Universitat Politècnica de Catalunya
Facultat de Matemàtiques i Estadística

Master in Advanced Mathematics and Mathematical Engineering

Master's thesis

# Finding Partite Hypergraphs Efficiently

**Ferran Espuña Bertomeu**

Supervised by Richard Lang

June 2025

Thanks to...

## Abstract

Turán-type problems are a central theme in extremal (hyper)graph theory, asking how many edges a $k$-graph $H$ can have, as a function of its number of vertices $n$, without containing a specific subgraph $G$. All graphs have a constant $\pi(G)$ called the *Turán density*, such that this number is $(\pi(G) + o(1))\binom{n}{k}$. This thesis deals with degenerate Turán problems ($\pi(G) = 0$) by focusing on finding complete balanced $k$-partite $k$-subgraphs, denoted $K\left(t, \overset{k}{\ldots}, t\right)$, within a $k$-uniform hypergraph ($k$-graph).

Classical existence theorems by Kővari, Sós, Turán (for $k = 2$) and Erdős (for $k \geq 2$) guarantee that $k$-graphs with constant edge density contain $K\left(t, \overset{k}{\ldots}, t\right)$ where $t$ grows with the number of vertices $n$ (typically $t = \Omega\left((\log n)^{1/(k-1)}\right)$). These proofs are non-constructive, and locating such dynamically sized subgraphs efficiently is challenging as brute-force search becomes superpolynomial.

This thesis presents a deterministic, polynomial-time algorithm that bridges this gap in the constant density case. Given a $k$-graph with $n$ vertices and $m$ edges, our algorithm finds a $K\left(t, \overset{k}{\ldots}, t\right)$ where $t$ explicitly depends on $n$, $k$, and the density $d = m/n^k$, matching the $\Omega((\log n/\log(1/d))^{1/(k-1)})$ scaling from existence proofs. Our method generalizes the bipartite case work by Mubayi and Turán, using a recursive strategy on link graphs. This provides an efficient, constructive proof for the existence of these growing $K\left(t, \overset{k}{\ldots}, t\right)$ structures, demonstrating they are algorithmically findable in polynomial time.

### Keywords

hypergraph, algorithm, graph, partite, extremal

# Contents

# 1. Introduction

Graph theory provides fundamental tools for modeling relationships and networks across diverse fields. A natural and powerful extension of graphs is the concept of *hypergraphs*, where edges can connect more than two vertices. Specifically, a $k$-uniform hypergraph, or $k$-graph, consists of a set of vertices and a collection of edges, each being a $k$-element subset of the vertices (for example, a 2-graph is simply a graph). These structures arise naturally in areas ranging from combinatorics and computer science to data analysis and computational biology.

A central theme in graph theory, extending readily to hypergraphs, is extremal (hyper)graph theory. This field seeks to understand the maximum or minimum size of a combinatorial structure satisfying certain properties. For instance, *Turán-type problems* ask how many edges a graph can have, as a function of its number of vertices $n$, without containing a specific subgraph $G$. The maximum such number of edges is called the *Turán number* and is denoted $\mathrm{ex}(n, G)$. A key result is Turán's Theorem [24], which determines $\mathrm{ex}(n, K_r)$ for all $n$ and $r \geq 2$, where $K_r$ is the complete graph on $r$ vertices. Furthermore, the Erdős–Stone–Simonovits Theorem [11] provides asymptotic estimates for $\mathrm{ex}(n, G)$ for any graph $G$ as $n \to \infty$, as a function of the chromatic number of $G$.

These theorems do not extend directly to hypergraphs, as the combinatorial structures become significantly more complex. The asymptotic behavior of $\mathrm{ex}(n, G)$ as $n \to \infty$ is often characterized by the *Turán density*, defined as

$$\pi(G) = \lim_{n\to\infty} \frac{\mathrm{ex}(n, G)}{\binom{n}{k}}.$$

A fundamental result distinguishes between two regimes: $\pi(G) = 0$ if and only if $G$ is $k$-partite (meaning its vertices can be partitioned into $k$ sets such that no edge has two vertices in the same set). Determining the exact value of $\pi(G)$ for non-$k$-partite $k$-graphs when $k > 2$ is a notoriously difficult open problem for many families, including even small hypergraphs like the complete 3-graph on 4 vertices, $K_4^{(3)}$ [15, 21].

This thesis focuses on the complementary *degenerate* case, where $\pi(G) = 0$. We are particularly interested in the problem of forbidding complete balanced $k$-partite $k$-graphs, denoted $K\left(t, \overset{k}{\ldots}, t\right)$, which consists of $k$ disjoint sets of $t$ vertices each, and all $t^k$ edges formed by selecting one vertex from each set. The classical Kővari–Sós–Turán Theorem [17, 14] provides the following upper bound for $k = 2$.

$$\mathrm{ex}(n, K(s, t)) = \mathcal{O}\left(n^{2-\frac{1}{\min\{s,t\}}}\right).$$

Erdős [8] generalized this type of bound to higher uniformities, showing that

$$\mathrm{ex}\left(n, K\left(t, \overset{k}{\ldots}, t\right)\right) = \mathcal{O}\left(n^{k-\frac{1}{t^{(k-1)}}}\right) \tag{1}$$

for $k \geq 2$.

Upper bounds for Turán numbers, like (1), often involve counting or probabilistic arguments, which are inherently non-constructive. They guarantee the existence of the desired subgraph $K\left(t, \overset{k}{\ldots}, t\right)$ in dense enough hypergraphs but do not typically provide an efficient algorithm to *find* such a subgraph. If we focus on a fixed guest $k$-graph $G = (V, E)$ and let the size $n$ of the host $k$-graph $H$ grow, this is not a problem. A brute-force search over all ordered sets of $|V|$ vertices in $H$ would yield a polynomial-time algorithm for finding a copy of $G$ in $H$, provided that we know that $H$ has more that $\mathrm{ex}(n, G)$ edges. However, the situation becomes more complex when we consider the case where $G$ is not fixed, but also grows with $n$.

We focus on the case where the edge density of $H$ is fixed, and $G = K\left(t, \overset{k}{\dots}, t\right)$ for some $t$ that grows with $n$. Careful analysis in the proof of the Erdős bound shows that this guarantees that $G$ is a subgraph of $H$ for some

$$t = \Omega\left((\log n)^{1/(k-1)}\right). \tag{2}$$

Running a greedy search checking all $\binom{n}{kt}$ sets of $kt$ vertices in $H$ would then **not** yield a polynomial-time algorithm, because $\binom{n}{kt}$ grows superpolynomially with $n$.

The main contribution of this thesis is bridging this gap by providing an efficient algorithmic solution. We develop and analyze a deterministic, polynomial-time algorithm that, given a $k$-graph $G$ with $n$ vertices and at least $dn^k$ edges, finds a complete balanced $k$-partite subgraph $K\left(t, \overset{k}{\dots}, t\right)$ within $G$, where

$$\left\lfloor \left(\frac{\log\left(n/2^{(k-1)}\right)}{\log(3/d)}\right)^{\frac{1}{k-1}} \right\rfloor,$$

matching the order of magnitude of (2). This result not only provides a constructive proof for the upper bounds of the type established by Erdős, but in fact reaches the best possible value of $t$ up to a constant factor depending on $d$, as can be show by probabilistic arguments (see Proposition 3.19 and the beginning of Section 4, where the algorithm is introduced). Our algorithm generalizes the approach used by Mubayi and Turán for the bipartite case ($k = 2$) [19]. It employs a recursive strategy that mirrors the inductive proof structure of the Erdős bound, iteratively reducing the uniformity $k$ by constructing appropriate link graphs.

This thesis is organized as follows. Section 2 formally introduces some families of hypergraphs (including complete $k$-partite hypergraphs), as well as basic operations like restrictions, links, and blow-ups. We also use this section to introduce asymptotic notation, which is used throughout the thesis. Section 3 provides an overview of relevant theoretical results for Turán-type problems, proving central theorems like the Turán Theorem (Theorem 3.2), the Kővari–Sós–Turán Theorem (Theorem 3.12), the mentioned Erdős bound (Theorem 3.14) and a more precise version of it (Theorem 3.16), thus showing that finding a complete balanced $k$-partite $k$-graph of part sizes in the order of (2) exists $H$ when it has constant positive density. Section 4 presents our main algorithm (Algorithm 2), provides a rigorous proof of its correctness and analyzes its polynomial runtime complexity (Theorem 4.1). Finally, Section 5 summarizes the main results of this thesis, and lays down some open problems for future research.

# 2. Notation and Basic Definitions

In this section, we introduce some basic definitions and results that are used throughout this thesis. We start with some preliminaries on hypergraphs, which are the main objects of study in this thesis.

## 2.1 Hypergraphs

**Definition 2.1.** For an integer $k \geq 1$ a *$k$-uniform hypergraph* (or *$k$-graph*, for short) is a tuple $G = (V, E)$ where $V$ is a finite set and $E \subset \binom{V}{k}$. We call the elements of $V(G) = V$ its *vertices* and those of $E(G) = E$ its *edges*. The value $k$ is called the *uniformity* of $G$.

*Remark* 2.2. In the definition above, if we let $k = 1$, we get a set of 1-sets of the vertex set $V$, which we can identify with a subset of $V$. If we let $k = 2$, we recover the usual definition of an undirected graph with no loops.

The following definition is a generalization of the notion of degree of a vertex in a graph.

**Definition 2.3.** Let $G = (V, E)$ be a $k$-graph and $v \in V$. The *degree* $d_G(v)$ of $v$ in $G$ is the number of edges containing $v$, that is

$$d_G(v) = |\{e \in E \mid v \in e\}|.$$

A useful operation is restricting a k-graph to a subset of its vertices. This yields a new k-graph, called the *subgraph induced by* the subset, which has the same uniformity.

**Definition 2.4.** Let $G = (V, E)$ be a $k$-graph and $T \subset V$. The *restriction* of $G$ to $T$ is the $k$-graph

$$G[T] = (T, E_T),$$

where

$$E_T = \{e \in E \mid e \subset T\}.$$

The following operation also lets us obtain graphs of a different uniformity from a subset of vertices of a $k$-graph.

**Definition 2.5.** Let $G = (V, E)$ be a $k$-graph. Let $1 \leq j \leq k-1$ be an integer and let $T \subset V$ be a set of vertices satisfying $k - j \leq |T|$. The *common $j$-link graph* of $T$ is the $j$-graph $L_G(T; j) = (V \setminus T, E')$, where

$$E' = \left\{ Y \in \binom{V \setminus T}{j} \,\middle|\, X \cup Y \in E \text{ for all } X \in \binom{T}{k-j} \right\}.$$

Figure 1 illustrates how to construct a common $j$-link graph from a $k$-graph $G$ in the case $k = 3$ and $j = 2$.

Next, we introduce $k$-graph homomorphisms, embeddings and isomorphisms, which allow us to relate $k$-graphs of the same uniformity to each other.

**Definition 2.6.** Let $G = (V, E)$ and $H = (W, F)$ be $k$-graphs and let $f : V \to W$ be a map between their vertex sets. If $A \subset E$ is a set of edges in $G$, we denote

$$f(A) = \{f(e) \mid e \in A\} = \{\{f(v) \mid v \in e\} \mid e \in A\}.$$

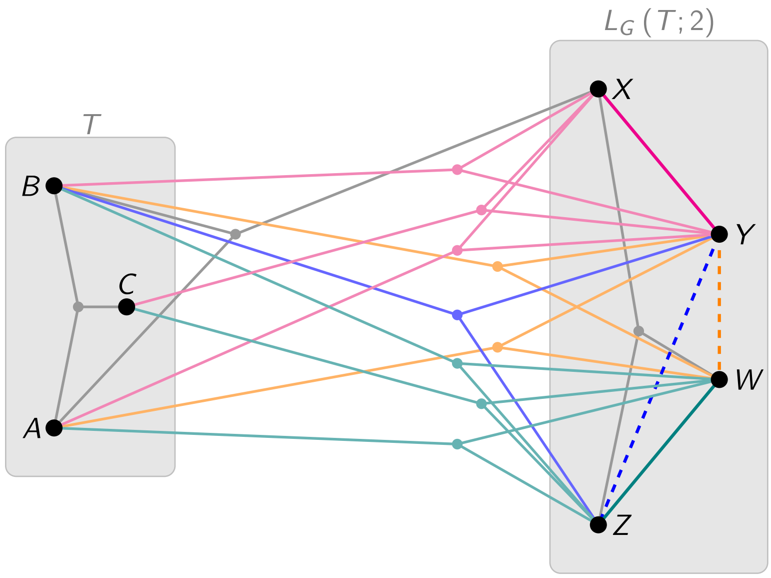


Figure 1: A 3-graph $G$ and the common 2-link graph $L_G(T;2)$ of the set $T = \{A, B, C\}$. Vertices are represented as black dots, and 3-edges of $G$ are represented as colored or gray small dots, connected by a line to the vertices they contain. Colored dots correspond to 3-edges with exactly $k-j = 3-2 = 1$ vertices in $T$, which are the only ones that can contribute to the common 2-link graph. Edges in the common 2-link graph are represented as solid lines connecting the corresponding vertices, in the same color as the 3-edges they come from. Dashed lines correspond to edge pairs in $V \setminus T$ that have some of the required 3-edges in $G$, but not all of them. The resulting link graph has vertex set $\{X, Y, Z, W\}$ and edge set $\{\{X, Y\}, \{W, Z\}\}$.

Then, $f$ is a *homomorphism* from $G$ to $H$ if

$$f(E) \subset E(H[f(V)]). \tag{3}$$

If such a homomorphism exists and is injective, we say that $f$ is an *embedding* of $G$ on $H$ and that $H$ *contains* $G$ as a subgraph. We write $G \subset H$. Otherwise, we say that $H$ is $G$*-free*.

If $f$ is an embedding of $G$ in $H$ and

$$f(E) = E(H[f(V)]), \tag{4}$$

we say that $f$ is an *induced* embedding and that $H$ contains $G$ as an *induced* subgraph. We write $G \subset_{\text{ind}} H$. If, in addition, $f$ is a bijection, we say that $f$ is an *isomorphism* and that $G$ is *isomorphic* to $H$. We write $G \cong H$.

*Remark* 2.7. Condition (3) of Definition 2.6 implies that $f$ is injective when restricted to each edge in $E$, because $G$ and $H$ have the same uniformity. However, it does not necessarily imply that $f$ is injective on all of $V$.

It can be checked that isomorphism of $k$-graphs is an equivalence relation. Furthermore, it is compatible with the subgraph ($\subset$) and induced subgraph ($\subset_{\text{ind}}$) relations, which are preorders. This allows us to discuss whether a $k$-graph $G$ is an (induced) subgraph of another $k$-graph $H$, *up to isomorphism* in both the *guest* $k$-graph $G$ and the *host* $k$-graph $H$. For more details, see Appendix A.

So far, we have not seen any concrete examples of $k$-graphs or their isomorphism classes. We now introduce an important family of them.

**Definition 2.8.** A $k$-graph $G = (V, E)$ is *complete* if $E = \binom{V}{k}$. We denote $G = K_V^{(k)}$.

If $K$ and $K'$ are complete $k$-graphs with the same number of vertices $r$, any bijection $f : V(K) \to V(K')$ is clearly an isomorphism between $K$ and $K'$. This allows us to talk, up to isomorphism, about *the* complete $k$-graph on $r$ vertices $K_r^{(k)}$. For example, in Figure 2 we show the complete 3-graph $K_4^{(3)}$.

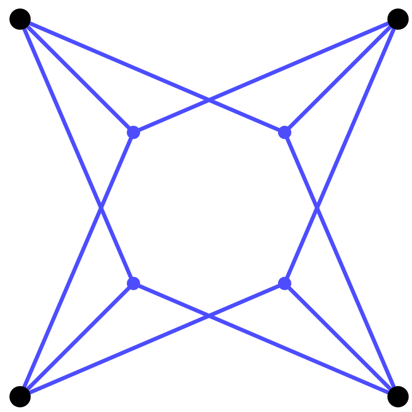

Figure 2: A complete 3-graph on 4 vertices.

*Remark* 2.9. A $k$-graph $H = (V, E)$ contains $G = K_r^{(k)}$ as a subgraph if and only if, for some subset $T \subset V$ of size $r$, (namely, the image of an embedding of $G$) $H[T] \subset H$ is complete. Such an embedding is always induced, as it is given by the identity map on $T$.

## 2.2 Partite Hypergraphs

One way to impose structure on a $k$-graph is to require that it is *partite*.

**Definition 2.10.** for an integer $r \geq k$, a $k$-graph $G = (V, E)$ is *$r$-partite* (or $(r, 1)$-*colorable*) if there exists a partition $V = V_1 \cup \cdots \cup V_r$ such that every edge $e \in E$ intersects every part $V_i$ in at most one vertex. We may write $G = (V_1, \ldots, V_r; E)$ and say that $G$ is a partite $k$-graph on $V_1, \ldots, V_r$. If $r$ is the minimum integer such that $G$ is $r$-partite, we say that $r = \chi_1(G)$ is the *partiteness* of $G$.

The above definition is a special case of the family of chromatic numbers of hypergraphs $\chi_\gamma(G)$, where the condition is that an edge of the hypergraph can intersect each part in at most $\gamma$ vertices [18]. If we set $\gamma = k - 1$, we recover the usual notion of chromatic number $\chi(G)$, in which we impose that edges are not fully contained in any part. In general, this is much weaker, but the two notions are the same when $k = 2$.

*Remark* 2.11. These chromatic numbers are monotone non-decreasing with respect to the subgraph relation. Indeed, if $f : G \to H$ is an embedding of $k$-graphs, and $H$ is $(r, \gamma)$-partite with parts $V_1, \ldots, V_r$, then $G$ is $(r, \gamma)$-partite with parts $f^{-1}(V_1), \ldots, f^{-1}(V_r)$. This is because $f$ is injective, so it preserves the number of vertices in each part for every edge. This in turn means that $\chi_\gamma(G) \leq \chi_\gamma(H)$.

If $G = (V_1, \ldots, V_k; E)$ is a $k$-partite $k$-graph, every edge intersects every part in exactly one vertex. This means that we can identify the edges with a subset of $V_1 \times \cdots \times V_k$. If it is clear from context, we may slightly abuse notation when talking about ordered and unordered sets of vertices, as in the definition below.

**Definition 2.12.** A $k$-partite $k$-graph $G = (V_1, \dots, V_k; E)$ is *complete* if $E = V_1 \times \cdots \times V_k$. That is, if all $(v_1, \dots, v_k) \in V_1 \times \cdots \times V_k$ satisfy $\{v_1, \dots, v_k\} \in E$. We denote $G = K(V_1, \dots, V_k)$.

In some cases, it is useful to generalize this notation to partite $k$-graphs where the number of parts is different from $k$.

**Definition 2.13.** Let $r \geq k \geq 1$. An $r$-partite $k$-graph $G = (V_1, \dots, V_r; E)$ is *complete* if

$$E = \bigcup_{\{i_1,\dots,i_k\} \in \binom{[r]}{k}} V_{i_1} \times \cdots \times V_{i_k}.$$

We denote $G = K^{(k)}(V_1, \dots, V_r)$.

If $V_1, \dots, V_r$ and $W_1, \dots, W_r$ are disjoint sets and $|V_i| = |W_i| = a_i$ for all $i$, then

$$K^{(k)}(V_1, \dots, V_r) \cong K^{(k)}(W_1, \dots, W_r).$$

An isomorphism is given by any bijection $f : V \to W$ (where $V = \bigcup_i V_i$, $W = \bigcup_i W_i$) such that $f(V_i) = W_i$ for all $i$. This allows us to talk, up to isomorphism, about *the* complete $r$-partite $k$-graph with part sizes $a_1, \dots, a_r$, which we denote by

$$K^{(k)}(a_1, \dots, a_r),$$

or, in the $k$-partite case, by

$$K(a_1, \dots, a_k) = K^{(k)}(a_1, \dots, a_k).$$

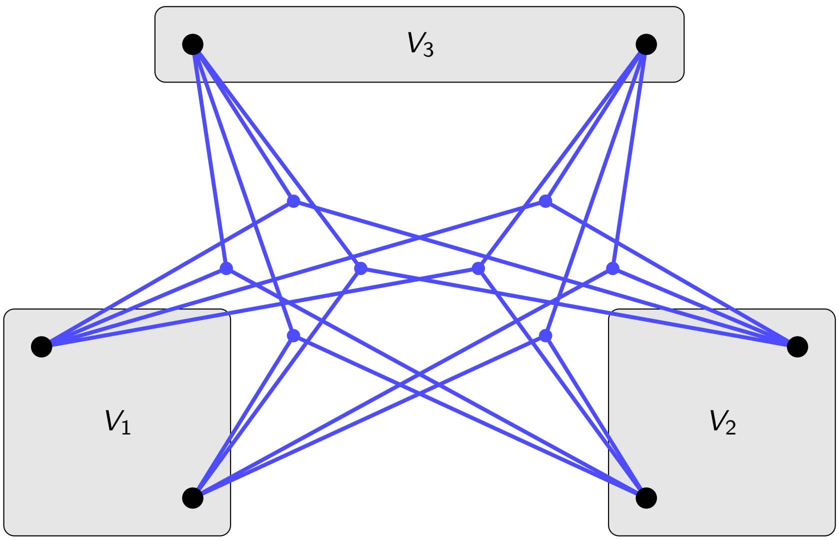


Figure 3: The complete 3-partite 3-graph $K(2,2,2)$, with parts $V_1, V_2, V_3$.

Figure 3 shows the complete 3-partite 3-graph with 2 vertices in each part, $K(2,2,2) = K^{(3)}(2,2,2)$. In contrast, Figure 4 shows the complete 3-partite 2-graph $K^{(2)}(2,2,2)$.

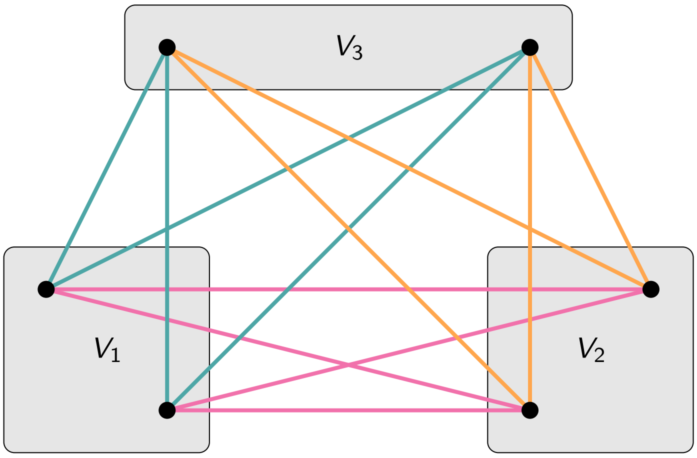


Figure 4: The complete 3-partite 2-graph $K^{(2)}(2,2,2)$, with parts $V_1, V_2, V_3$.

## 2.3 Blowing Up

Another interesting operation we can do with $k$-graphs is the following.

**Definition 2.14.** Let $G = (\{v_1, \dots, v_p\}, E)$ be a $k$-graph and let $t$ be a positive integer. The *t-blowup* $G(t)$ of $G$ is the $p$-partite $k$-graph

$$G(t) = (V_1, \dots, V_p; E')$$

where $|V_i| = t$ and

$$E' = \bigcup_{v_{i_1}, \dots, v_{i_k} \in E} V_{i_1} \times \cdots \times V_{i_k}.$$

In essence, we replace each vertex with a set of $t$ vertices, and each edge with a complete $k$-partite $k$-graph on the corresponding parts. For example, Figure 4 displays the 2-blowup of the triangle graph $K_3^{(2)}$, with the edges corresponding to the three original edges of the triangle shown in different colors. Any complete $p$-partite $k$-graph with equal part sizes can be seen as a blowup of a complete $k$-graph. In particular, the complete $k$-graph $K\left(t, \overset{k}{\dots}, t\right)$ (see Figure 3 for an example) is often known as a *blowup of an edge*, because it is a $t$-blowup of the $k$-graph on $k$ vertices with a single edge $K_k^{(k)}$. Similarly to the case of complete $k$-graphs and complete $k$-partite $k$-graphs, the $t$-blowup of a $k$-graph is unique up to isomorphism.

## 2.4 Asymptotic Notation

Throughout this thesis, we will often describe the growth rate of functions whose domain is the set of positive integers $\mathbb{N}$. Unless otherwise stated, the functions we consider map to non-negative real numbers $\mathbb{R}_{\geq 0}$. We now introduce some standard asymptotic notations.

**Definition 2.15.** Let $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a function. We define $\mathcal{O}(f)$ to be the set of functions $g : \mathbb{N} \to \mathbb{R}_{\geq 0}$ such that there exist a constant $C > 0$ and an integer $n_0 \in \mathbb{N}$ for which

$$g(n) \leq C \cdot f(n) \quad \text{for all } n \geq n_0.$$

As a common shorthand, we write $g(n) = \mathcal{O}(f(n))$ to mean $g \in \mathcal{O}(f)$. This notation is also used in arithmetic expressions. For example, if $h$ is the function $n \mapsto n^2$, an expression like $e^n + \mathcal{O}\left(n^2\right)$ denotes the set of functions $\{n \mapsto e^n + k(n) \mid k \in \mathcal{O}(h)\}$.

**Definition 2.16.** Let $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a function. We define $\Omega(f)$ to be the set of functions $g : \mathbb{N} \to \mathbb{R}_{\geq 0}$ such that there exist a constant $c > 0$ and an integer $n_0 \in \mathbb{N}$ for which

$$g(n) \geq c \cdot f(n) \quad \text{for all } n \geq n_0.$$

As shorthand, we write $g(n) = \Omega(f(n))$ for $g \in \Omega(f)$.

**Definition 2.17.** Let $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a function. We define $\Theta(f)$ to be the set of functions $g : \mathbb{N} \to \mathbb{R}_{\geq 0}$ such that there exist constants $c_1 > 0, c_2 > 0$ and an integer $n_0 \in \mathbb{N}$ for which

$$c_1 \cdot f(n) \leq g(n) \leq c_2 \cdot f(n) \quad \text{for all } n \geq n_0.$$

Equivalently, $\Theta(f) = \mathcal{O}(f) \cap \Omega(f)$, assuming $f, g$ are non-negative. As shorthand, we write $g(n) = \Theta(f(n))$ for $g \in \Theta(f)$.

**Definition 2.18.** Let $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a function. We define $o(f)$ to be the set of functions $g : \mathbb{N} \to \mathbb{R}_{\geq 0}$ such that for every constant $\varepsilon > 0$, there exists an integer $n_0 \in \mathbb{N}$ for which

$$g(n) \leq \varepsilon \cdot f(n) \quad \text{for all } n \geq n_0.$$

If $f(n) > 0$ for all $n \geq n_0'$ (for some $n_0'$), this condition is equivalent to $\lim_{n \to \infty} \frac{g(n)}{f(n)} = 0$. As shorthand, we write $g(n) = o(f(n))$ for $g \in o(f)$.

Unless otherwise specified, these notations describe the limiting behavior as $n \to \infty$. They are crucial for comparing growth rates, especially when exact formulas are complex or unknown.

# 3. Hypergraph Turán Problems

## 3.1 Turán-Type Problems

Now we can state the *forbidden subgraph problem* for $k$-graphs. Informally, given a $k$-graph $G$, and an integer $n \geq |V(G)|$, we want to find the smallest $M_0$ such that all $k$-graphs with $n$ vertices and $m > M_0$ edges contain $G$ as a subgraph.

**Proposition 3.1.** *Let $G = (V, E)$ be a $k$-graph with nonempty edge set and $n \geq |V|$ be an integer. Then there exists an integer $M_0 = ex(n, G) \in \left[0, \binom{n}{k}\right)$ such that the condition*

*"All $k$-graphs with $n$ vertices and $m$ edges contain $G$ as a subgraph."*

*is true for all $\binom{n}{k} \geq m > M_0$ and false for all $0 \leq m \leq M_0$.*

*Proof.* Note that, if such an $M_0$ exists, clearly it is unique. Also, the condition is clearly false for $m = 0$ and true for $m = \binom{n}{k}$ (the only $k$-graph $H$ with vertex set $W$, $|W| = n$ and $\binom{n}{k}$ edges is the one having all $k$-sets of vertices so any injective map $f : V \to W$ is an embedding of $G$ in $H$). We only need to show that if the condition is true for $m$ then it is true for all $m' \geq m$. Suppose it is true for $m$ and let $m' \geq m$. Let $H = (W, F)$ be a $k$-graph with $n$ vertices and $m'$ edges. We can take $F' \subset F$ with $|F'| = m$. By hypothesis, the $k$-graph $H' = (W, F')$ contains $G$ as a subgraph, and the identity map in $W$ is an embedding of $H'$ in $H$. Then, $G \subset H' \subset H$ implies $G \subset H$ by transitivity of the embedding relation (Proposition A.1).

□

We call the integer $\mathrm{ex}(n, G)$ the *Turán number* of $G$ on $n$ vertices. The Turán number is increasing both in $n$ and under $k$-graph inclusion. To see this, suppose that $H$ is a $k$-graph with $n$ vertices and $\mathrm{ex}(n, G)$ edges. The first property can be seen by adding to $H$ a new vertex $v$ with no edges containing it, obtaining $H'$. If $f : G \to H$ is an embedding and $v$ has a preimage $x$, it must be a vertex in $G$ with degree 0, so it can be replaced by any other vertex $w \neq v$ in $v(H)$ outside the image of $f$. Restricting this new map to $H$, we get that $G \subset H$, in contradiction to our assumption. For the second, suppose that $G \subset G'$. Because $H$ is $G$-free, it is also $G'$-free, which means that it has at most $\mathrm{ex}(n, G')$ edges. Therefore, $\mathrm{ex}(n, G) \leq \mathrm{ex}(n, G')$. As a consequence of this, we also get that the Turán number is invariant under isomorphism of $G$.

There are very few $k$-graphs $G$ for which an exact formula for $\mathrm{ex}(n, G)$ is known. Of these, the most famous family of examples are the complete 2-graphs $K_r^{(2)}$, for which Turán numbers were first studied by Turán [24] in 1941. The result is the following.

**Theorem 3.2** (Turán's Theorem)**.** *Let $r > 2$ be an integer and let $n \geq r$. Let $a_1, \ldots, a_{r-1}$ be integers such that $a_1 + \cdots + a_{r-1} = n$ and $\lfloor n/(r-1) \rfloor \leq a_i \leq \lceil n/(r-1) \rceil$ for all $i$. Then*

$$ex\left(n, K_r^{(2)}\right) = \sum_{\{x,y\} \in \binom{[r-1]}{2}} a_x \cdot a_y \tag{5}$$

*Furthermore, if $G$ is a 2-graph with $ex\left(n, K_r^{(2)}\right)$ edges and $G$ does not contain $K_r^{(2)}$ as a subgraph, then*

$$G \cong K^{(2)}(a_1, \ldots, a_{r-1}).$$

Before the proof of Turán's theorem, we introduce two lemmas.

**Lemma 3.3.** *Let $G = (V, E)$ be a 2-graph with $n$ vertices and $\mathrm{ex}\left(n, K_r^{(2)}\right)$ edges. If $x, y \in V$ are different vertices and $\{x, y\} \notin E$, then $d_G(x) = d_G(y)$.*

*Proof.* We argue by contradiction. Suppose, without loss of generality, that $d_G(x) > d_G(y)$. We argue that we can construct a 2-graph $G'$ with $n$ vertices and more edges than $G$ that does not contain $K_r^{(2)}$ as a subgraph, contradicting the definition of the Turán number.

The new graph $G' = (E', V')$ is constructed from $G$ by removing from $V$ the vertex $y$ (and all edges containing it) and adding a copy $x'$ of $x$, connected to the same vertices (that is, $\{x', v\} \in E'$ if and only if $\{x, v\} \in E$). Clearly, $|V'| = |V|$ and $|E'| = E - d_G(y) + d_G(x) > |E|$. To see that $G'$ does not contain $K_r^{(2)}$ as a subgraph, suppose that $G'[T']$ is complete for some $T' \subset V'$ of size $r$. Because $\{x, x'\}$ is not an edge in $G'$, $T'$ cannot contain both $x$ and $x'$. Because the edges not containing $x'$ are the same as in $G$, which contains no $K_r^{(2)}$, we deduce that $T'$ contains $x'$ and therefore does not contain $x$. Now, let $T = (T' \setminus \{x'\}) \cup \{x\} \subset V$, also of size $r$. We argue that the graph $G[T] = G'[T]$ must be complete, reaching a contradiction. If it were not, then there would exist $v \in T, v \neq x$, such that $\{x, v\} \notin E$. This implies that $\{x', v\} \notin E'$, but $v \in T' \setminus \{x'\}$, which contradicts our assumption that $G'[T']$ is complete. $\square$

**Lemma 3.4.** *Let $G = (V, E)$ be a 2-graph with $n$ vertices and $\mathrm{ex}\left(n, K_r^{(2)}\right)$ edges. Then, $G$ is a complete $p$-partite graph for some $p \geq 2$.*

*Proof.* Equivalently, we show that the relation defined by non-adjacency on $V$ (that is, $x \sim y$ when $\{x, y\} \notin E$) is an equivalence relation, so we can divide $V$ into equivalence classes by this relation, which means that $\{x, y\} \in E$ if and only if they are in different parts.

The reflexivity and symmetry of the relation are clear. Suppose, by way of contradiction, that there exist $x, y, z \in V$ such that $x \sim z$ and $y \sim z$, but $x \nsim y$. We now construct a different graph $G'$ with the same number of vertices as $G$ that also does not contain $K_r^{(2)}$ as a subgraph, reaching a contradiction. $G'$ is constructed from $G$ by removing the vertices $x$ and $y$ (and all the associated edges) and adding the two new vertices $z_1$ and $z_2$ and the edges $\{\{v, z_i\} \mid \{v, z\} \in E, i \in \{1, 2\}\}$.

First, we show that $G'$ contains no embedding of $K_r^{(2)}$. We make a similar argument as in the proof of Lemma 3.3. By way of contradiction, suppose that $G'[T']$ is complete for some $T' \subset V'$ of size $r$. Because $z$, $z_1$ and $z_2$ pairwise non-edges of $G'$, only one of them can be an image of a vertex in $K_r^{(2)}$. However, $G'[V \setminus \{x, y\}] \subset G$ has no embedding of $K_r^{(2)}$, so at least one of the vertices in $K_r^{(2)}$ must be mapped to $z_1$ or $z_2$. Without loss of generality, we can write $T' = \{x_1, x_2, x_3, \ldots, x_{r-1}, z_1\}$, with $x_i \notin \{z_2, z\}$ for all $i$. However, $\{z_1, x_i\}$ is an edge in $G'$ if an only if $\{z, x_i\}$ is an edge in $G$, which means that $G'[\{x_1, x_2, x_3 \ldots, x_{r-1}, z\}] = G[\{x_1, x_2, x_3 \ldots, x_{r-1}, z\}]$ is complete, contradicting our assumption.

Now, we show that $G'$ has more edges than $G$. By Lemma 3.3, $d_G(x) = d_G(z)$ and $d_G(y) = d_G(z)$, so the three vertices $x, y, z$ have the same degree $d$ in $G$. The edges containing $x$ and the edges containing $y$ intersect at exactly the edge $\{x, y\} \in E$. Therefore, by removing all of them from $G$ we are removing $2d - 1$ edges. Furthermore, for each edge containing $z$ we are adding two edges, and these sets of edges do not intersect because $z$ is not adjacent to $x$ or $y$ (so $\{z_1, z_2\} \notin E'$). We conclude that $G'$ has $|E'| = |E| - (2d - 1) + 2d = |E| + 1 > |E|$ edges, as desired. $\square$

Now we are ready to prove Turán's theorem.

*Proof of Theorem 3.2.* We have shown in Lemma 3.4 that $G = (V_1, \dots V_p; E)$ is complete. In fact, we can set $p = r - 1$: If $p < r - 1$, we can always add empty parts to $G$; and if it has more than $r - 1$ nonempty parts (without loss of generality, $x_1 \in V_1, \dots, x_r \in V_r$), then $G[\{x_1, \dots, x_r\}]$ is complete, which is a contradiction. Furthermore, any complete $(r-1)$-partite 2-graph is $K_r^{(2)}$-free, because $K_r^{(2)}$ is not $(r-1)$-partite.

This means that we only need to show that the choice of the part sizes $a_1, \dots, a_{r-1}$ summing to $n$ in the statement maximizes the expression (5). The condition that $\lfloor n/(r-1) \rfloor \leq a_i \leq \lceil n/(r-1) \rceil$ for all $i$ is equivalent to requiring that the part sizes are as equal as possible, that is, $|a_i - a_j| \leq 1$ for all $i, j$. Suppose, by way of contradiction and without loss of generality, that $a_1 > a_2 + 1$. Let $a_1' = a_1 - 1$, $a_2' = a_2 + 1$ and $a_i' = a_i$ for all $i \geq 3$. Then,

$$\begin{aligned} \sum_{\{x,y\} \in \binom{[r-1]}{2}} a_x' \cdot a_y' &= (a_1 - 1)(a_2 + 1) + (a_1 - 1) \sum_{i \geq 3} a_i + (a_2 + 1) \sum_{i \geq 3} a_i + \sum_{3 \leq x < y} a_x a_y \\ &= \sum_{\{x,y\} \in \binom{[r-1]}{2}} a_x \cdot a_y - a_2 + a_1 - 1 > \sum_{\{x,y\} \in \binom{[r-1]}{2}} a_x \cdot a_y, \end{aligned}$$

in contradiction to the number of edges in $G$ being maximal. □

Because of the difficulty of finding exact Turán numbers for $k$-graphs, we usually look for asymptotic approximations of them. In particular, we are interested in how the expression $\mathrm{ex}(n, G)$ grows with $n$ for any fixed $k$-graph $G$. This is known as the *Turán problem* for the $k$-graph $G$. For an example, we turn to the complete 2-graph $K_r^{(2)}$, for which we already have an exact formula. In expression (5), we can see that $a_i = n/(r-1) + \mathcal{O}(1)$ for all $i$. Therefore,

$$\mathrm{ex}\left(n, K_r^{(2)}\right) = \sum_{\{x,y\} \in \binom{[r-1]}{2}} a_x \cdot a_y = \binom{r-1}{2} \cdot \left(\frac{n}{r-1} + \mathcal{O}(1)\right)^2 = \frac{(r-2)}{2(r-1)} n^2 + \mathcal{O}(n). \tag{6}$$

Note that the maximum number of edges in a 2-graph on $n$ vertices is

$$\binom{n}{2} = \frac{1}{2} n^2 + \mathcal{O}(n).$$

The two quantities are comparable as they are both quadratic in $n$. This allows us to restate equation (6) as

$$\mathrm{ex}\left(n, K_r^{(2)}\right) = \frac{r-2}{r-1} \binom{n}{2} + \mathcal{O}(n) = \left(1 - \frac{1}{r-1} + o(1)\right) \binom{n}{2}, \tag{7}$$

which means that, asymptotically, the maximum *edge density* of a 2-graph on $n$ vertices without $K_r^{(2)}$ as a subgraph is $(r-2)/(r-1) < 1$, so we must exclude a nontrivial fraction of edges to avoid any particular complete 2-graph. The following general theorem greatly restricts the growth of Turán numbers for all $k$-graphs.

**Theorem 3.5.** *Let $G = (V, E)$ be a $k$-graph. The limit*

$$\pi(G) = \lim_{n \to \infty} \frac{ex(n, G)}{\binom{n}{k}} \tag{8}$$

*exists and is between 0 and 1. It is called it the* Turán density *of $G$.*

*Proof.* The sequence
$$a_n = \frac{\mathrm{ex}\,(n, G)}{\binom{n}{k}}$$
is bounded between 0 and 1 for all $n \geq |V(G)|$, by Proposition 3.1. Furthermore, it is less than 1 for all $n \geq |V(G)| + 1$. To see this, consider a $k$-graph $H = (W, F)$ with $n$ vertices and $\binom{n}{k} - 1$ edges. Its edge density is less than 1. Without loss of generality, we can suppose that $F = \binom{W}{k} \setminus \{\{x, y\}\}$. This means that $H[W \setminus \{x\}]$ is a complete $k$-graph on $n - 1$ vertices, which must contain $G$ as a subgraph.

We show that the sequence $(a_n)$ is non-increasing, so it must converge to a value $0 \leq \pi(G) < 1$. Let $n \geq |V(G)|$. There exists a $k$-graph $H = (W, F)$ with $n + 1$ vertices and $\mathrm{ex}\,(n + 1, G)$ edges that does not contain $G$ as a subgraph. For each vertex $w \in W$, the $k$-graph $H_w = H[W \setminus \{w\}]$ has $n$ vertices and does not contain $G$ as a subgraph. Therefore, it must contain at most $\mathrm{ex}\,(n, G)$ edges. Consider the set
$$\mathcal{P} = \{(w, e) \in W \times F \mid e \in E(H_w)\}\,.$$
Counting on the first coordinate, we get
$$|\mathcal{P}| = \sum_{w \in W} |E(H_w)| \leq (n + 1)\,\mathrm{ex}\,(n, G)\,. \tag{9}$$
On the other hand, for every edge $e \in F$, $e \in E(H_w)$ for all $w \in W \setminus e$. Therefore, counting on the second coordinate, we get
$$|\mathcal{P}| = (n + 1 - k)|F| = (n + 1 - k)\,\mathrm{ex}\,(n + 1, G)\,. \tag{10}$$
Combining equations (9) and (10), we get
$$(n + 1 - k)\,\mathrm{ex}\,(n + 1, G) \leq (n + 1)\,\mathrm{ex}\,(n, G)\,.$$
Going back to the sequence $a_n$, we can write
$$a_{n+1} = \frac{\mathrm{ex}\,(n + 1, G)}{\binom{n+1}{k}} \leq \frac{(n + 1)\,\mathrm{ex}\,(n, G)}{(n + 1 - k)\binom{n+1}{k}} = \frac{\mathrm{ex}\,(n, G)}{\binom{n}{k}} = a_n. \qquad \square$$

We can now summarize expression (7) as follows.

**Corollary 3.6.** *The Turán density of the complete* 2*-graph* $K_r^{(2)}$ *is*
$$\pi\left(K_r^{(2)}\right) = \frac{r - 2}{r - 1} = 1 - \frac{1}{r - 1}.$$

The first natural question that arises is for what $k$-graphs $G$ the Turán density $\pi(G)$ is positive (in which case, we call the corresponding Turán problem *non-degenerate* and consider it solved if we can calculate $\pi(G)$). The following gives a complete characterization.

**Proposition 3.7.** *Let $G = (V, E)$ be a $k$-graph. Then $\pi(G) = 0$ if and only if $G$ is $k$-partite.*

*Proof.* If $G$ is not $k$-partite, a construction similar to the one in the proof of Theorem 3.2 directly shows $\pi(G) > 0$. Indeed, for all $m$ the $k$-graph $K\left(m, \overset{k}{\ldots}, m\right)$ is $k$-partite so it cannot contain $G$ as a subgraph. Furthermore, its edge density is
$$\frac{m^k}{\binom{km}{k}} \geq \frac{1}{k^k}.$$

Because we can make $n = km = |V(K(m, \overset{k}{\ldots}, m))|$ as large as we want, the limit (8) bounded below by a positive constant. We defer the proof of the other direction to Section 3.2, where we study $k$-partite $k$-graph Turán problems in more depth (in particular, see Theorem 3.14). $\square$

In fact, non-degenerate Turán problems for 2-graphs are considered solved in this regard. The following theorem gives the Turán density of all 2-graphs.

**Theorem 3.8** (Erdős–Stone–Simonovits Theorem). *Let $G = (V, E)$ be a 2-graph and let $r = \chi(G)$. Then,*

$$\pi(G) = 1 - \frac{1}{r-1}.$$

We defer the proof of this theorem to the next subsection, where we will have more powerful tools at our disposal. Note that letting $G = K_r^{(2)}$ we recover Corollary 3.6 of Theorem 3.2.

We know that $k$-graphs asymptotically below the Turán number of a $k$-graph $G$ may not contain $G$ as a subgraph. We may also ask how many copies (different embeddings) of $G$ can be found in a $k$-graph $H$ exceeding the Turán density. The following surprising result [10] shows that the number of copies of $G$ in $H$ is asymptotically guaranteed to be very large. In essence, if a $k$-graph $H_n$ has $n$ vertices and $(\pi(G) + \Omega(1))\binom{n}{k}$ edges, then it must contain $\Omega\left(n^{|V|}\right)$ copies of $G$ as a subgraph, which correspond to a positive fraction of all the functions from $V(G)$ to $V(H)$.

**Theorem 3.9** (Supersaturation). *Let $G = (V, E)$ be a $k$-graph and let $\epsilon > 0$. There exists a positive integer $t = t(G, \epsilon)$ and a constant $\delta = \delta(G, \epsilon) > 0$ such that any $k$-graph $H$ with $n \geq t$ vertices and at least $(\pi(G) + \epsilon)\binom{n}{k}$ edges contains at least $\delta n^{|V|}$ copies of $G$ as a subgraph.*

*Proof.* Pick $t(G, \epsilon)$ large enough so that $\operatorname{ex}(t, G) \leq \left(\pi(G) + \frac{\epsilon}{2}\right)\binom{t}{k}$. Let $m \geq (\pi(G) + \epsilon)\binom{n}{k}$ be the number of edges of $H$. Notice that

$$\binom{n-k}{t-k} m = \sum_{T \in \binom{V}{t}} |E(H[T])|.$$

This is because, for each edge in $H$, we can choose a set $T \subset V$ containing it in $\binom{n-k}{t-k}$ ways. We define

$$P = \left\{ T \in \binom{V}{t} \middle| E(H[T]) > \left(\pi(G) + \frac{\epsilon}{2}\right)\binom{t}{k} \right\}.$$

If $T \in \binom{V}{t}$, the number of edges in $H[T]$ is at most $\binom{t}{k}$. Therefore,

$$\binom{n-k}{t-k}(\pi(G)+\epsilon)\binom{n}{k} \leq \binom{n-k}{t-k} m \leq |P|\binom{t}{k} + \left(\binom{n}{t} - |P|\right)\left(\pi(G) + \frac{\epsilon}{2}\right)\binom{t}{k}.$$

Rearranging and applying standard binomial coefficient identities, we can bound $|P|$ as

$$|P| \geq \frac{\epsilon}{2(1 - \pi(G) - \epsilon/2)}\binom{n}{t} \geq \frac{\epsilon}{2}\binom{n}{t}.$$

Now, for each $T \in P$, $H[T]$ contains $G$ as a subgraph. Furthermore, each copy of $G$ is in at most $\binom{n-|V|}{t-|V|}$ such sets. Therefore, the number of copies of $G$ in $H$ is at least

$$\frac{\epsilon}{2}\binom{n}{t}\frac{1}{\binom{n-|V|}{t-|V|}} = \frac{\epsilon}{2\binom{t}{|V|}}\binom{n}{|V|} \geq \frac{\epsilon}{2\binom{t}{|V|}|V|^{|V|}} n^{|V|}.$$

Picking $\delta = \frac{\epsilon}{2\binom{t}{|V|}|V|^{|V|}}$ gives us the desired result. $\square$

## 3.2 Degenerate Turán-Type Problems

We now turn our attention to Turán problems for $k$-partite $k$-graphs, which are the ones that have Turán density 0 (we will prove so in this section). All $k$-partite $k$-graphs with part sizes $b_1 \leq a_1, \dots, b_k \leq a_k$ are contained in $K(a_1, \dots, a_k)$ as subgraphs. This allows us to follow the same argument as in Proposition 3.1 to define the following.

**Definition 3.10.** Let $1 < t_1 \leq v_1, \dots, 1 < t_k \leq v_k$ be integers. Then the *generalized Zarankiewicz number* $z(v_1, \dots, v_k; t_1, \dots, t_k)$ is the largest integer $0 \leq z < \prod_i v_i$ for which there exists a $k$-partite $k$-graph $H$ with part sizes $|V_1| = v_1, \dots, |V_k| = v_k$ and $z$ edges such that no embedding $f$ of $K(T_1, \dots, T_k)$ with $|T_i| = t_i$ into $H$ exists satisfying $f(T_i) \subset V_i$ for all $i$.

From now on, every time we talk about embeddings from one $k$-partite $k$-graph $G = (T_1, \dots, T_k; E)$ to another $k$-partite $k$-graph $H = (V_1, \dots, V_k; F)$, we assume the condition $f(T_i) \subset V_i$. Similarly to the case of complete $k$-graphs, $H$ contains $K(t_1, \dots, t_k)$ as a subgraph if and only if for some sets $S_i \subset V_i$ of size $t_i$ for all $i$, $H[S_1 \cup \cdots \cup S_k] = K(S_1, \dots, S_K)$, and such an embedding is always induced. Definition 3.10 is useful for studying the Turán problem for $k$-partite $k$-graphs in the following way.

*Remark* 3.11. Finding Zarankiewicz numbers can help us upper bound the Turán number of $K(t_1, \dots, t_k)$. Assume that $H$ is a $K(t_1, \dots, t_k)$-free $n$-vertex $k$-graph with $m$ edges. pick $v_1, \dots, v_k$ such that $\sum_i v_i = n$ and $v_i \sim n/k$ (for example, $\lfloor n/k \rfloor \leq v_i \leq \lceil n/k \rceil$). Let $V_1, \dots, V_k$ be a uniform random partition of $V(H)$ with $|V_i| = v_i$. Assuming, for example, that $n \geq 2k$, for any edge $e \in E(H)$, the probability that $e$ is an edge in $K(V_1, \dots, V_k)$ is at least

$$k! \prod_i \frac{v_i}{n} \geq \frac{k!}{(2k)^k}$$

which is independent of $n$. Therefore, the expected number of edges satisfying this condition is a positive fraction of $m$. Applying the first moment method, there is at least one partition retaining $\frac{k!}{(2k)^k} m$ edges. This means that, if the size of each part is greater than $t_i$ (that is, $n \geq kt_i$ for all $i$) and $m$ is greater than $z(v_1, \dots, v_k; t_1, \dots, t_k) \cdot \frac{(2k)^k}{k!}$, then $H$ must contain $K(t_1, \dots, t_k)$ as a subgraph. All in all,

$$\mathrm{ex}(n, K(t_1, \dots, t_k)) \leq \frac{(2k)^k}{k!} \cdot z\left(\lceil n/k \rceil, \overset{k}{\dots}, \lceil n/k \rceil; t_1, \dots, t_k\right).$$

The problem of finding the Zarankiewicz number was first posed by K. Zarankiewicz in 1951 for the case of bipartite 2-graphs (that is, finding $z(u, w; s, t)$), in terms of finding all-1 sub-matrices in a 0-1 matrix. An upper bound for it in the case $u = w, s = t$ was found by Kővari, Sós and Turán [17] in 1954. This was generalized to arbitrary complete bipartite 2-graphs by Hyltén–Cavallius [14] in 1958. The result is stated and proved here for completeness.

**Theorem 3.12** (Kővari–Sós–Turán Theorem)**.** *Let $0 < s \leq u$ and $0 < t \leq w$ be integers. Then*

$$z(u, w; s, t) \leq (s-1)^{1/t}(w - t + 1)u^{1-1/t} + (t-1)u$$

*Proof.* Suppose, by way of contradiction, that we have a $K(s, t)$-free bipartite graph $H = (U, W; E)$ with $|U| = u$, $|W| = w$ and $|E| = z$ exceeding the bound stated above. Let us consider the set

$$P = \left\{ (x, T) \in U \times \binom{W}{t} \,\middle|\, \{x, y\} \in E \text{ for all } y \in T \right\}.$$

Counting on the first coordinate, we get

$$|P| = \sum_{x \in U} \binom{d_H(x)}{t} = \sum_{x \in U} \varphi(d_H(x)) \geq u\varphi(z/u) = u\binom{z/u}{t}, \tag{11}$$

where we define

$$\varphi(x) = \begin{cases} \binom{x}{t}, & \text{if } x \geq t-1; \\ 0, & \text{otherwise.} \end{cases}$$

The function $\varphi$ is convex, so we get the inequality in (11) as a consequence of Jensen's inequality. The other equalities come from the fact that $\varphi(d)$ agrees with $\binom{d}{t}$ for all integers $d \geq 0$; and that by our bound on $z$, $z \geq (t-1)u \implies z/u \geq t-1$.

If we had $s$ different elements of $P$ with the same second coordinate $T$, they would all necessarily have different first coordinates (say $S = \{x_1, \dots, x_s\}$). But now, by definition of $P$, for all $a \in S, b \in T$, we have $\{a, b\} \in E$, so $H[S \cup T] = K(S, T)$, contradicting the assumption that $H$ is $K(s, t)$-free. Therefore, there are at most $s-1$ different elements of $P$ for each $T \in \binom{W}{t}$:

$$|P| \leq (s-1)\binom{w}{t}. \tag{12}$$

Putting inequalities (11) and (12) together, we get

$$u\binom{z/u}{t} \leq (s-1)\binom{w}{t}. \tag{13}$$

Now, because we can see $E$ as a subset of $U \times W$, we get $z \leq uw \implies z/u \leq w$. We claim that this implies that

$$\frac{(z/u - (t-1))^t}{\binom{z/u}{t}} \leq \frac{(w-(t-1))^t}{\binom{w}{t}}, \tag{14}$$

because the function

$$g(x) = \frac{(x-(t-1))^t}{\binom{x}{t}}$$

is increasing for $x \geq t-1$. To see this, we expand the denominator into a product and absorb the $(x-(t-1))^t$ factor.

$$g(x) = \prod_{i=0}^{t-1} (x-(t-1))\frac{i+1}{x-i} = t! \prod_{i=0}^{t-1} \frac{x-(t-1)}{x-i}. \tag{15}$$

Every factor in the product on the right side of (15) is increasing in $x$ for $x \geq t-1 \geq i$, proving the claim. Multiplying inequalities (13) and (14) yields

$$u\,(z/u - (t-1))^t \leq (s-1)(w-(t-1))^t.$$

Then, algebraic manipulation then gives

$$z \leq (s-1)^{1/t}(w-t+1)u^{1-1/t} + (t-1)u,$$

in contradiction to our assumption. $\square$

*Remark* 3.13. Following Remark 3.11, we can use this bound to get an upper bound on the Turán number of $K(s,t)$:

$$\operatorname{ex}(n, K(s,t)) = \mathcal{O}\left((s-1)^{1/t}\left(\left\lceil\frac{n}{2}\right\rceil - t + 1\right)n^{1-1/t} + (t-1)\left\lceil\frac{n}{2}\right\rceil\right) = \mathcal{O}\left(n^{2-1/t}\right).$$

Note that if $s < t$, we get the better bound $\mathcal{O}\left(n^{2-1/s}\right)$ by interchanging the roles of $s$ and $t$.

In 1964, Erdős [8] generalized this result to arbitrary complete partite $k$-graphs in the following theorem.

**Theorem 3.14** (Erdős 1964 Theorem)**.** *For $k \geq 2$, $ex\left(n, K\left(t, \overset{k}{\ldots}, t\right)\right) = \mathcal{O}\left(n^{k-\frac{1}{t^{(k-1)}}}\right)$.*

This theorem is a consequence of the following lemma.

**Lemma 3.15.** *Let $k \geq 2$ be an integer. There exists a constant $\delta = \delta_k > 0$ such that, for all integers $t \leq w$,*

$$z\left(w, \overset{k}{\ldots}, w; t, \overset{k}{\ldots}, t\right) < \delta w^{k-\frac{1}{t^{(k-1)}}}.$$

*Proof.* We proceed by induction on $k$. For $k = 2$, this is obtained by setting $u = w$ and $s = t$ in Theorem 3.12. This yields

$$z(w,w,t,t) \leq (t-1)^{1/t}(w-t+1)w^{1-1/t} + (t-1)w < 2w^{2-1/t} + tw,$$

where the right inequality comes from $(t-1)^{1/t} < 2$ for all positive $t$. We now argue that $tw < 2w^{2-1/t}$, which will conclude the proof for $k = 2$ by setting $\delta_2 = 4$. Otherwise,

$$t \geq 2w^{1-1/t} \geq 2t^{1-1/t},$$

which is false for all positive $t$.

For $k > 2$, suppose that the lemma holds for $k-1$ and that a certain $\delta > 0$ does not meet our conditions. There exist integers $t \leq w$ and a $k$-partite $k$-graph $H = (W_1, \ldots, W_k; F)$ with part sizes $|W_i| = w$ and $z = |F| \geq \delta w^{k-\frac{1}{t^{(k-1)}}}$ edges such that no embedding of $K\left(t, \overset{k}{\ldots}, t\right)$ into $H$ exists. Consider, for each set $T \in \binom{W_k}{t}$, the associated $(k-1)$-link $L_H(T; k-1)$. We claim that it does not contain $K\left(t, \overset{k-1}{\ldots}, t\right)$ as a subgraph. If it did (say, $T_1 \times \cdots \times T_{k-1} \subset E(L_H(T; k-1))$), then $T_1 \times \cdots \times T_{k-1} \times T \subset F$ would contradict the assumption that $H$ does not contain $K\left(t, \overset{k}{\ldots}, t\right)$ as a subgraph. This means that

$$L_H(T; k-1) \text{ has at most } z' \text{ edges for all } T \in \binom{W_k}{t}, \tag{16}$$

where

$$z' = z\left(w, \overset{k-1}{\ldots}, w; t, \overset{k-1}{\ldots}, t\right) \leq \delta_{k-1} w^{k-1-\frac{1}{t^{(k-2)}}}.$$

Now, consider the bipartite graph $H' = (U, W; F')$, where

$$\begin{aligned} U &= W_1 \times \cdots \times W_{k-1}, \\ W &= W_k, \\ F' &= \{(X, y) \in U \times W \mid X \cup \{y\} \in F\}. \end{aligned}$$

Condition (16) is equivalent to saying that there is no embedding of $K(z'+1,t)$ onto $H'$ respecting the partitions. Furthermore, $H'$ has the same number of edges as $H$. Finally, we invoke Theorem 3.12 with $u=|U|=w^{k-1}$ and $s=z'+1$ to get

$$\delta w^{k-\frac{1}{t^{(k-1)}}} \leq |E| = |E'| \leq \left(\delta_{k-1} w^{(k-1)-\frac{1}{t^{(k-2)}}}\right)^{1/t} (w-t+1) w^{(k-1)(1-1/t)} + (t-1) w^{k-1}. \tag{17}$$

Approximating, this implies that

$$\delta w^{k-\frac{1}{t^{(k-1)}}} < \delta_{k-1}^{1/t} w^{k-\frac{1}{t^{(k-1)}}} + t w^{k-1} \leq \delta_{k-1} w^{k-\frac{1}{t^{(k-1)}}} + t w^{k-1}.$$

Similarly as before, one can check that $tw^{k-1} < 2w^{k-\frac{1}{t^{(k-1)}}}$. Therefore, we reach a contradiction whenever $\delta \geq \delta_{k-1}+2$, so setting $\delta_k = \delta_{k-1}+2$ gives us the desired result. In fact, the theorem works for $\delta_k = 2 \cdot k$. □

The proof of the Erdős 1964 Theorem is now straightforward.

*Proof of Theorem 3.14.* If $t=1$, there is nothing to prove, so suppose that $t \geq 2$. Following Remark 3.11, if $n \geq tk$ (and in particular $n \geq 2k$),

$$\mathrm{ex}\left(n, K\left(t, \overset{k}{\ldots}, t\right)\right) \leq \frac{(2k)^k}{k!} \cdot z\left(\lceil n/k \rceil, \overset{k}{\ldots}, \lceil n/k \rceil; t, \overset{k}{\ldots}, t\right) \leq 2k \cdot \frac{(2k)^k}{k!} \cdot \left\lceil \frac{n}{k} \right\rceil^{k-\frac{1}{t^{(k-1)}}} \leq \frac{k \cdot 4^k}{(k-1)!} \cdot n^{k-\frac{1}{t^{(k-1)}}}.$$

□

Because all $k$-partite $k$-graphs can be embedded in a $K\left(t, \overset{k}{\ldots}, t\right)$, Theorem 3.14 shows that the Turán density of all $k$-partite $k$-graphs is 0, completing the proof of Proposition 3.7. Note that the constant factor found in the bound of Theorem 3.14 does not depend on $t$. This lets us restate the theorem in the following stronger form.

**Theorem 3.16.** *Let $k \geq 2$ be an integer. There exist an integer $n_k$ and a positive constant $\gamma_k$ such that, for all $0 < \epsilon < 1$, all $k$-graphs with more than $\epsilon\binom{n}{k}$ edges contain $K\left(t_n, \overset{k}{\ldots}, t_n\right)$ as a subgraph, where*

$$t_n = \left\lfloor \left(\frac{\log n}{\log(\gamma_k/\epsilon)}\right)^{\frac{1}{k-1}} \right\rfloor.$$

*Proof.* Again, if $t_n = 1$, there is nothing to prove. Suppose that $t_n \geq 2$. Let us define

$$c_k = \frac{k \cdot 4^k}{(k-1)!}.$$

In the proof of Theorem 3.14, we have shown that

$$\mathrm{ex}\left(n, K\left(t_n, \overset{k}{\ldots}, t_n\right)\right) \leq c_k n^{k-\frac{1}{t_n^{(k-1)}}}$$

as long as $n \geq t_n k$. Suppose that $H$ is a $k$-graph with $n$ vertices and $m \geq \epsilon\binom{n}{k}$ edges. Suppose, furthermore, that $n \geq t_n k$ (we will later show that this condition can be made true independently of $n$ for our chosen $t_n$). This condition also implies that the number of edges of $H$ is at least

$$\epsilon\binom{n}{k} \geq \epsilon \frac{(n-k+1)^k}{k!} \geq \epsilon\left(\frac{1}{2}\right)^k \frac{1}{k!} n^k = (e_k \cdot \epsilon) n^k,$$

where $e_k = \frac{1}{k! \cdot 2^k}$ does not depend on $n$. We pick, for example,

$$\gamma_k = \frac{2c_k}{e_k} = 2 \cdot 8^k \cdot k,$$

so that

$$\text{ex}\left(n, K\left(t_n, \overset{k}{\ldots}, t_n\right)\right) \leq c_k n^{k - \frac{1}{t_n^{(k-1)}}} \leq c_k n^k \frac{\epsilon}{\gamma_k} = \frac{e_k \cdot \epsilon}{2} n^k \leq \frac{|E(H)|}{2} < |E(H)|,$$

which guarantees that $H$ contains $K\left(t_n, \overset{k}{\ldots}, t_n\right)$ as a subgraph. The only thing left to prove is that, for this choice of $t_n$, there exists $n_k$ such that $n \geq n_k$ implies $n \geq t_n k$. Indeed, we can pick any

$$n_k \geq \left(k\left(\frac{1}{\log(\gamma_k)}\right)^{\frac{1}{k-1}}\right)^2.$$

Using that $\gamma_k/\epsilon > \gamma_k$ (because $\epsilon < 1$), that $\gamma_k > 1$ (which can be easily checked from the definitions) and the inequality $\log n < \sqrt{n}$ yields

$$t_n k = \left\lfloor \left(\frac{\log n}{\log(\gamma_k/\epsilon)}\right)^{\frac{1}{k-1}} \right\rfloor k < \left\lfloor \left(\frac{\log n}{\log \gamma_k}\right)^{\frac{1}{k-1}} \right\rfloor k < \left(\frac{1}{\log \gamma_k}\right)^{\frac{1}{k-1}} k\sqrt{n} \leq \left(\frac{1}{\log \gamma_k}\right)^{\frac{1}{k-1}} k \frac{n}{\sqrt{n_k}} \leq n. \qquad \square$$

This is the result we prove constructively in Section 4. It is stronger than Theorem 3.14 because it bounds the Turán number of partite $k$-graphs uniformly, while obtaining the same order of magnitude. Indeed, suppose we have a fixed value for $t$. We may choose $\epsilon$ such that $t_n \geq t$. We only need that

$$\epsilon \geq \gamma_k \cdot n^{-\frac{1}{t^{(k-1)}}},$$

that is, the $k$-graph has at least

$$\gamma_k \cdot n^{-\frac{1}{t^{(k-1)}}} \cdot \binom{n}{k} = \mathcal{O}\left(n^{k - \frac{1}{t^{(k-1)}}}\right)$$

edges. Qualitatively, Theorem 3.16 states that we may find a blow-up of an edge as large as we wish, if we let the number of vertices grow while keeping the density constant. The following theorem generalizes this notion to blow-ups of arbitrary $k$-graphs.

**Theorem 3.17.** *Let $G = (V, E)$ be a $k$-graph and let $\epsilon > 0$. There exists a positive integer $n_0 = n_0(G, \epsilon)$ and a constant $\delta = \delta(G, \epsilon) > 0$ such that for all $n \geq n_0$,*

$$ex(n, G(t_n)) \leq (\pi(G) + \epsilon)\binom{n}{k},$$

*where*

$$t_n = \delta \cdot (\log n)^{\frac{1}{|V|-1}}.$$

*Proof.* We determine the value of $\delta$ later in the proof. Suppose that $H$ is a $k$-graph with $n$ vertices and at least $(\pi(G) + \epsilon)\binom{n}{k}$ edges. Theorem 3.9 states that if $n \geq n_0 \geq t(G, \epsilon)$, there are at least $\delta_1(G, \epsilon)n^{|V|}$ embeddings of $G$ into $H$.

Consider, as in Remark 3.11, a random partition of the vertices of $H$ into $|V|$ parts of size $\lfloor n/|V| \rfloor \leq |V_i| \leq \lceil n/|V| \rceil$. Suppose that we have an embedding $f$ of $G$ in $H$. Assuming that $n \geq n_0 \geq 2|V|$, the probability that for all vertices $v_i \in V$, $f(v_i) \in V_i$ is

$$\prod_{i=1}^{|V|} \frac{|V_i|}{n} \geq \left(\frac{1}{2|V|}\right)^{|V|}.$$

Therefore, for at least one such partition, there are at least

$$\delta_1(G,\epsilon) \cdot \left(\frac{1}{2|V|}\right)^{|V|} \cdot n^{|V|} = \delta_2(G,\epsilon) n^{|V|}$$

embeddings of $G$ in $H$ respecting the partition. Furthermore, these embeddings all have different image sets, which have one vertex in each part. Consider now the $|V|$-partite $|V|$-graph $H' = (V_1, \dots, V_{|V|}; F)$, where

$$F = \{f(V) \mid f \text{ is an embedding of } G \text{ in } H \text{ and } f(V_i) \subset V_i \text{ for all } i\}.$$

By Theorem 3.16, making $n_0$ large enough depending only on $\delta_2$ and $|V|$, there exists some $\delta_3 = \delta_3(G,\epsilon) = \delta(|V|, \delta_2(G,\epsilon)) > 0$ such that $H'$ contains $K\left(t_n, \overset{|V|}{\dots}, t_n\right)$ as a subgraph, where

$$t_n = \delta_3(G,\epsilon) \cdot (\log n)^{\frac{1}{|V|-1}}.$$

That is, there are $|V|$ sets $T_1, \dots, T_{|V|}$ of size $t_n$ such that if $\{v_{i_1}, \dots, v_{i_k}\} \in E$ is an edge of $G$, then $\{u_{i_1}, \dots, u_{i_k}\}$ is an edge of $H_n$ for all $u_{i_j} \in T_j$. This means that $G(t_n)$ is a subgraph of $H_n$, so picking $\delta = \delta_3$ gives the desired result.

$\square$

The following corollary highlights that degenerate Turán problems can be applied to solve non-degenerate ones.

**Corollary 3.18.** *Let $G$ be a $k$-graph and $t$ be a positive integer. Then, $\pi(G(t)) = \pi(G)$.*

In particular, this directly proves the Erdős–Stone–Simonovits theorem for 2-graphs.

*Proof of Theorem 3.8.* Let $G$ be a 2-graph with chromatic number $r$. For some $t \geq 1$, $G$ is a subgraph of $K^{(2)}(t, \overset{r}{\dots}, t) \cong K_r^{(2)}(t)$. Therefore,

$$\pi(G) \leq \pi\left(K_r^{(2)}(t)\right) = \pi\left(K_r^{(2)}\right) = 1 - \frac{1}{r-1}.$$

The reverse inequality follows from the same construction in the proof of Turán's theorem. Indeed, this construction has the desired density and avoids not only $K_r^{(2)}$ but also any $r$-partite 2-graph, and in particular $G$. $\square$

Knowing that the Turán density of $k$-partite $k$-graphs is 0 gives little information on the growth of their Turán numbers. In this case, we are usually satisfied with determining the growth up to a constant factor. So far, we have only proven upper bounds for this growth rate. General lower bounds are usually obtained by probabilistic arguments, which are often weak and worse that lower bounds obtained by other means (See Section 3.3 for a few examples of this in the literature). Here we present an example of a general probabilistic argument that applies to any guest $k$-graph.

**Proposition 3.19.** *Let $G = (V, E)$ be a $k$-graph with $\alpha = |V|$ vertices and $\beta = |E| > 1$ edges. Then, $ex(n, G) = \Omega\left(n^{k-\frac{\alpha-k}{\beta-1}}\right)$, and the constant factor depends only on the number of edges $\beta$ (and not on the number of vertices $\alpha$).*

*Proof.* Let $n \geq \alpha$. We use the so-called *random alteration* method to construct a $k$-graph $H_n$ with $n$ vertices that does not contain $G$ as a subgraph. We first define $R_n = (W, F)$ to be a random $k$-graph on a vertex set $W$ of size $n$, where each edge $e \in \binom{W}{k}$ is included in $E$ independently at random with a certain probability $p \in (0, 1)$. The expected number of edges in $R_n$ is

$$\mathbb{E}(|F|) = p\binom{n}{k}.$$

Let us now count the number of possible injective functions of $V(G)$ in $V$. They are defined by the (ordered) choice of the image of each vertex, so there are

$$\prod_{j=1}^{\alpha}(n - j + 1) \leq n^{\alpha}$$

of them. The probability that any particular injective function $f$ of $V(G)$ in $W$ is an embedding of $G$ in $R_n$ is calculated as the product of the probabilities that each image of an edge is an edge in $R_n$, because the presence of edges in $R_n$ is independent. Therefore,

$$\mathbb{P}(f \text{ is an embedding of } G) = p^{\beta}.$$

If we define $X$ to be the number of embeddings of $G$ in $R_n$, by linearity of expectation we get

$$\mathbb{E}(X) = \sum_{f}\mathbb{P}(f \text{ is an embedding of } G) \leq n^{\alpha}p^{\beta}.$$

We can now obtain a $G$-free $k$-graph $H_n$ by removing from $R_n$, for each embedding of $G$, the image of an edge of $G$ (these might coincide for different embeddings, but this only decreases the number of edges removed). The expected number of edges in $H_n$ is

$$\mathbb{E}(|E(H_n)|) \geq \mathbb{E}(|F|) - \mathbb{E}(X) \geq p\binom{n}{k} - n^{\alpha}p^{\beta}.$$

The right quantity is maximized by setting

$$p = \left(\frac{\binom{n}{k}}{\beta n^{\alpha}}\right)^{\frac{1}{\beta-1}}.$$

This yields

$$\mathbb{E}(|E(H_n)|) \geq m_0(n) = \left(\frac{\beta n^{\alpha} - 1}{(\beta n^{\alpha})^{\frac{\beta}{\beta-1}}}\right)\binom{n}{k}^{\frac{\beta}{\beta-1}} \geq \left(\frac{\beta - 1}{\beta^{\frac{\beta}{\beta-1}}}\right)n^{-\frac{\alpha}{\beta-1}}\binom{n}{k}^{\frac{\beta}{\beta-1}} \geq \left(\frac{\beta - 1}{(\beta k!)^{\frac{\beta}{\beta-1}}} - \frac{k^k}{n^k}\right)n^{k-\frac{\alpha-k}{\beta-1}}.$$

Therefore, the event that $|E(H_n)| \geq m_0(n)$ must have positive probability and, in particular, there exists one such $k$-graph $\widehat{H_n}$, which is $G$-free by construction. $\square$

## 3.3 Open Problems

In Sections 3.1 and 3.2, we have seen the solution for non-degenerate Turán problems for 2-graphs. Determining the Turán density of $k$-graphs for $k > 2$ is a much harder problem.

Famously, not even the Turán density of the tetrahedron 3-graph $K_4^{(3)}$ (pictured in Figure 2) or the unique 3-graph $K_{4-}^{(3)}$ obtained by removing one edge from it, are known. The best known bounds are

$$0.5555 = \frac{5}{9} \leq \pi\left(K_4^{(3)}\right) \leq 0.561666 \text{ [15, 1]}$$

and

$$0.2857 = \frac{2}{7} \leq \pi\left(K_{4-}^{(3)}\right) \leq 0.2871 \text{ [12, 1]}.$$

The lower bounds are obtained by explicit constructions of 3-graphs, and were conjectured to be optimal by Turán [15], while the upper bounds are obtained by the method of flag algebras [20], which is a powerful tool for studying Turán problems that automates the search for relevant inequalities.

One of the few examples of success in obtaining Turán densities of $k$-graphs with uniformity $k > 2$ is the case of the Fano plane $F_7^{(3)}$, a 3-graph with 7 vertices corresponding to the points of the projective plane over the field $\mathbb{F}_2$, and 7 edges corresponding to the projective lines. It is known that

$$\pi\left(F_7^{(3)}\right) = \frac{3}{4} \text{ [7]}.$$

There are even fewer solved cases for degenerate Turán problems than in the non-degenerate case. In general, there is a very large gap between the upper and lower bounds for the Turán numbers of degenerate $k$-graphs, even for complete $k$-partite $k$-graphs. For example, in the balanced case, where all $t_i$ are equal, we get

$$\operatorname{ex}\left(n, K\left(t, \overset{k}{\ldots}, t\right)\right) = \mathcal{O}\left(n^{k-\frac{1}{t^{(k-1)}}}\right) \tag{18}$$

from Theorem 3.14, but only

$$\operatorname{ex}\left(n, K\left(t, \overset{k}{\ldots}, t\right)\right) = \Omega\left(n^{k-\frac{k(t-1)}{t^k-1}}\right) \tag{19}$$

from Proposition 3.19. The exponent in (19) is always smaller than the one in (18), as long as $t \geq 2$ and $k \geq 2$.

In the case $k = 2$, it is known that, for $K(2, t)$ (for $t \geq 2$) and $K(3, 3)$, Theorem 3.12 is optimal in the sense that

$$\operatorname{ex}(n, K(2, t)) = \Theta\left(n^{\frac{3}{2}}\right) \text{ [9, 13]},$$

and also

$$\operatorname{ex}(n, K(3, 3)) = \Theta\left(n^{\frac{5}{3}}\right) \text{ [4]}.$$

The theorem is also optimal for $K(s, t)$ when $s \geq t! + 1$ [16]. Some progress has been made in the case of $K(s, t)$ when $s$ and $t$ have similar sizes, only for small values of $s \geq t \geq 4$. For example, Theorem 3.12 gives

$$\operatorname{ex}(n, K(5, 5)) = \mathcal{O}\left(n^{\frac{9}{5}}\right) = \mathcal{O}\left(n^{1.8}\right)$$

and, by Proposition 3.19, we get

$$\operatorname{ex}(n, K(5, 5)) = \Omega\left(n^{\frac{5}{3}}\right) = \Omega\left(n^{1.67}\right), \tag{20}$$

but (20) has been improved to

$$\mathrm{ex}\left(n, K(5,5)\right) = \Omega\left(n^{\frac{7}{4}}\right) = \Omega\left(n^{1.75}\right) \ [2].$$

Even less is known about degenerate problems for hypergraphs of higher uniformity. For example, not even the growth rate of the Turán number for the octahedron 3-graph ($K(2,2,2)$, pictured in Figure 3) is known. The best upper bound, again, comes from Theorem 3.14, which gives

$$\mathrm{ex}\left(n, K(2,2,2)\right) = \mathcal{O}\left(n^{\frac{11}{4}}\right) = \mathcal{O}\left(n^{2.75}\right),$$

while the best know lower bound is

$$\mathrm{ex}\left(n, K(2,2,2)\right) = \Omega\left(n^{\frac{8}{3}}\right) = \Omega\left(n^{2.67}\right) \ [6].$$

The main difficulty for degenerate problems is that sharp lower bounds for the Turán numbers often rely on specific geometric or algebraic constructions that work for very few cases, such as the ones cited for $K(2,2)$ and $K(3,3)$.

Theorem 3.17 is known not to be optimal. Erdős and Bollobás [3] in fact proved that the optimal growth rate of a guaranteed $t_n$-blow-up of a 2-graph $G$ in 2-graphs of constant density greater than $\pi(G)+\epsilon$ is

$$t_n = \delta(G,\epsilon)\cdot\log n.$$

A still open question is whether this can be extended to $k$-graphs. That is, is it true that $k$-graphs with $n$ vertices and $(\pi(G)+\epsilon)\binom{n}{k}$ edges contain $G(t_n)$ for some $t_n = \delta(G,\epsilon)(\log n)^{\frac{1}{k-1}}$? An even more general yet unresolved question is whether this is true for $k$-graphs with $n$ vertices and $\Omega(n^{|V(G)|})$ copies of $G$ [23].

# 4. Main Contribution

Let $H = (V, E)$ be a $k$-graph with $n$ vertices and $m$ edges. We describe a polynomial-time algorithm that finds a complete $k$-partite $k$-graph in $H$ with all part sizes at least

$$t = t(n, d, k) = \left\lfloor \left( \frac{\log\left(n/2^{(k-1)}\right)}{\log(3/d)} \right)^{\frac{1}{k-1}} \right\rfloor, \tag{21}$$

where

$$d = m/n^k \tag{22}$$

is the "un-normalized" density of $H$, which lies between 0 and $\frac{1}{k!}$ and is easier to work with for the arguments that follow. For the remainder of the section, we assume that $t \geq 2$ (otherwise, we may just select a set of $k$ vertices forming an edge in $H$). More precisely, we show the following.

**Theorem 4.1.** *There is an algorithm that, given a $k$-graph $H$ satisfying the conditions above, finds a $K\left(t, \overset{k}{\ldots}, t\right)$ embedded in $H$ with $t = t(n, d, k)$.*

*That is, the algorithm returns a tuple of sets $(V_1, \ldots, V_k) \subset \binom{V}{t}^k$ such that $V_1 \times \cdots \times V_k \subset E$. Furthermore, the algorithm's runtime is polynomial in $n$.*

*Remark* 4.2. The stated condition implies that the sets $V_1, \ldots, V_k$ are disjoint: If, for example, $v \in V_1 \cap V_2$ and for $3 \leq i \leq k$ $v_i \in V_i$ then $(v, v, v_3, \ldots, v_k) \in V_1 \times \cdots \times V_k$ has size $k-1$ as an unordered set so it cannot be an edge in $H$. This means that the inclusion map from $K(V_1, \ldots, V_k)$ to $V$ defines an embedding, as desired.

This gives a constructive proof of Theorem 3.16 (which implies the Erdős Theorem 3.14), by adjusting the value of $\gamma_k$ and $n_k$ appropriately. Furthermore, for a fixed value of $0 < d < \frac{1}{k!}$, the value of $t$ is in the best possible order of magnitude. Indeed, if

$$dn^k \geq \operatorname{ex}\left(n, K\left(t, \overset{k}{\ldots}, t\right)\right),$$

by Proposition 3.19, applied with $\alpha = tk$ and $\beta = t^k$, we have that

$$dn^k \geq \left( \frac{\beta - 1}{(\beta k!)^{\frac{\beta}{\beta-1}}} - \frac{k^k}{n^k} \right) n^{k - \frac{\alpha-k}{\beta-1}} \geq \left( \frac{t^k - 1}{t^k k!} - \frac{k^k}{n^k} \right) n^{k - \frac{k(t-1)}{t^k-1}}$$

One can check that for some $d < d' < \frac{1}{k!}$, for $n$ large enough depending only on $d$, and $k$, and assuming $t \geq 2$, this implies

$$k!d' \geq n^{-\frac{k(t-1)}{t^k-1}} \geq k!d' \geq n^{-\frac{2k}{t(k-1)}}.$$

Rearranging gives

$$t \leq \left( \frac{2k \log n}{\log(1/(k!d'))} \right)^{\frac{1}{k-1}} = \mathcal{O}\left(t(n, d, k)\right).$$

## 4.1 Previous Results

For $k = 2$, this problem was already solved by an algorithm of Mubayi and Turán [19], which we present here (Algorithm 1) for context and clarity. A slightly different value for $t$ is used because of different estimates in their proof of correctness. Specifically, $t$ is set to

$$t_2(n, d) = \left\lfloor \frac{\log(n/2)}{\log(2e/d)} \right\rfloor,$$

whereas we get

$$t(n, d, 2) = \left\lfloor \left( \frac{\log(n/2)}{\log(3/d)} \right) \right\rfloor.$$

The vertex set $V(H)$ is partitioned into two sets $U$ and $W$ such that there are many edges between them and the size of $W$ is logarithmic in $n$. This is achieved by selecting $W$ to be a set of vertices of highest degree (that is, no vertex in $U$ has a higher degree than any vertex in $W$). Then, by iterating over all $t$-subsets of $W$, such a set $T$ is found satisfying that the set $S$ of common neighbors of $T$ in $U$ has size at least $t$. In other words, $S \times T \subset E$ for $S, T \subset V$ of size at least $t$.

**Algorithm 1** Finding a balanced bipartite graph in a 2-graph

**Require:** A graph $H = (V, E)$ with $|V| = n$, $E = m$

1: $d \leftarrow m/n^2$
2: **assert** $d \geq 3n^{-1/2}$
3: $t \leftarrow \left\lfloor \frac{\log(n/2)}{\log(2e/d)} \right\rfloor$, $w \leftarrow \lfloor t/d \rfloor$
4: $W \leftarrow$ a set of $w$ vertices with highest degree in $H$
5: $U \leftarrow V \setminus W$
6: **for all** $T \in \binom{W}{t}$ **do**
7:     $S \leftarrow \{x \in U \colon \{x, y\} \in E \text{ for all } y \in T\}$
8:     **if** $|S| \geq t$ **then**
9:         **return** $(S, T)$
10:     **end if**
11: **end for**

The minimum density $d \geq 3n^{-1/2}$ in line 2 of Algorithm 1 is required because if $d = o\left(n^{-1/2}\right)$ then there may not even be a $K(2, 2)$ in $H$. If the set $S$ is too large, a subset of it of size $t$ can be returned instead. To see that the algorithm returns a pair of sets $(S, T)$, one uses the fact that there is large number of edges between $U$ and $W$ (proportional to the size of $W$). Then, a direct application of Theorem 3.12 with $u = |U| = n - w$ and $s = t$ shows that there is a $K(t, t)$ in the bipartite graph $(U, W; E \cap (U \times W))$. This in turn means that for some $T$, the size of $S$ is at least $t$ and the algorithm returns $(S, T)$. Finally, the algorithm runs in polynomial time because the number of iterations of the loop is

$$\binom{w}{t} \leq \left(\frac{ew}{t}\right)^t \leq \left(\frac{1}{d}\right)^t e^t < e^{t \log(1/d) + \log n} < e^{2 \log n} = n^2.$$

## 4.2 General Algorithm for Hypergraphs

We now present Algorithm 2, which is a generalization of Algorithm 1 to $k$-graphs. It follows the same structure as Algorithm 1, but it is defined recursively, resembling the induction step of Lemma 3.15. This is the algorithm mentioned in Theorem 4.1, and the main contribution of this work.

The main idea is to select a set $W \subset V$ of vertices of highest degree with

$$|W| = w = w(n, d, k) = \left\lceil \frac{2t(n, d, k)}{d} \right\rceil . \tag{23}$$

Then, for every $t$-subset $T$ of $W$, we compute the set $S$ of $(k-1)$-subsets of $V \setminus W$ that form an edge with every vertex in $T$. These are precisely the edges of $H' = L_{k-1}(H; T) = (V \setminus W, S)$. For a specific $T$, the set $S$ satisfies

$$|S| \geq s = s(n, d, k) = \left\lceil d^{t(n,d,k)} n^{(k-1)} \right\rceil . \tag{24}$$

As it turns out, $S$ is large enough (24) that applying the algorithm recursively to $H'$ yields a $K\left(t', \overset{k-1}{\dots}, t'\right)$ in $H'$ with $t' \geq t$. That is, a tuple $P' = (V_1, V_2, \dots, V_{k-1}) \in \mathcal{P}(V \setminus W)^{k-1}$ such that $|V_i| = t'$ and $V_1 \times \cdots \times V_{k-1} \subset S$.

If we now concatenate $P'$ with $T$ (choosing a subset of $X_i \subset V_i$ of size $t$ for each $i$ if necessary), we get a tuple $(X_1, \dots, X_{k-1}, T)$ of $t-$*sets* of $V$ which by the definition of $S$ satisfies $X_1 \times \cdots \times X_{k-1} \times T \subset E = E(H)$, so it forms a $K\left(t, \overset{k}{\dots}, t\right)$ in $H$.

**Algorithm 2** Finding a balanced partite $k$-graph in a $k$-graph

**function** FIND_PARTITE$(H, k)$
  **assert** $H$ is a $k$-graph
  **if** $k = 1$ **then**
    **return** $(\{x \colon \{x\} \in E(H)\})$
  **end if**
  $V \leftarrow V(H), E \leftarrow E(H), n \leftarrow |V|,\ m \leftarrow |E|,\ d \leftarrow m/n^k$
  $t \leftarrow t(n, d, k),\ w \leftarrow w(n, d, k),\ s \leftarrow s(n, d, k)$
  **assert** $t \geq 2$
  $W \leftarrow$ a set of $w$ vertices with highest degree in $H$
  $U \leftarrow \binom{V \setminus W}{k-1}$
  **for all** $T \in \binom{W}{t}$ **do**
    $S \leftarrow \{\, y \in U \colon \{x\} \cup y \in E$ for all $x \in T \,\}$
    **if** $|S| \geq s$ **then**
      $H' \leftarrow (V \setminus W, S)$
      $(V_1, \dots, V_{k-1}) \leftarrow$ FIND_PARTITE$(H', k-1)$
      **return** $(V_1, \dots, V_{k-1}, T)$
    **end if**
  **end for**
**end function**

## 4.3 Proof of Correctness

The implementation of Algorithm 2 and its proof of correctness are less cumbersome if we assume a 1-graph to be just a subset of a set and use it as the base case. We also make the simplification of not including in the algorithm pseudocode the size reduction of the sets obtained from the recursive call. As stated, Algorithm 2 in fact returns a complete $k$-partite subgraph with part sizes *at least* $t$, which can easily be post-processed if desired to get a complete *balanced* subgraph with part sizes $t$.

The aim of the rest of this section is to prove that Algorithm 2 is correct (as long as the condition $t \geq 2$ in line 8 is met on the first call) and runs in polynomial time. That is, to prove it meets the requirements of Theorem 4.1. From now on, we assume $k \geq 2$ and $t \geq 2$, unless stated otherwise. The following observation is useful for some of the bounds we have to prove.

*Remark* 4.3. The requirement $t \geq 2$ is met whenever

$$d \geq 3 \cdot 2^{\frac{k-1}{2^{(k-1)}}} n^{-\frac{1}{2^{(k-1)}}}, \tag{25}$$

However, $d$ satisfies

$$d = \frac{m}{n^k} \leq \frac{\binom{n}{k}}{n^k} < \frac{1}{k!}, \tag{26}$$

so we get the following minimum value of $n$.

$$n > \left(k! \cdot 3 \cdot 2^{\frac{k-1}{2^{(k-1)}}}\right)^{2^{(k-1)}} \geq 72. \tag{27}$$

From Inequality (25) we can also see that

$$d \geq 3\sqrt{\frac{2}{n}} \tag{28}$$

for all $k \geq 2$. For $k = 2$, this reads directly from the inequality. For $k > 2$, suppose that the bound is not met. Then,

$$3n^{-\frac{1}{4}} \leq 3n^{-\frac{1}{2^{(k-1)}}} < d < 3\sqrt{\frac{2}{n}},$$

which by algebraic manipulation implies $n < 4$, in contradiction to the minimum value for $n$ found in (27).

We start by proving that the selection of $t, w, s$ in line 7 of Algorithm 2 is sound, in the sense that we only consider subsets of sizes smaller than the corresponding supersets.

**Lemma 4.4.** *For $t, w, s$ as selected in line 7 of Algorithm 2, we have that $t \leq w \leq n$, $k-1 \leq n-w$ and $s \leq \binom{n-w}{k-1}$.*

*Proof.* It is clear from the definitions that $w \geq t$. To see that $w \leq n$, we in fact show that $w < \frac{n}{2}$. If not, then

$$\frac{n}{2} \leq w = \left\lceil \frac{2t}{d} \right\rceil \leq 1 + \frac{2t}{d} < 1 + \frac{\log(n/2)}{d}.$$

Now, using that $d \geq 3\sqrt{\frac{2}{n}}$ (28), this implies

$$\frac{n}{2} < 1 + \frac{2\log(n/2)\sqrt{n}}{3} < 1 + \frac{n}{4}.$$

Therefore, $n < 4$, in contradiction to our earlier result that $n > 72$ (27). It is also clear from Inequality (27) that $n > 2k$, so we also have $k < n/2$. Therefore, $k + w < n/2 + n/2 = n$, which implies $k - 1 < n - w$, as we wanted to show.

Finally, for sake of contradiction, suppose $s > \binom{n-w}{k-1}$. By the definition of $s$ (24) and the fact that $\binom{n-w}{k-1}$ is an integer, we have that $d^t n^{k-1} > \binom{n-w}{k-1}$. Then, using the fact that $w < \frac{n}{2}$, we get

$$\left(\frac{n}{2k}\right)^{k-1} \le \left(\frac{n-w}{k-1}\right)^{k-1} \le \binom{n-w}{k-1} < d^t n^{k-1},$$

which implies

$$\left(\frac{1}{2k}\right)^{k-1} < d^t \le \left(\frac{1}{k!}\right)^2.$$

In the last inequality, we have used that $t \ge 2$ and that $d \le \frac{1}{k!}$. Since $k!^2 \ge (2k)^{k-1}$ for all $k$, we have reached a contradiction. □

The next step is to show that there are many edges with exactly one vertex in $W$. More precisely, we have the following.

**Lemma 4.5.** *Given $W \subset V$ as defined in line 9 of Algorithm 2, There are at least $\frac{3}{2}dwn^{k-1}$ edges of $H$ with exactly one vertex in $W$.*

*Proof.* The degree sum over $V$ is $kdn^k$. By averaging, the degree sum over $W$ is at least $\frac{w}{n}kdn^k = wkdn^{k-1}$. For $2 \le j \le n$, consider the contribution to this sum by edges with exactly $j$ vertices in $W$. Each such edge contributes $j$ to the sum, and there are at most $\binom{w}{j}\binom{n-w}{k-j} \le \frac{w^j n^{k-j}}{j!} \le \frac{w^j n^{k-j}}{j}$ of them. Thus, the total contribution of these edges is at most $w^j n^{k-j} \le w^2 n^{k-2}$. The number of edges with only one vertex in $W$ is then at least

$$wkdn^{k-1} - (k-1)w^2 n^{k-2} = dwn^{k-1}\left(k - \frac{(k-1)w}{nd}\right).$$

Suppose, by way of contradiction, that $k - \frac{(k-1)w}{nd} < \frac{3}{2}$. Using that $\frac{k-1}{k-3/2} \le 2$ for $k \ge 2$, we arrive at

$$2 \ge \frac{nd}{w},$$

which implies

$$d \le \frac{2w}{n} = \frac{2\left\lceil \frac{2t}{d} \right\rceil}{n} < \frac{6t}{dn},$$

where the last inequality follows from the fact that $t > 1$ and $d \le 1$. Algebraic manipulation then yields

$$nd^2 < 6t.$$

We now closely follow the steps of Mubayi and Turán [19].

If $3\sqrt{\frac{2}{n}} \le d \le 3\sqrt{\frac{\log n}{n}}$, we get

$$18 \le nd^2 < 6t \le 6\frac{\log(n/2)}{\log(3/d)} < 6\frac{\log n}{\log\left(\sqrt{\frac{n}{\log n}}\right)} = 12\frac{\log n}{\log\left(\frac{n}{\log n}\right)} < 12\frac{\log n}{\log\left(\frac{n}{\log n}\right)} < 12\frac{\log n}{\log\left(n^{2/3}\right)} = 18,$$

which is a contradiction.

Otherwise, we have $d > 3\sqrt{\frac{\log n}{n}}$. This yields $9\log n \le nd^2 < 6t < 6\log n$ (again, a contradiction). □

We use this fact to show that for some $T \subset W$, there is a large number of $(k-1)$-subsets of $V \setminus W$ that form an edge with every vertex in $T$.

**Lemma 4.6.** *For some $T \in \binom{W}{t}$, the corresponding set $S$ defined in line 12 of Algorithm 2 has size at least $s$.*

*Proof.* We apply Theorem 3.12 to the 2-partite 2-graph

$$\mathcal{P} = (U, W; F),$$

where $F$ is defined as

$$F = \{(x, y) \in U \times W \mid \{x\} \cup y \in E\}.$$

By Lemma 4.5, $\mathcal{P}$ has at least $\frac{3}{2}dwn^{k-1}$ edges. By way of contradiction, suppose that the lemma is false. There are no sets $S \in \binom{U}{s}$, $T \in \binom{W}{t}$ such that $(x, y) \in E(\mathcal{P})$ for all $x \in S, y \in T$. In other words, there is no embedding of $K(s, t)$ in $\mathcal{P}$. By Theorem 3.12 applied with $u = \binom{n-w}{k-1}$, this implies that

$$\frac{3}{2}dwn^{k-1} \leq z\left(\binom{n-w}{k-1}; w, s, t\right) \leq (s-1)^{1/t}(w-t+1)\binom{n-w}{k-1}^{1-1/t} + (t-1)\binom{n-w}{k-1}.$$

We now substitute into the above expression $(s-1) \leq d^t n^{k-1}$ (which follows from $s = \lceil d^t n^{k-1} \rceil$) and $w > 0$. We get

$$\frac{3}{2}dwn^{k-1} < dn^{\frac{k-1}{t}} w \binom{n}{k-1}^{1-1/t} + t\binom{n}{k-1} \leq dn^{\frac{k-1}{t}} wn^{(k-1)(1-1/t)} + tn^{k-1}.$$

Finally, we substitute $t \leq \frac{1}{2}dw$, which follows from $w = \left\lceil \frac{2t}{d} \right\rceil$, obtaining

$$\frac{3}{2}dwn^{k-1} < dn^{\frac{k-1}{t}} wn^{(k-1)(1-1/t)} + \frac{1}{2}dwn^{k-1} = \frac{3}{2}dwn^{k-1},$$

which is a contradiction. □

This shows that we reach the recursive call in line 14 of Algorithm 2 at some iteration of the loop in line 11. The next step will be to show that this recursive call finds a $k-1$-partite $k-1$-graph in $H'$ of part sizes at least $t$. For this, we bound the density $d'$ of $H'$:

$$d' \geq \frac{s}{(n-w)^{(k-1)}} \geq \frac{d^t n^{(k-1)}}{n^{(k-1)}} = d^t,$$

and ensure that the associated part size

$$t' = t(n-w, d', k-1)$$

satisfies $t' \geq t$.

**Lemma 4.7.** *For all $k \geq 3$, $t' \geq t$.*

*Proof.* Substituting the new parameters into the definition, we get

$$t' = \left\lfloor \left( \frac{\log\left((n-w)/2^{(k-2)}\right)}{\log(3/d')} \right)^{\frac{1}{k-2}} \right\rfloor .$$

We start by using that $d' \geq d^t$ and that $w \leq n/2$:

$$t' \geq \left\lfloor \left( \frac{\log\left((n-w)/2^{(k-2)}\right)}{\log(3/d^t)} \right)^{\frac{1}{k-2}} \right\rfloor \geq \left\lfloor \left( \frac{\log\left(n/2^{(k-1)}\right)}{\log(3/d^t)} \right)^{\frac{1}{k-2}} \right\rfloor = \left\lfloor \left( \frac{\log\left(n/2^{(k-1)}\right)}{\log 3 - t\log d} \right)^{\frac{1}{k-2}} \right\rfloor .$$

Then, we substitute the definition of $t$, where removing the floor function maintains the inequality because the right hand side is decreasing in $t$ (recall $d \leq 1$):

$$t' \geq \left\lfloor \left( \frac{\log\left(n/2^{(k-1)}\right)}{\log 3 - \left(\frac{\log\left(n/2^{(k-1)}\right)}{\log(3/d)}\right)^{\frac{1}{k-1}} \log d} \right)^{\frac{1}{k-2}} \right\rfloor = \left\lfloor \left( \frac{\log\left(n/2^{(k-1)}\right)^{\left(1-\frac{1}{k-1}\right)}}{\frac{\log 3}{\log\left(n/2^{(k-1)}\right)^{\frac{1}{k-1}}} - \frac{\log d}{\log(3/d)^{\frac{1}{k-1}}}} \right)^{\frac{1}{k-2}} \right\rfloor . \tag{29}$$

Now we argue that $n/2^{k-1} \geq 3/d$. Otherwise, by Inequality (25), we would have

$$\frac{3}{n^{\frac{1}{2(k-1)}}} < d < \frac{3 \cdot 2^{(k-1)}}{n},$$

which implies

$$\sqrt{n} < n^{1-\frac{1}{2k-1}} \leq 2^{k-1} < k!,$$

so that

$$n < k!^2,$$

contradicting the minimum value of $n$ in Inequality (27).

This allows us to find a common denominator on the right side of (29):

$$t' \geq \left\lfloor \left( \frac{\log\left(n/2^{(k-1)}\right)^{\left(1-\frac{1}{k-1}\right)}}{\frac{\log 3 - \log d}{\log(3/d)^{\frac{1}{k-1}}}} \right)^{\frac{1}{k-2}} \right\rfloor = \left\lfloor \left( \frac{\log\left(n/2^{(k-1)}\right)^{\left(1-\frac{1}{k-1}\right)}}{\frac{\log(3/d)}{\log(3/d)^{\frac{1}{k-1}}}} \right)^{\frac{1}{k-2}} \right\rfloor = \left\lfloor \left( \frac{\log\left(n/2^{(k-1)}\right)}{\log(3/d)} \right)^{\frac{1}{k-1}} \right\rfloor = t.$$

□

This means that, assuming that the algorithm finds a $K\left(t', \overset{k-1}{\ldots}, t'\right)$ in $H'$ in the recursive call, it finds a $K\left(t, \overset{k}{\ldots}, t\right)$ in $H$. This argument only works if $k \geq 3$. For $k = 2$, the recursive call is handled by the base case in line 3 of Algorithm 2. Therefore, the part size of the (singleton) tuple returned by the recursive call is the number of (single-vertex) edges in $H'$, which is at least $s$. To ensure that the algorithm returns a $K(t,t)$ in this case, it suffices to show the following.

**Lemma 4.8.** *For $k = 2$, Algorithm 2 finds $s \geq t$.*

*Proof.* By way of contradiction, suppose that $t > s$. Substituting $k = 2$ into $s = \lceil d^t n^{k-1} \rceil$, we get $t > \lceil d^t n \rceil$ which implies

$$t > d^t n \geq d^{\frac{\log n}{\log(3/d)}} n = 3^{\frac{\log n}{\log(3/d)}} (d/3)^{\frac{\log n}{\log(3/d)}} n = 3^{\frac{\log n}{\log(3/d)}} \cdot \frac{1}{n} \cdot n = 3^{\frac{\log n}{\log(3/d)}} \geq 3^t,$$

which is false for all $t \geq 0$. $\square$

All in all, we can now state our main theorem.

**Theorem 4.9.** *Algorithm 2 finds a balanced partite $k$-graph in a $k$-graph $H$ with $n$ vertices and $m = dn^k$ with part size $t(n, d, k)$ in polynomial time, as long as $t(n, d, k) \geq 2$.*

*Proof.* To prove the correctness of the algorithm, we proceed by induction on $k$. If $k = 2$, it follows from Lemmas 4.6 and 4.8. Indeed, the algorithm finds $(V_1, T)$ with $|T| = t$ and $|V_1| \geq s \geq t$. Furthermore, $V_1$ is the set of vertices $x \in V \setminus W$ such that $\{x, y\} = \{x\} \cup \{y\} \in E$ for all $y \in T$. This means that $V_1 \times T \subset E(H)$.

If $k \geq 3$, Lemma 4.6 states that the algorithm reaches line 14 at some iteration of the loop. Furthermore, Lemma 4.7 states that the recursive call in line 14 has a part size $t' \geq t$. In particular, this means that $t' \geq 2$. Using the induction hypothesis for $k - 1$, this recursive call is successful and returns a tuple of sets $(X_1, X_2, \dots, X_{k-1}) \in \mathcal{P}(V)^{k-1}$ such that $|X_i| \geq t(n-w, d', k-1) \geq t$ for all $i$ and $X_1 \times \cdots \times X_{k-1} \subset E(H')$. However, by construction, $H' = L_{k-1}(H; T)$, which means that $T \times E(H') \subset E(H)$. All in all, the tuple $(X_1, \dots, X_{k-1}, T)$ returned in line 16 satisfies $X_1 \times \cdots \times X_{k-1} \times T \subset E = E(H)$, making the algorithm correct.

For the time complexity, note that all operations in the algorithm are in polynomial time, except for perhaps the for loop in line 11 and the recursive call in line 14.

We first argue that the for loop in line 11 runs in polynomial time. This is argued in the Mubayi and Turán paper [19], but we reproduce the argument here for completeness: The $t$-sets of $W$ can be enumerated in $\mathcal{O}\left(\binom{w}{t}\right)$ steps [22]. However, we can bound

$$\binom{w}{t} \leq \binom{2t/d + 1}{t} < \left(\frac{3et/d}{t}\right)^t = \left(\frac{3e}{d}\right)^t < e^{3t + t\log(1/d)} < e^{4\log n} = n^4.$$

Because there is only one recursive call, we can prove that it runs in polynomial time by induction on $k$. Clearly, if the algorithm runs in polynomial time for $k - 1$, it also runs in polynomial time for $k$. We can take as a base case $k = 1$, which has no recursive calls so it runs in polynomial time.

$\square$

# 5. Conclusions and Future Work

This thesis has focused on the algorithmic aspects of finding $k$-partite subgraphs in $k$-uniform hypergraphs, a problem central to degenerate Turán theory. We have presented a deterministic, polynomial-time algorithm (Algorithm 2) that, given a $k$-graph $G$ on $n$ vertices with $m$ edges, finds a complete balanced $k$-partite $k$-subgraph $K\left(t, \overset{k}{\ldots}, t\right)$. The part size $t$ achieved, given by Equation (21) as $t \approx (\log n/\log(1/d))^{1/(k-1)}$ where $d = m/n^k$, closely matches the parameters implicit in the non-constructive existence proofs of Erdős for such structures. This provides an efficient, constructive counterpart to these classical results, demonstrating that these subgraphs can indeed be located algorithmically within polynomial time. The recursive approach, which reduces the uniformity $k$ by analyzing appropriately defined link graphs, generalizes previous work for the $k = 2$ case by Mubayi and Turán. There are several avenues for future research stemming from this work:

1. **Tightening Bounds and Improving Practicality:** The proofs of correctness for Algorithm 2, particularly Lemmas 4.5 to 4.8, involve several inequalities. While sufficient to establish the polynomial runtime and the asymptotic nature of $t$, some of these bounds are quite loose (e.g., approximations of binomial coefficients, conditions for $w \leq n/2$, or the constants in the density arguments). A more meticulous analysis could potentially yield sharper constants in the definition of $t(n, d, k)$ or relax the minimum density requirements (Remark 4.3). This could make the algorithm applicable to sparser hypergraphs or guarantee larger $k$-partite structures for a given density, enhancing its practical significance for analyzing real-world hypergraphs which might not meet the currently proven, somewhat high, minimum vertex or density thresholds.

2. **Generalizing to unbalanced partite hypergraphs:** This approach taken in the proof of Theorem 3.14 can be generalized to give a lower bound on the number of copies (that is, embeddings with different image sets) of $K(t_1, \ldots, t_k)$ in a $k$-partite $k$-graph $G$ with different part sizes [5], therefore upper bounding all generalized Zarankiewicz numbers. Applying the same observations that we have made for the balanced case, we arrive at

$$\mathrm{ex}\left(n, K\left(t_1, \ldots, t_k\right)\right) = \mathcal{O}\left(n^{k - \frac{1}{\prod_{i=1}^{k-1} t_i}}\right). \tag{30}$$

   An algorithm finding these more general $k$-partite hypergraphs of growing sizes could be obtained by modifying the current algorithm and, for example, fixing the approximate ratio between the part sizes to be found.

3. **Exploring what happens when the density is not fixed:** There are versions of Theorem 3.16 that make $n_k$ depend on $\epsilon$ and show that the size of the guaranteed complete $k$-partite $k$-graph is at least $\delta n^{k - \frac{1}{t^{(k-1)}}}$, where $\delta \to \infty$ as $\epsilon \to 1$. This shows Algorithm 2 may not return a complete balanced $k$-partite $k$-graph of the best possible order of magnitude for a family of graphs with density tending to 1. An interesting question is whether the algorithm can be modified to find a partite hypergraph of approximately the right size in these circumstances, maybe losing the polynomial time complexity when this size is no longer logarithmic in $n$.

4. **Finding Blow-ups of General $k$-Graphs:** The presented algorithm is tailored to find blow-ups of a single edge, i.e., $K\left(t, \overset{k}{\ldots}, t\right)$. A natural extension would be to adapt this algorithmic framework to find $t_n$-blowups $H(t_n)$ of an arbitrary fixed $k$-graph $H$. As discussed in Theorem 3.17, $k$-graphs with

density $\pi(G)+\epsilon$ are known to contain $G(t_n)$ where $t_n = \delta(\log n)^{1/(|V(G)|-1)}$. Our current algorithm, if adapted, might yield a constructive proof for finding such blow-ups. However, for $k = 2$, it is known from Bollobás and Erdős [3] that the optimal growth for $t_n$ is $\delta \log n$, which is better than $(\log n)^{1/(|V(G)|-1)}$ if $|V(G)| > 2$. It remains open whether there is a way to adapt their proof into a constructive one, providing a polynomial-time algorithm.

5. **Implementation and Experimental Evaluation:** Implementing Algorithm 2 and evaluating its performance on various synthetic and real-world hypergraph datasets would be valuable. This could help identify practical bottlenecks and compare its findings with theoretical guarantees, especially concerning the constants involved in the calculation of $t$.

# 6. Bibliography

## References

[1] Rahil Baber and John Talbot. Hypergraphs do jump. *Combinatorics, Probability and Computing*, 20(2):161–171, 2011.

[2] Simeon Ball and Valentina Pepe. Asymptotic improvements to the lower bound of certain bipartite turán numbers. *Combinatorics, Probability and Computing*, 21(3):323–329, 2012.

[3] Béla Bollobás and Paul Erdős. On the structure of edge graphs. *Bull. London Math. Soc*, 5(317-321):31, 1973.

[4] William G Brown. On graphs that do not contain a thomsen graph. *Canadian Mathematical Bulletin*, 9(3):281–285, 1966.

[5] Matías Azócar Carvajal, Giovanne Santos, and Mathias Schacht. Canonical ramsey numbers for partite hypergraphs. *arXiv preprint arXiv:2411.16218*, 2024.

[6] David Conlon, Cosmin Pohoata, and Dmitriy Zakharov. Random multilinear maps and the erdős box problem. *arXiv preprint arXiv:2011.09024*, 2020.

[7] Dominique De Caen and Zoltán Füredi. The maximum size of 3-uniform hypergraphs not containing a fano plane. *Journal of Combinatorial Theory, Series B*, 78(2):274–276, 2000.

[8] P. Erdős. On extremal problems of graphs and generalized graphs. *Israel Journal of Mathematics*, 2(3):183–190, September 1964.

[9] Pál Erdős, Alfréd Rényi, and Vera T. Sós. On a problem of graph theory. *Studia Scientiarum Mathematicarum Hungarica*, 1:215–235, 1966.

[10] Paul Erdős and Miklós Simonovits. Supersaturated graphs and hypergraphs. *Combinatorica*, 3:181–192, 1983.

[11] Paul Erdös and Arthur H Stone. On the structure of linear graphs. 1946.

[12] Peter Frankl and Zoltán Füredi. An exact result for 3-graphs. *Discrete mathematics*, 50:323–328, 1984.

[13] Zoltán Füredi. New asymptotics for bipartite turán numbers. *Journal of Combinatorial Theory, Series A*, 75(1):141–144, 1996.

[14] C. Hyltén-Cavallius. On a combinatorical problem. *Colloquium Mathematicae*, 6(1):61–65, 1958.

[15] Peter Keevash. Hypergraph turan problems. *Surveys in combinatorics*, 392:83–140, 2011.

[16] János Kollár, Lajos Rónyai, and Tibor Szabó. Norm-graphs and bipartite turán numbers. *Combinatorica*, 16(3):399–406, 1996.

[17] T. Kővári, V. T. Sós, and P. Turán. On a problem of k. zarankiewicz. *Colloquium Mathematicae*, 3(1):50–57, 1954.

[18] Michael Krivelevich and Benny Sudakov. The chromatic numbers of random hypergraphs. *Random Structures & Algorithms*, 12(4):381–403, 1998.

[19] Dhruv Mubayi and György Turán. Finding bipartite subgraphs efficiently. *Information Processing Letters*, 110(5):174–177, 2010.

[20] Alexander A Razborov. Flag algebras. *The Journal of Symbolic Logic*, 72(4):1239–1282, 2007.

[21] Alexander A Razborov. On 3-hypergraphs with forbidden 4-vertex configurations. *SIAM Journal on Discrete Mathematics*, 24(3):946–963, 2010.

[22] Edward M Reingold, Jurg Nievergelt, and Narsingh Deo. *Combinatorial algorithms: theory and practice*. Prentice Hall College Div, 1977.

[23] Vojtěch Rödl and Mathias Schacht. Complete partite subgraphs in dense hypergraphs. *Random Structures & Algorithms*, 41(4):557–573, 2012.

[24] Paul Turán. Eine extremalaufgabe aus der graphentheorie. *Mat. Fiz. Lapok*, 48(436-452):61, 1941.

# A. Properties of Hypergraph Embeddings

**Proposition A.1.** *Hypergraph inclusion ($\subset$) and induced hypergraph inclusion ($\subset_{ind}$) are preorder relations on $k$-graphs.*

*Proof.* We need to show that the relations are reflexive and transitive. Reflexivity is clear, as the identity map is an induced embedding of a $k$-graph in itself. Let $G, H$, and $K$ be $k$-graphs with vertex sets $X, Y$, and $Z$ respectively. If $G \subset H$ via $f : X \to Y$ and $H \subset K$ via $g : Y \to Z$, then $g \circ f : X \to Z$ is injective and satisfies that for each edge $e \in E(G)$,

$$g \circ f(e) = \{g(f(x)) \mid x \in e\} = \{g(y) \mid y \in f(e)\} \in E(K),$$

because $f(e) \in E(H)$. Therefore, $G \subset K$ via $g \circ f$. If the embeddings are induced, and $e$ is an edge in $E(K[g \circ f(X)])$, then $e$ is also an edge in $E(K[g(Y)]) = g(E(H))$. Therefore, $e' = g^{-1}(e)$ is an edge in $H$. Furthermore, because $e = g(e') \subset g \circ f(X)$, we have that $e' \in E(H[f(X)]) = f(E(G))$, so $e \in g(f(E(G)))$. □

*Remark* A.2. In Definition 2.6, given that a map $f : V \to W$ is an embedding (and therefore injective), a different way to state that it is an induced embedding is to say that $f^{-1} : H[f(V)] \to G$ is also an embedding.

**Proposition A.3.** *The relation of isomorphism ($\cong$) is an equivalence relation on $k$-graphs.*

*Proof.* The relation is reflexive via the identity map. If $f : G \to H$ is an isomorphism, then $f^{-1} : H \to G$ is also an isomorphism, so the relation is symmetric. Finally, if $f : G \to H$ and $g : H \to K$ are isomorphisms, then $g \circ f : G \to K$ is also an isomorphism, because it is bijective, and by Proposition A.1, it is an embedding and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ is also an embedding. By Remark A.2, we are done. □

**Proposition A.4.** *Let $G, G', H, H'$ be $k$-graphs such that $G \cong G'$ and $H \cong H'$. Then,*

1. *$G \subseteq H$ if and only if $G' \subseteq H'$.*
2. *$G \subseteq_{ind} H$ if and only if $G' \subseteq_{ind} H'$.*

*Proof.* Because the isomorphism relation is symmetric, we only need to show one direction of each implication. let $f : V(G) \to V(H)$ be an embedding of $G$ in $H$, and let $g : V(G) \to V(G')$ and $h : V(H) \to V(H')$ be isomorphisms between the respective graphs. We claim that the composition

$$f' = h \circ f \circ g^{-1} : V(G') \to V(H')$$

is an embedding of $G'$ in $H'$. Injectivity is given by the injectivity of $h, f$, and $g^{-1}$. By Proposition A.3, we have that $g^{-1}$ is an isomorphism of $G'$ in $G$, and in particular an embedding. Therefore, by Proposition A.1, $f'$ is an embedding of $G'$ in $H'$, proving part (1). Suppose now that the embedding $f$ is induced. consider the maps

$$(f')^{-1} : f'(V(G')) \to V(G')$$

and

$$\varphi = g \circ f^{-1} \circ h^{-1} : h \circ f(V(G)) \to V(G'),$$

where we restrict the domain of $f^{-1}$ to $f(V(G))$. Because $g$ is a bijection, $V(G) = g^{-1}(V(G'))$ so the domain of $\varphi$ is $f'(V(G'))$. In fact, one can check that the two functions are identical. Because $f^{-1}$ is an embedding of $H[f(V(G))]$ in $G$, we can argue as in the first case that $(f')^{-1}$ is an embedding of $H'[f'(V(G'))]$ in $G'$, and therefore $f'$ is an induced embedding of $G'$ in $H'$. This concludes the proof of part (2). □

Propositions A.1, A.3 and A.4, establish that the properties of containing a $k$-graph $G$ as a (induced) subgraph within a $k$-graph $H$ depend only on the isomorphism classes of $G$ and $H$. Therefore, discussions of (induced) subgraph containment can be conducted up to isomorphism.