B) There are no more than $k^4(|S_i|+|S_{i-1}|)$ edges of $\mathcal{E}(\mathcal{H})\backslash\mathcal{E}_S$ intersecting S_i in two points and S_{i-1} in one point.

(C) $|S_i| > 4k^3n^{i/k}$

Note that Proposition 2.28 implies that $|S_k| > 4k^3n$, a contradiction. □

Thus all that remains is to prove Proposition 2.28.

Proof. We will prove the proposition by induction on i.

Base Case. Define a graph G on vertex set S_1. For each edge $h \in \mathcal{E}(\mathcal{H})\backslash\mathcal{E}_S$ with $x \in h$ let $h\backslash\{x\}$ be in $E(G)$. Now in G, there cannot be a path on $2k$ vertices $v_1, v_2, \ldots, v_{2k}$. Otherwise, there exists distinct edges $\phi(v_1), \phi(v_{2k}) \in \mathcal{E}_S$ adjoining v_1 and v_{2k} to x respectively in $\mathcal{H}$. But then $x, v_1, \ldots, v_{2k}$ form a $2k+1$ cycle in $\mathcal{H}$, a contradiction. The Erdős-Gallai Theorem says that $e(G) \leq (k-1)|\mathcal{V}(G)|$. Now there are exactly $|\{h \in \mathcal{E}_S : x \in h\}| = |S_1| = |\mathcal{V}(G)|$ edges of $\mathcal{H}$ containing x but not represented in G. We conclude that $|S_1| \geq d_{\mathcal{H}}(x)/k > 4k^3n^{1/k}$. Also the number of edges in $\mathcal{H}$ intersecting both S_1 and S_0 is no more than $k|S_1|$.

Inductive Step. Suppose there exists an i, $1 \leq i \leq k$ such that $\forall j \leq i, |S_j| > 4k^3n^{j/k}$ and that the number of edges in $\mathcal{E}(\mathcal{H})\backslash\mathcal{E}_S$ intersecting S_i in one point and S_{i-1} in two points is at most $k^4(|S_{i-1}|+|S_i|)$ and the number of edges in $\mathcal{E}(\mathcal{H})\backslash\mathcal{E}_S$ intersecting S_i in two points and S_{i-1} in one point is also at most $k^4(|S_{i-1}|+|S_i|)$.

We prove the size of S_{i+1} is greater than $4k^3n^{(i+1)/k}$ by double counting pairs (v,h) where $v \in h \cap S_i$ and $h \in \mathcal{E}(\mathcal{H})$. The first count is simple:

$\sum_{v\in S_i} d_{\mathcal{H}}(v) \geq \delta|S_i|$. The other count is not as simple. First we show there are not too many edges of $\mathcal{H}$ falling completely within S_i. This, together with the induction hypothesis, implies many edges must intersect both S_i and S_{i+1}. Finally, we show that these edges must be sparsely located in S_{i+1} which will imply the proof of the Proposition.

First we count pairs (v, h) with $v \in h \cap S_i, h \in \mathcal{E}_S$. It is clear from the definition of approximate neighborhood structures, that the number of such pairs is at most:

$$|\cup_{j<i} S_j| + 2|S_i| + |S_{i+1}| \tag{2.2}$$

Next, we count pairs (v, h) with $v \in h \cap S_i, h \in \mathcal{E}(\mathcal{H})\backslash\mathcal{E}_S$, and $h \cap (\cup_{j<i} S_j) \neq \emptyset$. The induction hypothesis says that these edges contribute at most:

$$3k^4(|S_{i-1}| + |S_i|) \tag{2.3}$$

Now we are ready to count pairs (v, h) with $v \in h \cap S_i, h \in \mathcal{E}(\mathcal{H})\backslash\mathcal{E}_S$ and $h \subset S_i$. We define a graph G_0 on S_i. One by one, for each edge $h \in \{h \in \mathcal{E}(\mathcal{H})\backslash\mathcal{E}_S : h \subset S_i\}$, pick a two-set from h which has least multiplicity so far in G_0 to be in $E(G_0)$. Now applying Lemma 2.23 to G_0, we find that, ignoring multiplicities, $e(G_0) \leq 2|S_i|\binom{k}{2}$. Then we apply Lemma 2.21 to find that each edge in G_0 has multiplicity at most $\binom{2k-1}{2} + 2$. We conclude that $|\{h \in \mathcal{E}(\mathcal{H}) : h \subset S_i\}| = e(G_0) \leq 2|S_i|\binom{k}{2}(\binom{2k-1}{2} + 2)$. Thus the number of pairs (v, h) coming from such edges is at most:

$$6|S_i|\binom{k}{2}(\binom{2k-1}{2} + 2) \tag{2.4}$$

We now count pairs (v, h) with $v \in h \cap S_i, h \in \mathcal{E}(\mathcal{H}) \backslash \mathcal{E}_S$, and $|h \cap S_{i+1}| = 1$. Note that $\{h \in \mathcal{E}(\mathcal{H}) \backslash \mathcal{E}_S : h \cap S_i \neq \emptyset$ and $|h \cap S_{i+1}| = 1\} = \{h \in \mathcal{E}(\mathcal{H}) \backslash \mathcal{E}_S : |h \cap S_{i+1}| = 1$ and $|h \cap S_i| = 2\}$. Let $\mathcal{E}^* = \{h \in \mathcal{E}(\mathcal{H}) \backslash \mathcal{E}_S : |h \cap S_{i+1}| = 1$ and $|h \cap S_i| = 2\}$. We define a graph G_1 on S_i. For each edge $h \in \mathcal{E}^*$, let $e(h) = h \cap S_i$ be an edge in G_1. Let G_1^* be the simple graph obtained from G_1. As we do in Lemma 2.23, we can partition $E(G_1^*)$ into k parts such that there is no path of length $2k-1$ within any of the parts. Applying the theorem of Erdős and Gallai, (Theorem 1.9) we conclude that if G^+ is one of these k parts, then $e(G^+) \leq (k-1)|V(G^+)|$. We now apply the following theorem of Nash-Williams[86].

Theorem 2.29. *A graph G is decomposable into l forests iff for any $X \subset V(G)$, $e(G|_X) \leq l(|X|-1)$.*

Thus we can further partition each of these k parts into at most $k-1$ edge disjoint forests. In particular, we can partition $E(G_1^*)$ into no more than $k(k-1) < k^2$ edge disjoint forests. Let F_1 be the forest such that the number of edges of G_1 which are multiple with some edge of F_1 is greatest. We will bound the number of edges of G_1 by bounding the number of edges of G_1 that are multiples of edges of F_1.

To this purpose, we define an auxiliary bipartite graph G_A on parts $V_1 = S_{i+1}$ and $V_2 = E(F_1)$. Recall that for each $e \in E(G_1)$, there is an associated vertex $t = t(e) \in S_{i+1}$. For each edge $e \in E(G_1)$ where e is a multiple of an edge in F_1, add $(t(e), e)$ to the edges of G_A. Then the number of edges in

G_A equals the number of edges in G_1 that are multiples of edges of F_1. To count the number of edges in G_A, we will use a technique similar to that in Lemma 2.23.

We define partitions on $V_2 = E(F_1)$ related. To do this, we need to associate each edge $e \in E(F_1) = V_2$ to a vertex in S_i. For each component of F_1, pick a vertex in that component, and orient the edges of the component away from the distinguished vertex. Let $p, s : E(F_1) \to S_i$ be maps such that $\forall e \in E(F_1)$, $s(e)$ is the vertex towards which e is oriented, and $p(e)$ is the the vertex away from which e is oriented. Now for $j \leq i$, note that the nonempty sets $\pi^{i-j}(y)$ for $y \in S_{i-j}$ partition the vertex set S_i. For each such nonempty set, consider its inverse image under the map s; these resulting sets partition $V_2 = E(F_1)$. Call this partition $\mathcal{Q}_j$. We are now ready to prove the following claim which is similar to Lemma 2.23 in technique.

Claim 2.30. *Let $e_a, e_b \in E(F_1)$. Let l be maximal such that e_a and e_b are in the same partition class of $\mathcal{Q}_l$. If the minimal degree in G_A is at least k, then there can be no path of length $2(k - i + l)$ in G_A from e_a to e_b.*

Proof. Suppose not. Let P be such a path. Let $v_a = s(e_a)$ and $v_b = s(e_b)$. As l was chosen maximally and $v_a, v_b \in S_i$, we must have that l is the largest integer satisfying $\pi^{i-l}(u) = \pi^{i-l}(v)$. Thus there is a path of length $2(i - l)$ in $\mathcal{H}|_{\mathcal{E}_S}$ from v_a to v_b. We will use the path P in G_A to find a path from v_a to v_b in $\mathcal{H}|_{\mathcal{E}(\mathcal{H})\setminus\mathcal{E}_S}$ of length $2(k - i + l) + 1$ using only vertices in S_i and S_{i+1}. It is clear that once we produce such a path, we will have arrived at a contradiction as $\mathcal{H}$ is $\mathcal{C}_{2k+1}$-free.

Write $P = (e_a = e_1, u_1, e_2, u_2, \ldots, e_{k-i+l-1}, u_{k-i+l-1}, e_b = e_{k-i+l})$. Since F_1 is a forest, $\exists e_m \in \{e_1, e_2 \ldots, e_{k-i+l}\}$ such that $\forall j \leq k-i+l, s(e_j) \neq p(e_m)$. As the minimum degree in G_A is at least k, $\exists u^* \in N_{G_A}(e_m)$ with $u^* \notin \{u_1, u_2, \ldots, u_{k-i+l-1}\}$. Note that $(e_1 \cup u_1), (u_1 \cup e_2), \ldots, (u_{m-1} \cup e_m), (e_m \cup u^*), (e_m \cup u_m), (u_m \cup e_{m+1}), \ldots, (u_{k-i+l-1} \cup e_{k-i+l}) \in \mathcal{E}(\mathcal{H}) \backslash \mathcal{E}_S$ form a path from v_a to v_b on the following vertices of $S_i \cup S_{i+1}$:

$$s(e_1), u_1, s(e_2), u_2, \ldots, u_{m-1}, p(e_m), s(e_m), u_m, \ldots, u_{k-i+l-1}, s(e_{k-i+l})$$

We remark that if $m = 1$, then the order of the vertices changes slightly:

$$s(e_1), p(e_1), u_1, s(e_2), u_2, \ldots, u_{m-1}, p(e_m), s(e_m), u_m, \ldots, u_{k-i+l-1}, s(e_{k-i+l})$$

Thus we have found the desired path in $\mathcal{H}$ and reached our contradiction. □

Claim 2.31. $e(G_A) < k^2(|S_i| + |S_{i+1}|)$

Proof. Suppose not. Then the average degree in G_A is at least $2k^2$. Thus we can safely delete vertices from G_A of degree less than k^2 without lowering the average. Applying Claim 2.30 and then Lemma 2.24, we get a contradiction.

□

To summarize Claims 2.30 and 2.31, we have showed $|\mathcal{E}_1| = e(G_1) < k^4(|S_i| + |S_{i+1}|)$. This proves part (A) of Proposition 2.28. Since each edges in $\mathcal{E}_1$ intersects S_i in two points, the number of pairs (v, h) coming from the edges of $\mathcal{E}_1$ is at most:

$$2k^4(|S_i| + |S_{i+1}|) \tag{2.5}$$

We are now ready to count pairs (v, h) with $v \in h \cap S_i$ and $|h \cap S_{i+1}| = 2$. Let $\mathcal{E}_2 = \{h \in \mathcal{E}(\mathcal{H}) \backslash \mathcal{E}_S : |h \cap S_i| = 1 \text{ and } |h \cap S_{i+1}| = 2\}$. We will count these edges in almost exactly the same way we counted the edges of $\mathcal{E}_1$. Let G_2 be a graph on S_{i+1}. For $h \in \mathcal{E}_2$ let $e(h) = h \cap S_{i+1}$ be in $E(G_2)$. Let G_2^* be the simple graph obtained from G_2. As before, we decompose G_2^* into at most $k(k-1) < k^2$ edge disjoint forests. Let F_2 be the forest such that the number of edges in G_2 that are multiple with edges in F_2 is greatest. By bounding the number of edges of G_2 multiple with edges of F_2, we will get a bound on the total number of edges in G_2. Let G_B be an auxiliary bipartite graph on parts $V_1 = S_i$ and $V_2 = E(F_2)$. Recall that for each edge $e \in E(G_2)$, there is an associated vertex $t(e) \in S_i$. For each $e \in E(G_2)$ where e is a multiple of some edge in F_2, let $(e, t(e))$ be an edge of G_B. Then the number of edges in G_B is equal to the number of edges in G_2 which are multiple with some edge of F_2.

We orient the edges of F_2 as we did with F_1. For each component in F_2, pick a vertex and orient the edges of the component away from this distinguished vertex. Let $p, s : E(F_2) \to S_{i+1}$ be maps such that $\forall e \in E(F_2)$, $s(e)$ is the vertex towards which e is oriented and $p(e)$ is the vertex away from which e is oriented. We are now ready to prove a variant of Claim 2.30 and Lemma 2.23 for the bipartite graph G_B.

Claim 2.32. *Let $u, v \in S_i$. Let l be maximal such that $\pi^{i-l}(u) = \pi^{i-l}(v)$. If the minimum degree of G_B is at least k, then there can be no path of length $2(k-i+l)$ in G_B from u to v.*

Proof. Suppose not. Let P be such a path. Note that by assumption, there is a path of length $2(i-l)$ from u to v in $\mathcal{H}|_{\mathcal{E}_S}$. This second path lies completely within $(\cup_{j<i}S_j)$ except for the endpoints u and v. We will use the path P in G_B to find a third path in $\mathcal{H}|_{\mathcal{E}(\mathcal{H})\backslash\mathcal{E}_S}$ from u to v of length $2(k-i+l)+1$ contained entirely within $S_i \cup S_{i+1}$. The existence of such a path guarantees a $\mathcal{C}_{2k+1}$ in $\mathcal{H}$, a contradiction. Now we produce such a path.

Write $P = (u = u_1, e_1, u_2, e_2, \ldots, e_{k-i+l-1}, u_{k-i+l})$. As F_2 is a forest, $\exists e_m \in \{e_1, e_2 \ldots, e_{k-i+l}\}$ such that $\forall j \leq k-i+l, s(e_j) \neq p(e_m)$. As the minimum degree in G_B is at least k, $\exists u^* \in N_{G_B}(e_m)\backslash\{u_1, u_2, \ldots, u_{k-i+l}\}$. Now $(u_1 \cup e_1), (e_1) \cup u_2), \ldots, (u_m \cup e_m), (u^* \cup e_m), (e_m \cup u_{m+1}), \ldots, (e_{k-i+l-1} \cup u_{k-i+l}) \in \mathcal{E}(\mathcal{H})\backslash\mathcal{E}_S$ form a path in $\mathcal{H}$ of length $2(k-i+l)+1$ on the vertices:

$$u_1, s(e_1), u_2, \ldots, u_m, s(e_m), p(e_m), u_{m+1}, \ldots, s(e_{k-i+l-1}), u_{k-i+l}$$

Thus we have found the desired path, a contradiction. □

Claim 2.33. $e(G_B) < k^2(|S_i| + |S_{i+1}|)$

Proof. Suppose not. Then the average degree in G_B is at least $2k^2$. Thus we can safely delete vertices from G_B of degree less than k^2 without lowering the average. Applying Claim 2.32 and then Lemma 2.24 to get a contradiction. $\leftrightarrows$

In Claims 2.32 and 2.33, we have showed that $|\mathcal{E}_2| = e(G_2) < k^4(|S_i| + |S_{i+1}|)$. This proves part (B) of Proposition 2.28. As each of the edges in $\mathcal{E}_2$ intersects S_i in only one point, the number of pairs (v, h) coming from edges in $\mathcal{E}_2$ is at most:

$$k^4(|S_i| + |S_{i+1}|) \tag{2.6}$$

We are now ready to calculate the size of S_{i+1}. From Equations (2.2) through (2.6), we have:

$$\begin{aligned}\sum_{v \in S_i} d(v) \leq 3k^4(|S_{i-1}| + |S_i|)& \\ + |\cup_{j<i} S_j| + 2|S_i| + |S_{i+1}|& \\ + 6|S_i|\binom{k}{2}\left[\binom{2k+1}{2} + 2\right]& \\ + 2k^4(|S_i| + |S_{i+1}|)& \\ + k^4(|S_i| + |S_{i+1}|)& \\ \leq (3k^4 + k)|S_{i-1}|& \\ + \left[6k^4 + 6\binom{k}{2}\binom{2k-1}{2} + 12\binom{k}{2} + 2\right]|S_i|& \\ + (3k^4 + 1)|S_{i+1}|&\end{aligned} \tag{2.7}$$

As $\delta > 4k^4 n^{1/k} + 15k^4 + 10k^2$ is the minimum degree of $\mathcal{H}$, we have:

$$\begin{aligned}|S_{i+1}| &> \frac{|S_i|(\delta - 12k^4 - 10k^2) - |S_{i-1}|(3k^4 + k)}{3k^4 + 1} \\ &> 4k^3 n^{i/k}\end{aligned} \tag{2.8}$$

This proves part (c), the final part of Proposition 2.28. □

2.3 Non-Uniform Hypergraphs

We now turn to non-uniform hypergraphs avoiding cycles of a given length. Here, the only interesting question remaining given Theorem 2.15, is how large can a hypergraph be if it has arbitrarily large edges, and contains no $\mathcal{C}_{2k+1}$? Also, if we allow arbitrarily large edges, we are interested in not just the number of possible edges in such a hypergraph, but in a more exact notion of the size of the hypergraph. Namely, we would like to bound the size of $\sum_{e\in\mathcal{E}(\mathcal{H})}|e| = \sum_{v\in\mathcal{V}(\mathcal{H})} d(v)$. The following theorem gives an answer to this question which is consistent with the previous results.

Theorem 2.34 (Győri-Lemons[66]). *Fix $k > 1$ and let $\mathcal{H}$ be a (multi) hypergraph with all of its edges of size at least $4k^2$ such that $\mathcal{H}$ contains no cycle of length $2k + p$ where either $p = 0$ or $p = 1$. Then*

$$\sum_{v\in\mathcal{V}(\mathcal{H})} d(v) \leq 8k^4 n^{(k+1)/k} + (68k^5 + 24k^4)n$$

The proof of 2.34 broadly speaking uses the same technique as that in the proof of Theorem 2.2 - the bound for 3-uniform hypergraphs avoiding a given odd cycle. However, many of the technical details are different; in many ways this is a simpler proof as we can assume all the edges are quite large. Before we prove this theorem, we note a special case:

Theorem 2.35 (Győri-Lemons). *Let $\mathcal{H}$ be a hypergraph containing no $\mathcal{C}_4$. Then*

$$\sum_{e\in\mathcal{E}(\mathcal{H})} (|e| - 3) = O(n^{3/2})$$

Theorem 2.35 is an interesting special case of Theorem 2.34 because we have a completely different proof for Theorem 2.35. This second proof 2.35 is much simpler (being less general) but it also stands out as nice example of the substitution method. That is, the idea of the proof is to replace each hyperedge in the hypergraph with a suitable graph and then analyze the resulting graph.

We note that the constant factor we obtained for an upper bound in Theorem 2.35 is most likely not sharp - neither is the one in Theorem 2.34; we use a much weaker (larger) constant than for instance Lazebnik and Verstraëte (see Theorem 2.9). However, our result is more general in the sense that it covers non-uniform hypergraphs; and it seems to be difficult to achieve sharp bounds for non-uniform hypergraphs. Given the next theorem, it seems very likely to us that Theorem 2.35 could be extended to show similar bounds for hypergraphs avoiding a $\mathcal{C}_{2k}$ for fixed k. We also note that Theorem 2.35 is sharp up to the constant factor as the following construction similar in spirit to Construction 2.14 shows.

Construction 2.36. Let G be a complete $\mathcal{C}_4$-free bipartite graph on $V = (V_1, V_2)$ with $|V_1| = |V_2| = n/4$. We define a 4-uniform hypergraph, $\mathcal{H}$, in the following way. We define this on V_1 together with 3 copies of V_2. For each edge $e = (v_1, v_2)$ in G we take the corresponding 4-edge containing v_1 and the 3 copies of the v_2. It is clear from the construction that $\mathcal{H}$ has no $\mathcal{C}_4$. Moreover, $\mathcal{H}$ has approximately $(1/2)(n/4)^{3/2}$ edges.

Proof of Theorem 2.35. Let $\mathcal{H}$ be a $\mathcal{C}_4$-free hypergraph on n vertices. Without loss of generality, we may assume all the edges in $\mathcal{H}$ are of size at least 4. We will define a graph G on $\mathcal{V}(\mathcal{H})$ from $\mathcal{H}$ such that the number of edges in

G is proportional to $\sum_{e\in E(\mathcal{H})}(|e|-3)$. The idea is to replace each hyperedge with a star graph (a connected tree where all except one of the vertices have degree one) on the vertices of the hyperedge. For such star-graphs, we call the vertices of degree one, the *rays* of the star, and the remaining vertex the *center* of the star. We then give each star a unique color. Thus the problem will be reduced to finding the max number of edges in a regular graph with such a "star" coloring where there is no cycle on 4 colors. Of course such a cycle would correspond to a forbidden 4-cycle in $\mathcal{H}$. (Note that a cycle on 4 colors in G could be of length between 4 and 8.)

Claim 2.37. *We can replace each edge $h \in \mathcal{E}(\mathcal{H})$ where $|h| \geq 7$ with a star-graph of size at least $\frac{1}{2}(|h|-3)$ on the vertices of h such that each graph edge occurs with multiplicity 1. Here the size of the star-graphs refers to the number of edges (equivalently the number of rays) in the star.*

Proof. We replace the hyperedges of size at least 7 one by one with star graphs of unique color in the following way. Suppose we are replacing some hyperedge $h \in \mathcal{E}(\mathcal{H})$ with a star-graph. Let $k = |h|$. Now there may already be some graph edges defined on the k vertices of the hyperedge h. However, there cannot be too many: there can be no path (on these graph edges) of length 5 within h. Such a path would necessarily span at least 3 different colors and thus would correspond to a path (in $\mathcal{H}$) of length at least 3 within h. But then this is clearly (together with h) a 4-cycle in $\mathcal{H}$. Thus we can use the Erdős-Gallai Theorem (Theorem 1.9) to bound the number of such graph edges within h from above by

$$q\binom{5}{2}+\binom{r}{2} \tag{2.9}$$

where $k = 5q + r$, $0 \leq r \leq 4$. This implies that the number of pairs of vertices in h which are not yet graph edges is at least

$$\binom{k}{2}-\left(q\binom{5}{2}+\binom{r}{2}\right) \tag{2.10}$$

And we conclude that there are is at least one vertex within h with non-degree at least

$$\left(\frac{2}{k}\right)\left[\binom{k}{2}-\left(q\binom{5}{2}+\binom{r}{2}\right)\right] \geq k-5 \tag{2.11}$$

Here we are simply considering the graph edges defined so far within the hyperedge h. That is, we can always define a star of size $k-5$ without repeating any graph edges in h no matter how the previous stars were defined. Now as $|h| = k \geq 7$, we have $k-5 \geq \frac{1}{2}(k-3)$. So for each such edge h, we can find stars of size $\geq \frac{1}{2}(k-3)$ within h. □

Before we proceed, we need a definition.

Definition 2.38. Two vertices are *covered by color* i if they are both leaves of the star of color i.

Claim 2.39. *We can replace each edge $h \in \mathcal{E}(\mathcal{H})$ where $|h| \geq 7$ with a star-graph of size at least $\frac{1}{6}(|h|-3)$ on the vertices of h such that each graph edge occurs with multiplicity 1 and such that any pair of vertices is covered by at most 3 colors.*

Proof. Again, we replace the hyperedges one by one. Let h be an edge of size at least 7 in $\mathcal{H}$. From Claim 2.37 we can find a star-graph within h of size at least $\frac{1}{2}(k-3)$. We will find a substar of this star graph satisfying the conditions of Claim 2.39. We define a an auxiliary graph G_h on the $\frac{1}{2}(k-3)$ rays of the star. In G_h, two vertices are connected iff they are already covered three times. Then the components of G_h are either of size 3 or are stars. Otherwise, suppose there is a component of size at least 4 that is not a star. In this case, there is a path of length at least 3 within this component - that is, there is a 3-path within h on 3 different colors, a contradiction as $\mathcal{H}$ is quadrilateral-free. Thus there is an independent set in G_h of size at least $\frac{1}{3}|\mathcal{V}(G_h)|$. We then replace h with this substar consisting of the center and this independent set; this star has size at least $\frac{1}{3} * \frac{1}{2}(|e|-3)$, and no two leaves in the star are covered by more than three colors. □

We now replace each hyperedge in $\mathcal{H}$ of size at least 7 with such a star. For each hyperedge of size less than 7, we put in exactly one graph edge such that it does not overlap any already defined graph edges and so that the two vertices of this graph edge are not covered by more than three colors. It is clear that if the edge has size at least 3, both of these conditions can be satisfied. In this way the graph G is defined.

Now we estimate the size of the graph G. We can find a subgraph G_1 of G which is bipartite and has at least $\frac{1}{2}e(G)$ edges. Again, we can find a further subgraph G_2 of G_1 such that the centers of all the stars in G_2 are contained within one of the partition classes such that G_2 has at least $\frac{1}{2}e(G_1)$ edges. Call the partition class of G_2 containing the centers of the stars V_1 and let V_2 be the other partition class containing the rays of the stars. Now we claim that G_2 contains no $K_{2,5}$ with 5 vertices in V_1 and 2 in V_2. Otherwise, such a $K_{2,5}$ must contain a 4-colored $K_{2,2}$ since no two vertices in V_2 are covered by more than 3 colors. But such a 4-colored $K_{2,2}$ implies that our original hypergraph $\mathcal{H}$ had a 4-cycle. Thus there is no such $K_{2,5}$ in our graph. We are now prepared to estimate the number of edges in G_2. We know:

$$\sum_{x\in V_2}\binom{d(x)}{2} \leq 5\binom{|V_1|}{2} \tag{2.12}$$

and $|V_1| = |V_2|$ gives

$$\frac{n}{2}d(x)^2 \leq 5(\frac{n}{2})^2 \tag{2.13}$$

We conclude:

$$d(x) \leq (\frac{5n}{2})^{1/2} \text{ and } e(G_2) \leq (\frac{5}{8}n^3)^{1/2} \tag{2.14}$$

Thus:

$$\sum_{h\in\mathcal{H}}(|h|-3) \leq 12\sqrt{\frac{5}{8}}n^{3/2} \tag{2.15}$$

□

This proof utilizes the simple trick of replacing hypergraph edges with a nicely chosen graph (in this case a star-graph.) In this way we are able to convert a relatively hard hypergraph problem into a much easier graph problem. Unfortunately, it is unclear how to generalize this proof technique for other even cycles, or even how to proceed in the case of a forbidden $\mathcal{C}_6$.

Thus we consider a different approach; that of Theorem 2.2 to prove Theorem 2.34.

These proofs are not exactly the same; the main outline of the proof is the same, but the technical details in Theorem 2.34 are a bit more complicated. We will use the following lemma in the proof.

Lemma 2.40. *Let $\mathcal{H}$ be a multi-hypergraph with all of its edges of size at least $2l$ for some fixed constant l. Suppose also that $\mathcal{H}$ has no path on l vertices. Then the average degree in $\mathcal{H}$ is less than $3l$.*

Proof. We prove this by induction on the number of vertices of $\mathcal{H}$. Clearly, if $\mathcal{H}$ has n vertices ($l \leq n < 4l$) and the average degree is $3l$, there will be a path of length l in $\mathcal{H}$. Otherwise, let $n > 4l$ be the number of vertices of $\mathcal{H}$, and suppose that the average degree in $\mathcal{H}$ is at least $3l$. Let $\mathcal{P}$ be a longest path in $\mathcal{H}$. Let j be the length of this path (so $j < l$.) Let the E_1 be one of the tail edges of the path. Note that the number of vertices in E_1 not in the interior of the path is at least $|E_1| - (j-1) \geq 2l - j + 1 \geq l$. Each of these vertices is contained only within the edges of $\mathcal{P}$ - each has degree at most j. Pick l of these vertices and delete them and the edges of $\mathcal{P}$ which subsequently have less than $2l$ vertices from $\mathcal{H}$ (forming a new hypergraph $\mathcal{H}^1$). Then $\mathcal{H}^1$ has at most $j(3l-1)$ fewer incidences than does $\mathcal{H}$. And so $\sum_{e \in \mathcal{E}(\mathcal{H}^1)} |e| \geq \sum_{e \in \mathcal{E}(\mathcal{H})} |e| - j(3l-1) \geq 3ln - j(3l-1) \geq 3l(n-l)$. Thus the average degree in $\mathcal{H}^1$ is still at least $3l$ and by induction it must have a path of length l. $\square$

Proof of Theorem 2.34. Suppose the theorem is not true. Let $\mathcal{H}$ be a hypergraph on n vertices containing no $\mathcal{C}_{2k+p}$ where either $p = 0$ or $p = 1$. Suppose also that all the edges of $\mathcal{H}$ have size at least $4k^2$ and that $\sum_{h\in\mathcal{E}(\mathcal{H})} |h| > 8k^4n^{(k+1)/k} + (68k^5 + 24k^4)n$. We derive a contradiction by showing $\mathcal{H}$ must have more than n vertices.

First note that we may assume the minimum degree of $\mathcal{H}$ to be at least $2k^2n^{1/k}+17k^3+6k^2$. Otherwise we can delete those vertices of smaller degree and any edges that become smaller than $4k^2$ without decreasing the average degree in $\mathcal{H}$. The resulting hypergraph will still be a counterexample to Theorem 2.34 and will have minimum degree $\delta \geq 2k^2n^{1/k} + 17k^3 + 6k^2$ as desired.

Pick $\in \mathcal{V}(\mathcal{H})$ and let $(\{S_i\}, \mathcal{E}_S, \pi, \phi)$ be a greedily defined approximate k-neighborhood structure of $\mathcal{H}$ around x. For $1 \leq i \leq k$ let $\mathcal{E}_i = \phi^{-1}(S_i)$. Note that these sets partition $\mathcal{E}_S$.

Proposition 2.41. $|S_i| \geq n^{i/k}$ *for* $i = 0, 2, \ldots, k$.

Before we prove Proposition 2.41, we note that Theorem 2.34 easily follows from Proposition 2.41. As the sets S_i are pairwise disjoint, Proposition 2.41 implies $|\mathcal{V}(\mathcal{H})| > n$, a contradiction. □

Proof of Proposition 2.41. We prove this by induction on i. Clearly the base case holds. Suppose the proposition holds for some i, $0 \leq i < k$. We will show that $|S_i| \geq n^{1/k}|S_{i-1}|$. To do this we simply double count the pairs (v, h) where $v \in h \cap S_{i-1}$ and $h \in \mathcal{E}(\mathcal{H})$. We distinquish between four types of these pairs:

(A) pairs (v, h) with $h \in \mathcal{E}_j$ for $j < i$.

(B) pairs (v, h) with $h \in \mathcal{E} \backslash (\cup_{j \leq i} \mathcal{E}_j) = \mathcal{E} \backslash (\mathcal{E}_S)$.

(C) pairs (v, h) with $h \in \mathcal{E}_i)$ and $|h \cap S_{i-1}| > 2k^2$.

(D) pairs (v, h) with $h \in \mathcal{E}_i)$ not of type (C).

Note that that every pair (v, h) with $v \in h \cap S_{i-1}$ is exactly one of these four types. In particular, there are no edges $h \in \mathcal{E}_j$ for $j > i$ intersecting S_{i-1}. Also note that for pairs (v, h) of type (D) we have $|h \cap S_i| > 2k^2$. The heart of Proposition 2.41 lies in the following claims:

Claim 2.42. *There are no more than $14k^3|S_{i-1}|$ pairs (v, h) of type (A).*

Claim 2.43. *There are no more than $(3k^3 + 6k^2)|S_{i-1}|$ pairs (v, h) of type (B) and of type (C).*

Letting y represent the number of pairs of type (D), we find:

$$(17k^3 + 6k^2)|S_{i-1}| + y \geq \delta |S_{i-1}| \tag{2.16}$$

Which implies $y \geq [\delta - (17k^3 + 6k^2)]\,|S_{i-1}| \geq 2k^2 n^{1/k}|S_{i-1}|$. If (v, h) is a pair of type (D) then h intersects S_{i-1} in at most $2k^2$ places, thus there must be at least $n^{1/k}|S_{i-1}|$ edges h such that $\exists v \in h$ with (v, h) a pair of type (D). But clearly $|S_i| = |\mathcal{E}_i^*| \geq n^{1/k}|S_{i-1}|$. $\square$

Thus, all that remains to prove Proposition 2.41 is to prove the above two claims.

Proof of Claim 2.42. Fix $j < i$. Let $\mathcal{F} = \{h \in \mathcal{E}_j : |h \cap S_{i-1}| > 4k^2\}$. Clearly the number of pairs of type (A) is no more than

$$\sum_{h \in \mathcal{F}} |f \cap S_{i-1}| + 4k^2 |S_{i-1}| \tag{2.17}$$

Let $\mathcal{F}_1 = \{h \cap S_{i-1} : h \in \mathcal{F}\}$ and for $f \in \mathcal{F}_1$, and let $\hat{f}$ represent the edge from which f was derived in $\mathcal{F}$. Let $\mathcal{H}_1$ be the hypergraph consisting of the edges $\mathcal{F}_1$ and vertex set S_{i-1}. Clearly $\mathcal{H}_1$ may have multiple edges. We now attempt to find a linear bound (in terms of $|S_{i-1}|$) for $\sum_{f \in \mathcal{F}_1} |f|$.

Suppose without loss of generality that $\sum_{f \in \mathcal{F}_1} |f| \geq 4k^4 |S_{i-1}|$. Then we may delete those vertices of $\mathcal{H}_1$ of degree less than $2k^2$ and any edges that are smaller than $2k^2$ without decreasing the average degree of $\mathcal{H}$. As we will be bounding the size of $\mathcal{H}_1$ by a linear function of its vertex set, we can thus assume that $\mathcal{H}_1$ has minimum degree $2k^2$ and that all the edges are of size at least $2k^2$.

We now define an improper coloring of the vertices of $\mathcal{H}_1$ based on the vertices of S_j to which they are connected. For a vertex $v \in \mathcal{F}(\mathcal{H}_1)$, there are $m = m(v) \geq 2k^2$ edges $f_1, \ldots, f_m$ edges in $\mathcal{E}(\mathcal{H}_1)$ containing v. These derive from the edges $\hat{f}_1, \ldots \hat{f}_m \in \mathcal{F} \subseteq \mathcal{E}_j \subseteq \mathcal{E}(\mathcal{H})$. Thus there are vertices $w_l(v) = \phi^{-1}(\hat{f}_l) \in S_j$ and $u_l(v) = \pi(w_l) \in S_{j-1}$ for $1 \leq l \leq m$. We remark that the u_l's may not be pairwise distinct, however throughout we will be refering to the number of u_l with certain properties (or in certain sets) and in all of these cases, we mean the number of u_l including (counting) multiplicities.

We define sets $Q_0(v), \ldots, Q_{j-1} \subseteq S_{j-1}$ in the following way. Let $Q_0(v) = S_{j-1}$. Suppose we have defined Q_l for some l. Then pick $Q_{l+1}(v)$ from the family $\{\pi^{-[(j-1)-(l+1)]}(y) : y \in S_{l+1} \text{ and } \pi^{-[(j-1)-(l+1)]}(y) \subseteq Q_l(v)\}$ such that $|Q_{l+1} \cap \{u_1, \ldots, u_m\}|$ is maximal.

We are now ready to define the color of the vertex v. Let $l < j-1$ be the smallest integer such that $|Q_l \cap \{u_1, \ldots, u_m\}| \geq 2k(k-l)$ and such that more than $2k$ of the vertices $\{u_1, \ldots, u_m\}$ are contained in $Q_l \backslash Q_{l+1}$. Color v with color $l+1$. If no such integer l exists, then $|Q_{j-1} \cap \{u_1, \ldots, u_m\}| \geq 2k(k-(j-1)) \geq 4k$ and we color v color j.

We now color the edges of $\mathcal{H}_1$. Recall that for $f \in \mathcal{F}_1$, we have $|f| \geq 2k^2$. We define sets $Q_0(f), \ldots, Q_{j-1}(f)$ similarly to the prequel. Let $Q_0(f) = S_{j-1}$. Suppose we have defined $Q_l(f)$ for some $0 \leq l < j-1$. Let $Q_{l+1}(f)$ be the set occurring most frequently in the family $\{Q_{l+1}(v) : v \in f, Q_l(v) = Q_l(f)$ and the color of $v \geq m+2\}$. After finding these sets, we can now color the edge f. Let $m < j-1$ be the smallest integer such that $|\{v \in f : Q_{m-1}(v) = Q_{m-1}(f)\}| \geq 2k(k-(m-1))$ and $|\{v \in f : Q_{m-1}(v) = Q_{m-1}(f), v$ is of color m or, $Q_m(v) \neq Q_m(f)\}| \geq 2k$. If no such integer exists, let $m = j-1$. Color f with color m. Delete from the edge $f \in \mathcal{E}(\mathcal{H}_1)$ the vertices for which $Q_{m-1}(f) \neq Q_{m-1}(v)$. In this way we decrease the size of individual edges of $\mathcal{H}_1$, but not the number of vertices in $\mathcal{H}_1$. Altogether, we diminish the sum $\sum_{f \in \mathcal{F}_1} |f|$ by less than $4k^2|\mathcal{E}(\mathcal{H}_1)| \leq 4k^2|S_{j-1}|$. Also note that every edge will have at least $2k(k-(j-2)) \geq 4k$ vertices remaining after these deletions.

Now we claim that there is no path P on the edges of $\mathcal{H}$ of color l and of length $2(k-j+l-1)+p$. Otherwise let e_1 and e_2 be the first and last edges in this path respectively. As e_1 has at least $4k$ vertices, we can pick v_1 not in the path P. Now as e_2 has color l, we most have $|\{v \in e_2 : Q_{l-1}(v) = Q_{l-1}(e_2),\ v$ has color l, or $Q_m(v) \neq Q_m(e_2)\}| \geq 2k$. By choice of $Q_m(e_2)$, this implies there are at least $2k$ vertices v in e_2 with $Q_m(v) \neq Q_m(v_1)$. In particular we can find such a vertex v_2 not in the path P. Note that we have produced vertices v_1 and v_2 such that there is a path in $\mathcal{H}_1$ of length $2(k-j+l-1)+1$ between them and such that $Q_{l-1}(v_1) = Q_{l-1}(v_2)$ but $Q_l(v_1) \neq Q_l(e_2)$. We will now find a distinct path in $\mathcal{H}$ between these vertices of length $2(j-l+1)$.

By definition of $Q_l(v_1)$, we can find a vertex $u_1(v_1) \in Q_l(v_1)$ such that the associated edge $\pi^{-1}(\phi^{-1}(u_1))$, is not contained in the path P. Also as v_2 has color l, there must be at least $2k$ vertices $u(v_2) \in Q_{l-1}(v_2)$ such that $u \notin Q_l(v_1)$. In particular we can pick $u_2(v_2)$ to be such a vertex such that the associated edge $\pi^{-1}(\phi^{-1}(u_2))$ is not in the path P. Then it is easy to see that there is a path of length $2(j-l)$ between the vertices $u_1(v_1)$ and $u_2(v_2)$ in $\mathcal{H}$. The vertices of this path are: $u_1, \pi(u_1), \ldots, \pi^{j-l}(u_1) = \pi^{j-l}(u_2), \ldots, \pi(u_2), u_2$. The edges of this path are from the sets $\mathcal{E}_{j-l}, \ldots, \mathcal{E}_{j-1}$. But as the edges of the hypergraph $\mathcal{H}_1$ are derived from edges of $\mathcal{H}$ within the set $\mathcal{E}_j$, and as the vertices are distinct, we can combine the path P with this second path to get a cycle of length $2k+p$ in $\mathcal{H}$; a contradiction. Thus by Lemma 2.40, we conclude that the sum of the sizes of the edges of color l in $\mathcal{H}_1$ is strictly less than $3[2(k-j+l)+p-1]|S_{i-1}|$. Summing over l, we find

$$\sum_{f\in\mathcal{E}(\mathcal{H}_1)} |f| \le 6k^2|S_{i-1}| \tag{2.18}$$

We conclude that

$$\sum_{h\in\mathcal{F}} |h \cap S_{i-1}| \le 14k^2|S_{i-1}| \tag{2.19}$$

This concludes the proof of Claim 2.42 as summing over $j < i$, we have that the number of pairs (v, h) of type (A) is less than $14k^3|S_{i-1}|$ $\square$

The proof of Claim 2.43 closely mimics the proof of Claim 2.42. However, there are some technical details which are different; for the sake of clarity and completeness we include this proof here.

Proof of Claim 2.43. Let $\mathcal{F} = \{h \in \mathcal{E}(\mathcal{H}) : \exists v \in \mathcal{V}(\mathcal{H})$ where (v, h) is a pair of type (B) or (C)$\}$. Let $\mathcal{F}_2 = \{f \cap S_{i-1} : f \in \mathcal{F}\}$ and let cH_2 be the hypergraph on the vertex set S_{i-1} with edges $\mathcal{F}_2$. As before, we color the edges of $\mathcal{H}_2$.

For an edge $f \in \mathcal{E}(\mathcal{H}_2)$, let $Q_0(f) = S_{i-1}$. We define the sets $Q_1(f) \ldots, Q_{i-1}(f)$ in the following way. Suppose we have defined Q_l for some $0 \le l < i-1$. Then pick Q_{l+1} from within the family $\{\pi^{-[(i-1)-(l+1)]}(y) : y \in S_{l+1}$ and $\pi^{-[(i-1)-(l+1)]}(y) \subseteq Q_l(f)\}$ to maximize $|Q_{l+1}(f) \cap f|$. Having defined the sets $Q_m(f)$ for $0 \le m < i$, we are ready to assign a color to the edge f. Let l be the smallest integer such that $|Q_l(f) \cap f| \ge 2k(k-l)$ and $|f \cup (Q_l(f) \backslash Q_{l+1}(f))| \ge 2k$. Note as the edge f has size at least $2k^2$ such an integer must exist. Color the edge f with color l and delete from f those vertices not in $Q_l(f)$.

Claim 2.44. *Coloring (and deleting vertices from) each edge in this manner, we decrease the total size of the hypergraph $\mathcal{H}_2$ by less than $2k^2|\mathcal{E}(\mathcal{H}_2)|$.*

For fixed l, we claim that there can be no path of length $2(k - [(i - 1) - l]) + p$ in $\mathcal{H}_2$ on edges of color l. Suppose such a path P exists, and let e_1 and e_2 be the first and last edges respectively of the path. Then as each edge of $\mathcal{H}_2$ has size at least $4k$, we can find a vertex $v_1 \in e_1$ that is not contained in the path P. As e_2 has color l, there are at least $2k$ vertices of e_2 which do not belong to the set $\pi^{-[(i-1)-l+1]}(\pi^{(i-1)-(l+1)}(v_1))$. In particular, we can pick such a vertex v_2 from e_2 which does not belong to the path P. But then we have $\pi^{(i-1)-l}(v_1) = \pi^{(i-1)-l}(v_2)$ and $\pi^{(i-1)-(l+1)}(v_1) \neq \pi^{(i-1)-(l+1)}(v_2)$. This implies there is a path from v_1 to v_2 of length $2(i-1-l)$ within the vertices of the S_m for $m < i - 1$ and on the edges of $\mathcal{E}_m$ for $m < i - 1$. Namely, the path is: $v_1, \pi(v_1), \dots, \pi^{(i-1)-(l+1)}(v_1), \pi^{(i-1)-l}(v_1) = \pi^{(i-1)-l}(v_2), \pi^{(i-1)-(l+1)}(v_2), \dots, \pi(v_2), v_2$. But this together with the path P is a $2k + p$ cycle in $\mathcal{H}$, a contradiction. Thus we can apply Lemma 2.40 to find that the sum of the edge sizes of color l is strictly less than $3[2(k - (i - 1 - l)) + p]|S_{i-1}|$. Summing over all the edges in $\mathcal{H}_2$ we get:

$$\sum_{f \in \mathcal{E}(\mathcal{H}_2)} |f| \leq 6k^2|S_{i-1}| \tag{2.20}$$

As every edge in $\mathcal{H}_2$ has at least $4k$ vertices, there can be no more than $1.5k|S_{i-1}|$ edges in $\mathcal{H}_2$. This implies that in Remark 2.44, we diminished the size of $\mathcal{H}_2$ by less than $3k^3|S_{i-1}|$. Therefore, we have

$$\sum_{h\in\mathcal{F}} |h \cap S_{i-1}| \leq (3k^3 + 6k^2)|S_{i-1}| \tag{2.21}$$

Thus the number of pairs (v, h) of types (B) and (C) is no more than $(3k^3 + 6k^2)|S_{i-1}|$. $\square$

3 Extremal Set Systems

We look at a different type of problem - one that can be easily expressed in the language of forbidden substructures, but which has its origin in set theory. These problems are fundamentally different in flavor and in proof technique. Almost all of the problems which we will look at in this section are extensions, generalizations, and outgrowths of the celebrated Erdős-Ko-Rado[30] Theorem:

Theorem 3.1 (Erdős-Ko-Rado)**.** *Suppose $\mathcal{H}$ is a k-uniform hypergraph on n vertices such that any two members intersect in at least t points. Then*

$$e(\mathcal{H}) \leq \binom{n-t}{k-t}$$

holds for $n > n_0(k,t)$. Moreover, equality holds iff $\mathcal{H}$ consists of all the k-edges containing a fixed t-element set.

This seminal paper introduced the so called 'shifting technique' which has subsequently become a standard proof technique in extremal set theory. Erdős, Ko, and Rado showed that $n_0(k,1) = 2k$. Later, Frankl[41] found the exact bound for all $t \geq 15$ and Wilson[100] proved that for all t, $n_0(k,t) = (k-t+1)(t+1)$. This bound is best possible; for smaller values of n, there are bigger t-intersecting hypergraphs. For instance, let

$$\mathcal{H}_r = \{H \in \binom{[n]}{k} : |H \cap [t+2r] \geq t+r\}$$

Frankl[41] conjectured that the families $\mathcal{A}_r$ are the only (up to isomorphism) extremal t-intersecting families. It is easy to see that the family $\mathcal{H}_r$ is the largest if r satisfies:

$$(k-t+1)(2+\frac{t-1}{r+1}) \leq n < (k-t+1)(2+\frac{t-1}{r}) \tag{3.1}$$

Frankl and Füredi[49] proved the family $\mathcal{H}_r$ was the largest for r satisfying (3.1) for $n > (k-t+1)c\sqrt{t/\log t}$ and $t \geq 1+cr(r)$. Later this was improved by Ahlswede and Khachatrian[1] who showed it was true for all n.

All of these papers utilized the shifting technique as their main approach. Here we describe the basic idea. Suppose $\mathcal{H}$ is a t-intersecting hypergraph (the edges pairwise intersect in at least t points) and let $1 \leq i < j \leq n$. Let $P_{ij} : \mathcal{H} \to \binom{[n]}{k}$ be the mapping

$$P_{ij}(H) = \begin{cases} (H \backslash \{j\}) \cup \{i\} & \text{if } i \notin H, j \in H, (H \backslash \{j\}) \cup \{i\} \notin \mathcal{H}, \\ H & \text{otherwise} \end{cases} \tag{3.2}$$

A hypergraph $\mathcal{H}$ is said to be *left-shifted* if for all $1 \leq i < j \leq n$, $P_{ij}(\mathcal{H}) = \mathcal{H}$. It is easy to see that shifting a hypergraph preserves the property of being t-intersecting. That is, if $\mathcal{H}$ is t-intersecting, then so is $P_{ij}(\mathcal{H})$. The main idea of the above proofs is to show that if $\mathcal{H}$ is a t-intersecting, left-shifted hypergraph, then there exists an r such that for all $H_1, H_2 \in \mathcal{H}$, $|H_1 \cap H_2 \cap [t+2r]| \geq t$.

We can express the Erdős-Ko-Rado theorem using the language of Turán type problems easily: Let $\mathcal{H}_k^l = \{[k], [k-l+1, 2k-l]\}$, that is two k-sets which intersect in l points. Then Theorem 3.1 says that $ex(n, \mathcal{H}_k^{<t}) = \binom{n-t}{k-t}$ for $n > n_0(k,t)$.

3.1 Erdős-Ko-Rado Type Problems

An important generalization of the Erdős-Ko-Rado problem is the following: what can we say about $\mathcal{H}$ if we want to restrict the sizes of the pairwise intersections of $\mathcal{H}$? Let $L \subseteq [k-1]$, and suppose

$$\forall H_1, H_2 \in \mathcal{H},\ |H_1 \cap H_2| \in L \tag{3.3}$$

The Erdős-Ko-Rado Theorem tells us the maximal size of such a hypergraph $\mathcal{H}$ if the set of possible intersection sizes $L = \{t+1, \ldots, k-1\}$. In general, we write

$$m(n,k,L) = ex(n, \mathcal{H}_k^l : 0 \leq l < k,\ l \notin L)$$

If we only allow small intersections (as opposed to only large intersections in the Erdős-Ko-Rado Theorem) then the problem of finding the extremal hypergraph is a packing problem. Specifically, if $L = \{0, 1, \ldots, t-1\}$ then $\mathcal{H}$ satisfying 3.3 is called a (n,k,t)-packing. In 1964, Erdős and Hanani[29] conjectured that for fixed k and t,

$$m(n,k,\{0,1,\ldots,t-1\}) = (1-o(1))\binom{n}{t}/\binom{k}{t}$$

They proved this for $t = 2$ and for infinitely many values of k when $t = 3$. The conjecture was proven correct for all n, k, and t by Rödl[87] in 1985. The first general result for $m(n,k,L)$ for any L was found in 1975 by Ray-Chaudhuri and Wilson[20].

Theorem 3.2. *$m(n,k,L) \leq \binom{n}{s}$ if $s = |L|$.*

Later, in 1991, Alon, Babai, and Suzuki[4] proved Theorem 3.2 using the so called linear algebra method. This elegant proof technique has become quite important in the study of extremal systems. The idea is to associate appropriate linearly independent polynomials to each hyperedge in our hypergraph. Then, if we have chosen a nice finite dimensional polynomial space, and we know its dimension, we can get a bound on the number of edges in the hypergraph. Here we sketch Alon, Babai, and Suzuki's proof. We use the following notation. If f is a polynomial in n variables, $x_1, \dots, x_n$ then there is a multilinear polynomial f^* obtained from f by replacing each occurrence of x_i^j with x_i (for all $j \geq 2$ and all $1 \leq i \leq n$).

Proof. Associate to $H \in \mathcal{H}$ the polynomial

$$g_H = \prod_{l \in L} (\sum_{i \in H} x_i - l)$$

and let g_H^* be the corresponding linear polynomial. Now clearly for $H_1, H_2 \in \mathcal{H}$ we have $g_{H_1}^*(H_2) = 0$ if $H_1 \neq H_2$. Otherwise, $g_{H_1}^*(H_1) > 0$. (Here and from now on, we consider the edge H_i to represent the vector with components $x_j = 1$ iff $j \in H$ and $x_j = 0$ iff $j \notin H$). Let $h = \sum_{i \in [n]} x_i - k$. Clearly, $h(H) = 0$ for all $H \in \mathcal{H}$. Now for $I \subset [n]$ we write x_I for the polynomial $\prod_{i \in I} x_i$. It is easy to show that the set of polynomials

$$\{g_H^* : H \in \mathcal{H}\} \cup \{(x_I h)^* : I \subset [n], |I| \leq s - 1\}$$

is linearly independent. All of these polynomials are multilinear and thus live in a space of dimension $\sum_{i=0}^{s}\binom{n}{i}$. We are done as

$$|\{(x_I h)^* : I \subset [n], |I| \leq s-1\}| = \sum_{i=0}^{s-1}\binom{n}{i}$$

□

Theorem 3.2 was improved by Frankl and Wilson[56] who showed that if there is an integer valued polynomial f of degree d and a prime p such that for all $l \in L$, $p, |, f(l)$ but $p, \nmid, f(k)$ then

$$m(n,k,L) \leq \binom{n}{d}$$

On a similar note, Babai and Frankl[5] proved that if the greatest common divisor of L does not divide k then $m(n,k,L) \leq n$. In 1986, Frankl[43] showed that for all rationals $1 \leq r$, there exist k and L such that $m(n,k,L) = \Theta(n^r)$.

We now consider a different generalization of the Erdős-Ko-Rado Theorem, which can be considered a perturbation result. Namely, how large can a k-uniform intersecting hypergraph $\mathcal{H}$ be if we require $\cap_{H\in\mathcal{H}} H = \emptyset$? (Such hypergraphs are called *non-trivial.*) Let $\mathcal{H}^* = \{H \in \binom{[n]}{k} : 1 \in H, H \cap [2,k+1] \neq \emptyset\} \cup [2,k+1]$. Hilton and Milner[68] proved this is the maximal non-trivial intersecting family and moreover that for $k \geq 4$, it is the only extremal family.

Theorem 3.3. *If $\mathcal{H}$ is a k-uniform non-trivial intersecting hypergraph then*

$$|\mathcal{H}| \leq |\mathcal{H}^*| = \binom{n-1}{k-1} - \binom{n-k-1}{k-1} + 1$$

To emphasize the importance of the shifting method, we note that Frankl and Füredi[46] gave a short proof of this theorem using the shifting technique.

Note that in the Erdős-Ko-Rado Theorem, the construction is a 'star' hypergraph; there exists one vertex contained in every hyperedge. Similarly in the Hilton-Milner construction, there is one vertex contained in every hyperedge but one. We consider more results in this direction. The set $C \subset [n]$ is a *transversal* for the hypergraph $\mathcal{H}$ if for all $H \in \mathcal{H}, H \cap C \neq \emptyset$. The minimal size of a transversal is denoted $\tau(\mathcal{H})$ and is called the *transversal number* of the hypergraph $\mathcal{H}$. Then we can ask how big can an intersecting family be if it has transversal number τ? (Note that Theorem 3.1 answers this for $\tau = 1$ while Theorem 3.3 answers this for $\tau = 2$.) Frankl, Ota and Tokushige[51] answer the general question by considering the following function. For a hypergraph $\mathcal{H} \subset \binom{[n]}{k}$, and an integer $t \geq 1$, let

$$C_t(\mathcal{H}) = \{C \in \binom{[n]}{t} : \ \forall H \in \mathcal{H}, C \cap H \neq \emptyset\}$$

Of course, if t is less than the transversal number of $\mathcal{H}$, $C_{t(\mathcal{H})}$ will be empty. Now define

$$p_t(k) = \max\{|C_t(\mathcal{H})| : \ \mathcal{H} \subset \binom{[n]}{k} \text{ is intersecting and } \tau(\mathcal{H}) \geq t\}$$

Gyárfás[59] proved that $|C_t(\mathcal{H})| \leq k^t$ for hypergraphs $\mathcal{H}$ which are not necessarily intersecting. For $k \geq 2$, this bound is only achieved when $\mathcal{H}$ consists of t disjoint edges. Frankl, Ota, and Tokushige[51] found the correct values of $p_1(k), p_2(k)$, and $p_3(k)$. The significance of this function is explained by their theorem:

Theorem 3.4. *Let $\mathcal{F}$ be a k-uniform intersecting family with a minimal transversal of size τ with $k > k_0(\tau)$. Then*

$$|\mathcal{F}| \leq p_{\tau-1}(k)\binom{n-\tau}{k-\tau} + O(n^{k-\tau-1})$$

Thus they have determined the maximal size of hypergraphs with transversal number 3 and transversal number 4. Erdős and Lovász[33] constructed k-uniform hypergraphs with transversal number k of size $k!(\mathrm{e}-1)$. Lovász[80] conjectured this was maximal but Frankl, Ota, and Tokushige[51] have managed to construct a 4-uniform hypergraph with transversal number 4, and 42 edges (one more edge than in the Erdős-Lovaász construction). Frankl, Ota, and Tokushige conjecture that their construction is best possible, but it remains an open problem. They also conjecture that

Conjecture 3.5. $p_t(k) = k^t - \binom{t}{2}k^{t-1} + O(k^{t-2})$ *holds for* $k \geq k_0$.

3.2 The Unbalance of a Hypergraph

We now examine one last Erdős-Ko-Rado-type problem. Define the unbalance of a hypergraph $\mathcal{H}$ as $u(\mathcal{H}) = |\mathcal{H}| - d(\mathcal{H})$, where $d(\mathcal{H})$ is the maximal degree size in $\mathcal{H}$. The problem of finding a hypergraph with largest possible unbalance is something of a perturbation problem of the Erdős-Ko-Rado Theorem. We want to find a large intersecting hypergraph with relatively small maximum degree. Dinur and Friedgut[22] proved that if $\mathcal{H}$ is an intersecting hypergraph then $u(\mathcal{H}) = O(\binom{n-2}{k-2})$. The proof used analysis and Katona[71] asked if there was a simple purely combinatorial proof. Here we present the combinatorial proof from [78] of a sharp bound on the maximum unbalance of k-uniform intersecting families. Our proof relies heavily upon the classification of families based on their transversal number.

Observe that in the case of equality in Theorem 3.1 we have a family with unbalance equal to 0. Families with transversal number equal to 1 are referred to as *trivial* families. In [68], Hilton and Milner proved that the largest nontrivial k-uniform family has size $\binom{n-1}{k-1} - \binom{n-k-1}{k-1} + 1$ when $n \geq 2k$. It is worth noting that this family has only slightly larger unbalance; it has an unbalance of 1. We will use the following to construct families with maximal unbalance.

Definition 3.6. A $(2l+1)$-*kernal system* is the following family. Take a (2l+1) element subset of the ground set. This is the *kernel* of the family. The family consists of all k-sets containing at least $(l+1)$ vertices from the kernel.

The 3-kernel system has $3\binom{n-3}{k-2}+\binom{n-3}{k-3}$ sets and max degree $2\binom{n-3}{k-2}+\binom{n-3}{k-3}$; it has unbalance $\binom{n-3}{k-2}$. We will show that this is the maximal unbalance a k-uniform family can achieve. In fact, we will show that every family achieving this bound is either isomorphic to the 3-kernel system or to any sub-family of the 3-kernel system which contains all the k-sets intersecting the kernel in exactly 2 elements. Clearly adding or deleting those sets containing the entire kernel has no effect on the unbalance of the family. We prove that this is true for all $n > 6k^3$. For $3k-2 \geq n$ it is easy to check that the 5-kernel system has a larger unbalance than the the 3-kernel system.

We conjecture that the 3-kernel system is best for $n > 3k - 2$.

Theorem 3.7 (Lemons-Palmer[78]). *If $\mathcal{F} \subset \binom{[n]}{k}$ is intersecting and $n > n_0(k) = 6k^3$ the unbalance of $\mathcal{F}$ is $u(\mathcal{F}) \leq \binom{n-3}{k-2}$. Equality holds iff $\mathcal{F}$ is a subfamily of the 3-kernel system which contains all the k sets intersecting the vernal in two points.*

To prove the theorem, we investigate the cases when $\tau(\mathcal{F}) = 1$, $\tau(\mathcal{F}) = 2$, and $\tau(\mathcal{F}) \geq 3$. The second case will give us our 3-kernel system. We then refine this approach, considering both subcases $\tau = 3$ and $\tau \geq 4$ to improve our bound of $n_0(k)$. In [42], Frankl gives the exact bounds on the size of families with $\tau(\mathcal{F}) = 3$. Similarly, in [52], Frankl, Ota, and Tokushige found the exact bounds on the size of families with $\tau(\mathcal{F}) = 4$. However, we include our calculations here both for the sake of completeness (as the calculations are fairly simple) and because our results are true for all n.

Proof of Theorem 3.7. Let $\mathcal{F}$ be an intersecting family, and let T be a transversal of minimal size for $\mathcal{F}$. We consider three cases.

1. $|T| = 1$. Then there exists a vertex x that is contained in every $F \in \mathcal{F}$. So, $d(\mathcal{F}) = d(x) = |\mathcal{F}|$ implying $u(\mathcal{F}) = 0$.

2. $|T| = 2$. Let $T = \{x, y\}$. Let $\mathcal{F}_x = \{F \in \mathcal{F} : F \cap T = \{x\}\}$. Similarily, let $\mathcal{F}_y = \{F \in \mathcal{F} : F \cap T = \{y\}\}$. Now, $|\mathcal{F}| = d(x) + d(y) - d(xy)$ and clearly

$d(\mathcal{F}) \geq \max\{d(x), d(y)\}$. Thus,

$$\begin{aligned}
u(\mathcal{F}) &= |\mathcal{F}| - d(\mathcal{F}) \\
&\leq d(x) + d(y) - d(xy) - \max\{d(x), d(y)\} \\
&= \min\{d(x), d(y)\} - d(xy) \\
&= \min\{|\mathcal{F}_x|, |\mathcal{F}_y|\}
\end{aligned}$$

Let $\mathcal{F}_x - x$ and $\mathcal{F}_y - y$ be the families achieved by removing x, respectively y, from every set in $\mathcal{F}_x$ (respectively $\mathcal{F}_y$.) Clearly $|\mathcal{F}_x - x| = |\mathcal{F}_x|$ and $|\mathcal{F}_y - y| = |\mathcal{F}_y|$. Now $(\mathcal{F}_x - x)$ and $(\mathcal{F}_y - y)$ are $(k-1)$-uniform families on the ground set $X \backslash \{x, y\}$. As $\mathcal{F}$ is intersecting, we must have that $\forall \in \mathcal{F}_x - x$ and $\forall f \in \mathcal{F}_y - y, h \cap f \neq \emptyset$. That is, $\mathcal{F}_x - x$ and $\mathcal{F}_y - y$ are *cross-intersecting.* We can now apply a theorem of Frankl and Tokushige[53]:

Theorem 3.8. *Suppose $\mathcal{F}, \mathcal{G} \in \binom{X}{k}$, are cross-intersecting families with $|X| = n \geq 2k$. If $|\mathcal{F}| \geq \binom{n-1}{k-1} - \binom{n-k-1}{k-1} + 1$ and $\mathcal{F}$ is nontrivial, then $|\mathcal{G}| \leq \binom{n-1}{k-1} - \binom{n-k-1}{k-1} + 1$* □

If at least one of $(\mathcal{F}_x - x), (\mathcal{F}_y)$ is nontrivial, then we can apply this directly, so that,

$$\begin{aligned}
u(\mathcal{F}) &\leq \min\{|\mathcal{F}_x|, |\mathcal{F}_y|\} \\
&\leq \binom{(n-2)-1}{(k-1)-1} - \binom{(n-2)-(k-1)-1}{(k-1)-1} + 1 \\
&= \binom{n-3}{k-2} - \binom{n-k-2}{k-2} + 1 \\
&< \binom{n-3}{k-2}
\end{aligned}$$

On the otherhand, if both $\mathcal{F}_x$ and $\mathcal{F}_y$ are trivial and maximal, Theorem 3.1 implies that there is a z_1 and z_2 contained in all the sets of $\mathcal{F}_x - x$ and all the sets of $\mathcal{F}_y - y$ respectively. In this case we have $|\mathcal{F}_x| = |\mathcal{F}_y| = \binom{(n-2)-1}{(k-1)-1} = \binom{n-3}{k-2}$. Since $\mathcal{F}_x - x$ and $\mathcal{F}_y - y$ are cross-intersecting, z_1 must be the same as z_2 and thus $\mathcal{F}$ must be one of the subfamilies of the 3-kernel system with kernel $\{x, y, z_1\}$ which contains all the edges intersecting the kernel in two points.

3. $|T| \geq 3$. Let $x, y, z \in T$. By the minimality of $|T|$ for each $v \in T, \exists F \in \mathcal{F}$ such that $F \cap T = \{v\}$ (otherwise v need not be included in T.) Let $F, G \in \mathcal{F}$ such that $F \cap T = \{x\}$ and $G \cap T = \{y\}$. Let $L = F \cap G$. Now for each point $v \in L, \exists F_v \in \mathcal{F}$ such that $F_v \cap \{x, y\} \neq \emptyset$ and $v, z \notin F$. Otherwise each set from $\{F : F \cap \{x, y, z\} \neq \emptyset\}$ meets at least one point from $\{v, z\}$. But this means that x and y can be replaced by z to make a smaller transversal, contradicting the minimality of T. Thus the sets F_v must exist. Now we can find a simple upper bound for $d(z)$ by overcounting the number of sets that contain z.

Consider a set H containing z. H must intersect both F and G. We distinguish two subcases.

3.1. H intersects both F and G in $v \in L$. The point v can be chosen in $|L| < k$ ways. Now H must also intersect F_v (recall that $v \notin F_v$). The intersection of H and F_v can be chosen in $|F_v| = k$ ways. The remaining $k - 3$ points of H can be chosen in $\binom{n-3}{k-3}$ ways. This gives no more than $k^2\binom{n-3}{k-3}$ choices for H.

3.2. H intersects F and G in distinct points outside of L. The intersection of H and F can be chosen in $k - |L| < k$ ways. The intersection of H and G can be chosen in $k - |L| < k$ ways. The remaining $k - 3$ points of H can be chosen in $\binom{n-3}{k-3}$ ways. This gives no more than $k^2\binom{n-3}{k-3}$ choices for H.

In total we have $d(z) \le 2k^2\binom{n-3}{k-3}$. It should be noted that in general this will be a gross overcount. The choice of z from T was arbitrary, so $\forall v \in T, d(v) \le 2k^2\binom{n-3}{k-3}$. Clearly $|T| \le k$, so we have $|\mathcal{F}| \le \sum_{v\in T} d(v) \le 2k^3\binom{n-3}{k-3}$. For $n > 2k^3(k-2)+k$ we have $2k^3\binom{n-3}{k-3} < \binom{n-3}{k-2}$.

With a little closer analysis, we can improve $n_0(k)$ to $6k^3$. The above argument shows that if $\tau(\mathcal{F}) = 3$, then $|\mathcal{F}| < 6k^2\binom{n-3}{k-3}$. This is less than $\binom{n-3}{k-2}$ for $n > 4k^3 - 12k^2 + k$. Thus all we need to do is bound the size of $\mathcal{F}$ when $\tau(\mathcal{F}) \ge 4$.

Suppose $\mathcal{F}$ is intersecting with $\tau(\mathcal{F}) \ge 4$. Then there is a minimal transversal T of the family containing at least four elements. Let them be t_1, t_2, t_3, and t_4. By minimality of T, for $1 \le i \le 3$ there are edges in the family F_i such that $T \cap F_i = \{t_i\}$. We estimate the degree of t_4. Each edge, H containing t_4 must intersect all the F_i for $i \le 3$. There are 3 possibilities.

1. Suppose we pick a point $x \in \bigcap_{i\le 3} F_i$ to be in H (if such an x exists). Now x cannot replace t_1, t_2, and t_3, in the transversal T (otherwise T would not be minimal) so there must exist 3 edges G_i, $1 \le i \le 3$, meeting $\{t_1, t_2, t_3\}$ such that: $x, t_4 \notin G_i$, and $\bigcap_{i\le 3} G_i = \emptyset$. Otherwise all sets containing t_1, t_2, and t_3, also contain one of x, t_4 or a point $z \in \bigcap_{i\le 3} G_i$. But then we can replace t_1, t_2 and t_3 in the transversal with x and z, violating the minimality

of T. Now H must intersect each of the G_i; we must choose at least one point from one of the pairwise intersections and one from the remaining G_i. The remaining $k-4$ points of H we choose arbitrarily. In this way we can construct no more than $k^3\binom{n-4}{k-4}$ edges containing t_4.

2. Suppose we pick a point x in the intersection of exactly two of the sets F_i - say F_1 and F_2 (the two sets can be chosen in 3 different ways) to be in H. Now H must intersect F_3. Let it do so in the point y. Then as above, there must exist an edge G meeting $\{t_1, t_2, t_3\}$ such that $x, y, t_4 \notin G$. Otherwise x and y could replace t_1, t_2 and t_3 in T, a contradiction by the minimality of T. Now H must also intersect G. Thus the number of such edges is no more than $3k^3\binom{n-4}{k-4}$ edges containing t_4.

3. Suppose we pick points $x_i \in F_i \setminus F_j$ for $1 \leq i, j \leq 3$, and $j \neq i$ to be in H. We then choose the remaining vertices of H arbitrarily. Again we have the number of such edges is at most $3k^3\binom{n-4}{k-4}$ edges containing t_4.

Altogether we estimate that degree of t_4 to be no more than $7k^3\binom{n-4}{k-4}$. We conclude that $|\mathcal{F}| \leq 6k^4\binom{n-4}{k-4}$. We note that this is less than $\binom{n-3}{k-2}$ for $n > 6k^3$ as desired. □

It is not known if $n_0(k) = O(k^3)$ is best possible. An easy lower bound is $n_0(k) \geq 3k-2$. As mentioned above, the unbalance of $\mathcal{F}$, a 5-kernel system, is greater than that of the 3-kernel system for such values of n and k. We have $|\mathcal{F}| = 10\binom{n-5}{k-3} + 5\binom{n-5}{k-4}$ and $d(\mathcal{F}) = 6\binom{n-5}{k-3} + 4\binom{n-5}{k-4}$ yields $u(\mathcal{F}) = 4\binom{n-5}{k-3} + \binom{n-5}{k-4}$. This is larger than $\binom{n-3}{k-2}$ when $n < 3k-2$.

4 Further Questions

Here we consider some of the questions raised by our results and possibilities for further research. Most of our results concerned hypergraphs but there was one directly about graphs, namely Lemma 1.10. We previously mentioned that it would be interesting to know if this theorem is sharp. It is also interesting that we do not know what the corresponding statement should be regarding paths of even length. Another possible extension of our work concerning graphs is the following: our proof technique in section 2.2.2 might be translatable to graphs - and if so, perhaps could give a better bound to the Bondy-Simonivits Even Cycles Theorem. The best that we could hope for is an improvement in the bound by a factor of k but this would be interesting.

Concerning hypergraphs, the most interesting next step (regarding our main theorems) would be to prove the sharpness of our bounds - however it seems like this will be no easier than it has been in the case of graphs; a major breakthrough is needed. On the other hand, the Berge approach that we have followed here seems to be relatively unstudied (at least as Turán problems.) As far as we know, only the extremal hypergraphs avoiding Berge-cycles of a certain length have been studied so far. Certainly other graphs could yield interesting questions in the hypergraph setting as well. For instance, we are unaware of any results concerning the Berge-type generalization of complete graphs to hypergraphs. Finding the extremal hypergraphs containing no such Berge-type complete hypergraph may be easier than finding the extremal hypergraphs avoiding Turán-type complete hypergraphs. An even easier problem would be to determine the extremal size of hypergraphs with large edges containing no short path; in other words, the generalization of Lemma 2.40 should be a simple problem. Again, as far as we know, these problems have not been looked at yet, despite their simplicity. In general it would be interesting to know which Berge type Turán problems are degenerate and which are not. Clearly the extremal hypergraphs avoiding Berge-cycles is an extremely degenerate problem; we have upper bounds that are the same as the bounds for the corresponding degenerate problems in graph theory. However, it is easy to imagine there are problems that are only mildly degenerate; that is that there are k-uniform hypergraphs avoiding a Berge-type structure with $\Theta(n^r)$ edges where $k > r > 2$. For instance finding extremal k-uniform hypergraphs avoiding a Berge-type complete graph on l vertices with $k > l$ is almost certainly such a 'mildly degenerate' problem.

References

[1] Ahlswede Khachatrian, The complete intersection theorem for systems of finite sets, *European J. Combinatorics* **18**, (1997) n. 2, 125-136.

[2] Ajtai, M., Komlós, J., and Szemerédi, E. A note on Ramsey numbers, *J. Combinatorial Theory (A)* **29**, (1980) 354-360.

[3] Ajtai, M., Komlós, J. and Szemerédi, E. A dense infinite Sidon sequence, *European J. Combinatorics* **2**, (1981) 1-11.

[4] Alon, N., Babai, L. and Suzuki, H. Multilinear polynomials and Frankl-Ray Chaudhuri-Wilson type intersection theorems, *J. Combinatorial Theory (A)* **58**, (1991) 165-180.

[5] Babai, L. and Frankl, P. Note on set intersections, *J. Combinatorial Theory (A)* **28**, (1980), 103-105.

[6] Benson, C. T. Minimal regular graphs of girrths eight and twelve, *Canadian J. Math.* **18**, (1966) 1091-1094.

[7] Berge, C. Hypergraphs, Combinatorics of Finite Sets. North-Holland, Amsterdam, 1989.

[8] Berge, C. Hypergraphs, *Selected Topics in Graph Theory 3* (eds. Beineke, L. W, and Wilson, R. J.), Academic Press Limited, (1998), 189-207.

[9] Bollobás, B. Three graph without two triples whose symmetric difference is contained in a third, *Discrete Math.* **8**, (1974) 21-24.

[10] Bollobás, B., Győri, E. Pentagons vs. triangles, *Discrete Math.* (to appear.)

[11] Bondy, J. A. and Simonovits, M. Cycles of even length in graphs, *J. Combinatorial Theory (B)* **16**, (1974) 97-105.

[12] Brown, W. G. On graphs that do not contain a Thomsen graph, *Canad. Math. Bull.* **9**, (1966) 281-285.

[13] Brown, W. G. On an open problem of Paul Turán concerning 3-graphs, *Studies in Pure Math.*, Birkhäuser, Basel-Boston, Mass., (1983) 91-93.

[14] Brown, W. G., Erdős, P., and Sós, V. T. On the existence of triangulated spheres in 3-graphs, and related problems, *Period. Math. Hungar.* **3**, (1973) 221-228.

[15] Brown, W. G., Erdős, P., and Sós, V. T. Some extremal problems on r-graphs, *New Directions in the Theory of Graphs* (ed. F. Harary), Academic Press, New York, (1973) 53-63.

[16] de Caen, D. Extensions of a theorem of Moon and Moser on complete subgraphs, *Ars Combin.* **16**, (1983) 5-10.

[17] de Caen, D. Uniform hypergraphs with no blocks containing the symmetric difference of any two other blocks, *in* Proc. 16th Southeastern Conf. on Combinatorics, Graph Theory and Computing, *Congressus Numerantium* **47**, (1985) 249-253.

[18] de Caen, D. On upper bounds for 3-graphs without tetrahedra, *Congressus Numerantium* **62**, (1998) 193-202.

[19] de Caen, D. and Füredi, Z. The maximum size of 3-uniform hypergraphs not containing a Fano plane, *J. Combinatorial Theory (B)* **78**, (2000) 274-276.

[20] Ray-Chanduri, D. K. and Wilson, R. M. On t designs, *Osaka J. Math.* **12**, (1975) 737-744.

[21] Chung, F. and Lu, L. An upper bound for the Turán number $t_3(n, 4)$, *J. Combinatorial Theory (A)* **78**, (1999) 381-389.

[22] Dinur, I. and Friedgut, E. Intersecting families are essentially contained in Juntas. *Submitted* (2007)

[23] Erdős, P. On sequences of integers no one o which divides the product of two others and somee related problems, *Izvestiya Naustno-Issl. Inst. Mat. i Meh. Tomsk* **2**, (1938) 74-82.

[24] Erdős, P. Graph theory and probability, *Canadian J. Math.* **11**, (1959) 34-38.

[25] Erdős, P. Some recent results on extremal problems in graph theory, *Theory of Graphs* (ed. P. Rosenthal), Gordon and Breach, New York, and Dunod, Paris, (1967) 117-123.

[26] Erdős, P. On some new inequalitites concerning extremal properties of graphs, *Theory of Graphs* (ed. P. Erdős and G. Katona), Academic Press, New York, (1968) 77-81.

[27] Erdős, P. Some unsolved problems in graph theory and combinaotrial analysis, *Combinatorial Mathematics and Its Applications* (Proc. Conf. Oxford, 1969), Academic Press, (1971) 97-109.

[28] Erdős, P. On some problems in graph theory, combinatorical analysis and combinatorial number theory, *Graph Theory Combin.* (ed. B. Bollobás) 1-17.

[29] Erdős, P. and Hanani, H. On a limit theorem in combinatorial analysis, *Publ. Math. Debrecen* **10**, (1963) 10-13.

[30] Erdős, P., Ko, C., and Rado, R. An intersecting theorem for finite sets, *Quart. J. Math. Oxford, Ser. (2)* **12**, (1961) 313-320.

[31] Erdős, P. and Gallai, T. On maximal paths and circuits of graphs, *Acta Math. Acad. Sci. Hungar.* **10**, (1959) 337-356.

[32] Erdős, P., Hajnal, A., Simonovits, M., Sós, V. T., Szemeerédi, E. More results on Ramsey-Turán type problems, *Combinatorica* **3** (1983), n. 1, 69-81.

[33] Erdős, P. and Lovász, L. Problems and rsults on 3-chromatic hypergraphs and some related questions, *Infinite and Finite Sets (Proc. Col-*

loq. Math. Soc. J. Bolyai) **10** (ed. Hajnal, A.) Amsterdam: North-Holland (1975) 609-627.

[34] Erdős, P., Ko, C. and Rado, R. intersection theorems for systems of finite sets, *Quart. J. Math. Oxford (2)* **12**, (1961) 313-320.

[35] Erdős, P., Rényi, A., and Sós, V. T. On a problem of graph theory, *Studia Sci. Math. Hungar.* **1**, (1966) 215-233.

[36] Erdős, P. and Sós, V. T. Some remarks on Ramsey's and Turán's theorem, *Combinatorial Theory and its Applications, II* (ed. P. Erdős et al.), North-Holland, Amsterdam, (1970) 395-404.

[37] Erdős, P. and Spencer, J. *Probabilistic Methods in Combinatorics,* Academic Press, London - New York, Akadémiai Kiadó, Budapest, (1974).

[38] Erdős, P. and Stone, A. M. On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52**, (1946) 1087-1091.

[39] Erdős, P. and Simonovits, M. A limit theorem in graph theory, *Studia Sci. Math. Hungar.* **1**, (1966) 51-57.

[40] Faudree, R. J. and Schelp, R. H. Ramsey type results, *Infinite and Fine Sets, II* (ed. A. Hajnal et al.), North-Holland, Amsterdam, (1975) 657-665.

[41] Frankl, P. The Erdős-Ko-Rado theorem is true for $n = ckt$, *Combinatorics*, Proc. Fifth Hungarian Colloq. Combin., Keszthely, Hungary,

1976 (Hajnal, A. et al. Eds.), *Proc. Colloq. Math. Soc. János Bolyai* **18** (1978) North-Holland, Amsterdam, 365-375.

[42] Frankl, P. On intersecting families of finite sets, *Bull. Austral. Math. Soc.* **17**, (1980) 363-272.

[43] Frankl, P. All rationals occur as exponents, *J. Combinatorial Theory (A)* **42**, (1986) 200-206.

[44] Frankl, P. and Füredi, Z. A new geralization of the Erdös-Ko-Rado theorem, *Combinatorica* **3**, (1983) 341-349.

[45] Frankl, P. and Füredi, Z. Union-free hypergraphss and probability theory, *European J. Combinatorics* **5**, (1984) 127-131. Erratum, *ibid* **5**, (1984) 395.

[46] Frankl, P. and Füredi, Z. Non-trivial intersecting families, *J. Combinatorial Theory (A)* **41**, (1986) 150-153.

[47] Frankl, P. and Füredi, Z. An exact result for 3-graphs, *Discreete Math.* **50**, (1984) 323-328.

[48] Frankl, P. and Füredi, Z. Exact solution of some Turán-type problems, *J. Combinatorial Theory (A)* **45**, (1987) 226-262.

[49] Frankl, P. and Füredi, Z. Beyond the Erdős-Ko-Rado theorem, *J. Combinatorial Theory (A)* **56**, (1991) 182-194.

[50] Frankl, P., Füredi, Z. and Simonyi, G. *private communication*

[51] Frankl, P., Ota, K. and Tokushige, N. Covers in uniform intersecting families and a counterexample to a conjecture of Lovász, *J. Combinatorial Theory (A)* **74**, (1996) 33-42.

[52] Frankl, P., Ota, K. and Tokushige, N. Uniform intersecting families with covering number four, *J. Combinatorial Theory (A)* **71**, (1995) 127-145.

[53] Frankl, P. and Tokushige, N. Some inequalities concerning cross intersecting sets, *Combinatorics, Probability, and Computing* **7**, (1998) 247-260.

[54] Frankl, P. and Rödl, V. Hypergraphs do not jump, *Combinatorica* **4**, (1984) 149-159.

[55] Frankl, P. and Rödl, V. Some Ramsey-Turán type results for hypergraphs, *Combinatorica* **8**, (1988) 323-332.

[56] Frankl, P. and Wilson, R. M. Intersection theorems with geometric consequences, *Combinatorica* **8**, (1981) 357-368.

[57] Füredi, Z. New asymptotics for bipartite Turán numbers, *J. Combinatorial Theory (A)* **75**, (1996) 141-144.

[58] Giraud, G. Remarques sur deux problèmes extrémaux, *Discrete Math.* **84**, (1990) 319-321.

[59] Gyárfás, A. Partition covers and blocking sets in hypergraphs (in Hungarian). *MTA SZTAKI Tanulmányok* (1977) 71.

[60] Gyárfás, A., Jacobson, M. S., Kézdy, A. E. and Lehel, J. Odd cycles and Θ-cycles in hypergraphs, *Discrete Mathematics* **306**, (2006) 2481-2491.

[61] Győri, E. On the number of edge disjoint triangles in graphs of given size, *Combinatorics* (ed. A. Hajnal, L. Lovász, and V. T. Sós), Colloq. Math. Soc. János Bolyai, **52**, North-Holland, Amsterdam, (1988) 267-276.

[62] Győri, E. On the number of C_5's in a triangle free graph, *Acta Math. Acad. Sci. Hungar.* **10**, (1989) 101-102.

[63] Győri, E. On the number of edge disjoint cliques in graphs of a given size, *Combinatorica* **11**, (1991) 231-243.

[64] Győri, E. C_6-free bipartite graphs and the product representation of squares, *Graphs and combinatorics* (Marseille, 1995). *Discrete Math.* **165/166** (1997), 371-375.

[65] Győri, E. Triangle-Free Hypergraphs, *Combinatorics, Probability and Computing* **15**, (2006) 185-191.

[66] Győri, E and Lemons, N. Hypergraphs avoiding cycles of a given length, *in preperation*

[67] Győri, E. and Li, H. A note on the number of triangles in $\mathcal{C}_{2k+1}$ graphs, *in preperation*

[68] Hilton, A. J. W., and Milner, E. C. Some intersection theorems for systems of finite sets, *Quart. J. Math. Oxford (2)*, **18**, (1967) 369-384.

[69] Imrich, W. Explicite construction of graphs without small cycles, *Combinatorica*, **4**, (1984) 53-59.

[70] Katona, G. O. H. Extremal problems for hypergraphs, *Combinatorics* Vol II (ed. M. Hall et al.), Math. Centre Tracts **56**, Amsterdam, (1974) 13-42.

[71] Katona, G.O.H. *private communication*

[72] Katona, G.O.H., Nemetz, T., and Simonovits, M. On a problem of Turán in the theory of graphs, *Math. Lapok* **15**, (1964) 228-238.

[73] Kostochka, A. A class of constructions for Turán's (3,4)-problem, *Combinatorica* **2**, (1982) 187-192.

[74] Kostochka, A. and Verstraëte, J. Even cycles in hypergraphs, *J. Combinatorial Theory (B)* **94**, (2005) 173-182.

[75] Kővári, T., Sós, V. T., and Turán, P. On a problem of Zarankiewicz, *Colloq. Math.* **3**, (1954) 50-57.

[76] Lazebnik, F., Verstraëte, J. On hypergraphs of girth five, *Electronic J. Combinatorics* **10**, (2003) 15 (Research Paper 25).

[77] Lazebnik, F., Ustimenko, V. A., and Woldar, A. J. A new series of dense graphs of high girth, *Bull. Amer. Math. Soc. (N.S.)* **32**, (1995), no.1, 73-79.

[78] Lemons, N. and Palmer, C. The unbalance of set systems, *Graphs and Combinatorics* accepted for publication.

[79] Lubotzky, A. Philips, R., and Sarnak, P. Ramanujan graphs, *Combinatorica* **8**, (1988) 261-277.

[80] Lovász, L. On minimax theorems of combinatorics, *Mathematikai Lapok* **26**, (1975) 369-384.

[81] Mantel, W. Problem 28, *Wiskundige Opgaven* **10**, (1907) 60-61.

[82] Margulis, G. A. Explicite construction of graphs without short cycles and low density codes, *Combinatorica* **2**, (1982) 71-78.

[83] Mubayi, D. On hypergraphs with every four points spanning at most two triples, *Electronic J. Combinatorics* **10**, (2003) (10).

[84] Mubayi, D. and Rödl, V. On the Turán number of triple systems, *J. Combinatorial Theory (A)* **100**, (2002) 136-152.

[85] Motzkin, T.S. and Straus, E. G. Maxima for graphs and a new proof of a theorem of Turán, *Canadian J. of Math.* **17**, (1965) 535-540.

[86] Nash-Williams, C. St. J. A. Decomposition of finite graphs into forests. *J. London Math. Soc.* **39**, (1964) 12.

[87] Rödl, V. On a packing and covering problem, *European J. Combinatorics* **6**, (1985) 69-78.

[88] Ruzsa, I. and Szemerédi, E. Triple systems with no six points carrying three triangles, *Combinatorics, II* (ed. A. Hajnal and V. T. Sós), North-Holland, Amsterdam, (1978) 939-945.

[89] Sidorenko, A. F. On the maximal number of edges in a uniform hypergraph that does not contain prohibited subgraphs, *Math. Notes* **41**, (1987) 247-259. Russian orginal: *Mat. Zametki* **41**, No. 3, 433-455.

[90] Sidorenko, A. F. Asymptotic solutions for a new class of forbidden r-graphs, *Combinatorica* **9**, (1989) 207-215.

[91] Simonovits, M. A method for solving extremal problems in graph theory, *Theory of Graphs* (ed. P. Erdős and G. Katona), Academic Press, New York, (1968) 279-319.

[92] Simonovits, M. Extremal graph problems with symmetrical extremal graphs, additional chromatic conditions, *Discrete Math.* **7**, (1974) 349-376.

[93] Simonovits, M. Extremal graph theory, *Selected Topics in Graph Theory 2* (ed. L. W. Beineke and R. J. Wilson), Academic Press, New York, (1983) 161-200.

[94] Sos, V. T. Some remarks on the connection of graph theory, finite geometry and block designs, *Teorie Combinatorie*, Tomo II., Accad. Naz. Linzei, Roma, (1976) 2223-233.

[95] Talbot, J. Chromatic Turán problems and a new upper bound for the Turán density of K_4^-, *European J. Combinatorics* **28**, (2007) 2125-2142.

[96] Todorov, D. T. On some Turán numbers, *Mathematics and mathematical education* B'lgar. Akad. Nauk, Sofia, (1984) 179-186.

[97] Turán, P. On an extremal problem in graph theory, *Mat. Fiz. Lapok* **48**, (1941) 436-452.

[98] Turán, P. On the theory of graphs, *Colloq. Math.* **3**, (1954) 19-30.

[99] Wenger, R. Extremal graphs with no C_4's, C_6's, or C_{10}'s, *J. Combinatorial Theory (B)* **52**, (1991) 113-116.

[100] Wilson, R. M. The exact bound in the Erdős-Ko-Rado theorem, *Combinatorica* **4**, (1984) 247-257.