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

Keywords

hypergraph, algorithm, graph, partite, extremal

Contents

1. Introduction

TODO: Write introduction

2. Basic Definitions and Results

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 finite *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

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

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

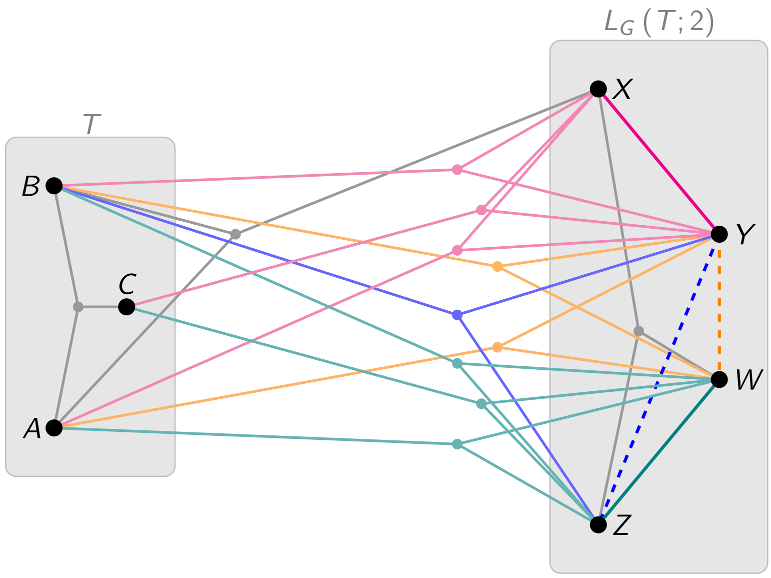

Figure 1: A 3-graph G and the common 2-link graph $L_G(T;2)$ of the set $T = \{A, B, C\}$. The link graph has vertex set $\{X, Y, Z, W\}$ and edge set $\{\{X, Y\}, \{W, Z\}\}$.

Remark 2.6. Definition 2.5 might appear confusing because E' is defined as a set of subsets of V, while the link graph has vertex set $V \setminus T$. In fact, it is impossible for an element $Y \in E'$ to contain a vertex $v \in T$. To see this, suppose that $v \in Y$. We can pick a $(k-j)$-set $X \in \binom{T}{k-j}$ such that $v \in X$. Then $v \in X \cap Y$ so $|X \cup Y| < j + (k-j) = k$, contradicting the fact that $X \cup Y$ is an edge in G.

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

Definition 2.7. 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

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

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

$$f(E) \subset E(H[f(V)]). \tag{1}$$

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 $G \subset H$. If, furthermore,

$$f(E) = E(H[f(V)]), \tag{2}$$

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 2.8. Condition (1) of definition 2.7 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.

Remark 2.9. In Definition 2.7, 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.

For the notions that we have introduced to be useful, we need to show some basic properties.

Proposition 2.10. *Graph inclusion* ($\subset$) *and induced graph 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 $e \in g(f(E(G)))$. □

Proposition 2.11. *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 2.10, it is an embedding and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ is also an embedding. By remark 2.9, we are done. □

Proposition 2.12. *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 2.11, we have that g^{-1} is an isomorphism of G' in G, and in particular an embedding. Therefore, by Proposition 2.10, 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 φ 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). $\square$

Propositions 2.11, 2.10, and 2.12 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.

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 2.13. 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 lets us talk, up to isomorphism, about *the* complete k-graph on r vertices $K_r^{(k)}$. For example, in Figure 2 we show the complete graph $K_4^{(3)}$.

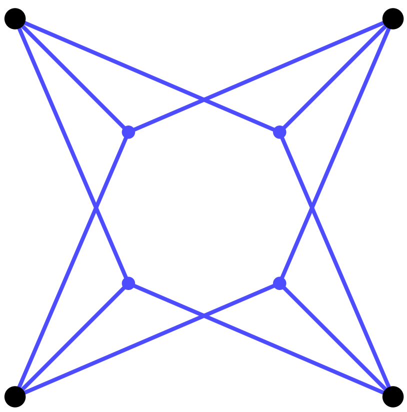

Figure 2: A complete 3-graph on 4 vertices.

Remark 2.14. 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 k-graphs and the chromatic number

We now introduce the notion of partite k-graphs.

Definition 2.15. for an integer $p \geq k$, a k-graph $G = (V, E)$ is *p-partite* if there exists a partition $V = V_1 \cup \cdots \cup V_p$ such that every edge $e \in E$ intersects every part V_i in at most one vertex. We may write $G = (V_1, \ldots, V_p; E)$ and say that G is a partite k-graph on $V_1, \ldots, V_p$.

A different view of p-partite k-graphs is to think of them as vertex colored k-graphs, in the following sense.

Definition 2.16. A *proper vertex coloring* of a k-graph $G = (V, E)$ in p colors is a map $\chi : V \to [p]$ such that χ is injective on every edge. The elements $\chi^{-1}(i)$ for $1 \leq i \leq p$ are called the *color classes* of G. We say that G is *p-vertex-colorable* or *p-colorable*.

The two notions are equivalent in the following sense.

Proposition 2.17. *A k-graph $G = (V, E)$ is p-partite if and only if there exists a vertex coloring of G in p colors.*

Proof. Let $G = (V, E)$ be a k-graph. if G is p-partite (say, $G = (V_1, \dots, V_p; E)$), then we can construct a vertex coloring of G in p colors by assigning to the vertices in V_i the color i. The coloring is proper by Definition 2.15. Conversely, let $\chi : V \to [p]$ be a proper vertex coloring of G in p colors. Then the color classes form a partition of V into p sets where, by the injectivity of χ on edges, every edge $e \in E$ intersects every color class in at most one vertex. □

To measure how easily colorable a k-graph is, we define the following.

Definition 2.18. Let $G = (V, E)$ be a k-graph. The *chromatic number* $\chi(G)$ of G is the smallest integer $p \geq k$ such that G is p-colorable (or, equivalently p-partite).

Remark 2.19. The chromatic number is increasing with respect to inclusion. This is because if

$$f : V(G) \to V(H)$$

is an embedding of G in H, and

$$\chi : V(H) \to [p]$$

is a proper vertex coloring of H, then

$$\chi \circ f : V(G) \to [p]$$

is a proper vertex coloring of G. Therefore, if H is p-colorable, so is G, and $\chi(G) \leq \chi(H)$.

If $G = (V_1, \dots, V_k; E)$ 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 2.20. 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 2.21. Let $p \geq k \geq 1$. A p-partite k-graph $G = (V_1, \dots, V_p; E)$ is *complete* if

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

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

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

$$K^{(k)}(V_1, \dots, V_p) \cong K^{(k)}(W_1, \dots, W_p).$$

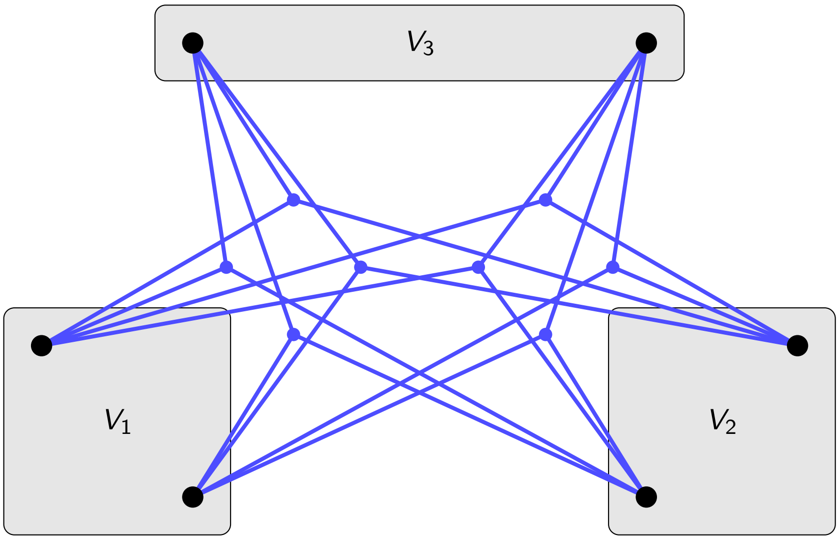

Figure 3: The complete 3-partite 3-graph $K(2,2,2)$, with parts V_1, V_2, V_3.

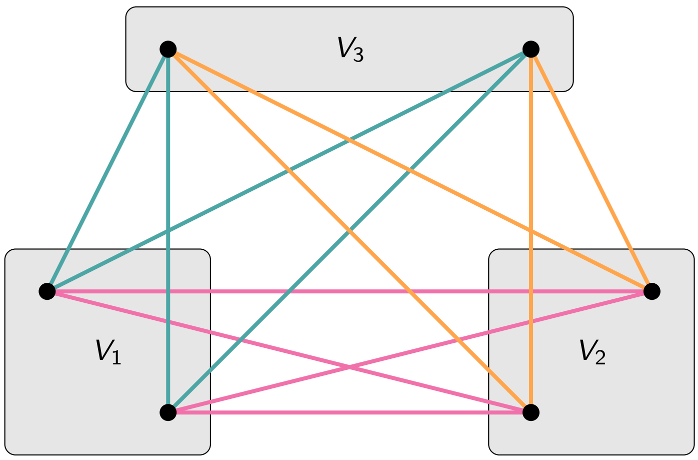

Figure 4: The complete 3-partite 2-graph $K^{(2)}(2,2,2)$, with parts V_1, V_2, V_3.

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

$$K^{(k)}\left(a_1, \dots, a_p\right),$$

or, in the k-partite case, by

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

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 $K^{(3)}(2, 2, 2)$.

2.3 Turán Problem for k-graphs

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 2.22. *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 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 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 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 2.10). □

We call the integer $\mathrm{ex}(n, G)$ the *Turán number* of G on n vertices. It is clearly 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 extremal numbers were first studied by Turán [7] in 1941. The result is the following.

Theorem 2.23 (Turán Theorem)**.** *Let $r > 2$ be an integer and let $n \geq r$. Let $a_1, \dots, a_{r-1}$ 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*

$$ex\left(n, K_r^{(2)}\right) = \sum_{\{x,y\} \in \binom{[r-1]}{2}} a_x \cdot a_y \tag{3}$$

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

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

To prove this theorem, we suppose that $G = (V, E)$ is a 2-graph with n vertices and $\mathrm{ex}\left(n, K_r^{(2)}\right)$ edges, and use the following two lemmas.

Lemma 2.24. *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, against 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 $v \in T' \setminus \{x'\}$, against the completeness of $G'[T']$. □

Lemma 2.25. *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 $\{\{v, z_i\} | \{v, z\} \in E, i \in \{1, 2\}\}$.

First, we show that G' contains no embedding of $K_r^{(2)}$. We make a similar argument as in the proof of Lemma 2.24. 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 $T' = \{x_1, x_2, x_3, \dots, x_{r-1}, z_1\}$, 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 $G'[\{x_1, x_2, x_3 \dots, x_{r-1}, z\}] = G[\{x_1, x_2, x_3 \dots, x_{r-1}, z\}]$ is complete, against our assumption.

Now, we show that G' has more edges than G. By Lemma 2.24, $d_G(x) = d_G(z)$ 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 we wanted to show. □

Now we are ready to prove Theorem 2.23.

Proof of Turán Theorem. We have shown in Lemma 2.25 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 the first is $(r-1)$-colorable and the second is not.

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 (3). The imposition that $\lfloor n/(r-1) \rfloor \leq a_i \leq \lceil n/(r-1) \rceil$ for all i is equivalent to saying that the part sizes are as equal as possible, that is, $|a_i - a_j| < 1$ 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 with the number of edges in G being maximal. $\square$

Because of the difficulty of finding exact extremal 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 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 (3), we can see that $a_i = n/(r-1) + \mathcal{O}(1)$ for all i. Therefore,

$$\mathrm{ex}\left(n, K_r^{(2)}\right) = \sum_{\{x,y\} \in \binom{[r-1]}{2}} a_x \cdot a_y = \binom{r-1}{2} \cdot \left(\frac{n}{r-1} + \mathcal{O}(1)\right)^2 = \frac{(r-2)}{2(r-1)} n^2 + \mathcal{O}(n). \tag{4}$$

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 lets us restate equation (4) as

$$\mathrm{ex}\left(n, K_r^{(2)}\right) = \frac{r-2}{r-1}\binom{n}{2} + \mathcal{O}(n) = \frac{r-2}{r-1}\binom{n}{2} + o\left(\binom{n}{2}\right), \tag{5}$$

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

$$\pi(G) = \lim_{n \to \infty} \frac{ex(n, G)}{\binom{n}{k}} \tag{6}$$

exists and is between 0 *and* 1*. We call 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 2.22. Furthermore, it is less than 1 for all $n \geq |V(G)| + 1$. To see this, consider a 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) it is non-increasing, so it must converge to a value $0 \leq \pi(G) < 1$. Let $n \geq |V(G)|$. There exists a graph $H = (W, F)$ with $n+1$ vertices and $\mathrm{ex}\,(n+1, G)$ edges does not contain G as a subgraph. For each vertex $w \in W$, the 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

$$\mathcal{P} = \{(w, e) | w \in W, e \in E(H_w)\}\,.$$

Counting on the first coordinate, we get

$$|\mathcal{P}| = \sum_{w \in W} |E(H_w)| \leq (n+1)\,\mathrm{ex}\,(n, G)\,. \tag{7}$$

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

$$|\mathcal{P}| = (n+1-k)|F| = (n+1-k)\,\mathrm{ex}\,(n+1, G)\,. \tag{8}$$

Combining equations (7) and (8), 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

$$a_{n+1} = \frac{\mathrm{ex}\,(n+1, G)}{\binom{n+1}{k}} \leq \frac{(n+1)\,\mathrm{ex}\,(n, G)}{(n+1-k)\binom{n+1}{k}} = \frac{\mathrm{ex}\,(n, G)}{\binom{n}{k}} = a_n.$$

□

We can now summarize expression (5) as follows.

Corollary 2.27. *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 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 2.28. *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 2.23 directly shows $\pi(G) > 0$. Indeed, for all $m$the 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 $km = \left|V\left(K\left(m, \overset{k}{\ldots}, m\right)\right)\right|$ as large as we want, the limit (6) bounded below by a positive constant. We defer the proof of the other direction to subsection 2.4, where we study k-partite k-graph Turán problems in more depth (in particular, see Theorem 2.35). □

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 2.29 (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 end of this section, where we will have more powerful tools at our disposal. Note that letting $G = K_r^{(2)}$ we recover Corollary 2.27 of Theorem 2.23.

2.4 Degenerate Turán Problems

Remark 2.30. All k-partite k-graphs with part sizes $b_1 \le a_1, \dots, b_k \le a_k$ are contained in $K(a_1, \dots, a_k)$ as subgraphs. This lets us follow the same argument as in Proposition 2.22 to define the following.

Definition 2.31. Let $0 < t_1 \le v_1, \dots, 0 < t_k \le v_k$ be integers. Then the *generalized Zarankiewicz number* $z(v_1, \dots, v_k; t_1, \dots, t_k)$ is the largest integer $0 \le z < \prod_i v_i$ 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 $|T_i| = t_i$ in it 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 onto another we assume the condition $f(T_i) \subset V_i$.

Remark 2.32. Finding this number can help us upper bound the extremal number of $K(t_1, \dots, t_k)$ asymptotically: 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 $\lfloor n/k \rfloor \le v_i \le \lceil n/k \rceil$) Let $V_1, \dots, V_k$ be a random partition of $V(H)$ with $|V_i| = v_i$. for an edge $e \in E(H)$, the probability that e is an edge in $K(V_1, \dots, V_k)$ is greater than

$$k! \prod_i \frac{v_i}{n} \sim \frac{k!}{k^k},$$

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, we can conclude that

$$\operatorname{ex}(n, K(t_1, \dots, t_k)) = \mathcal{O}\left(z\left(\lceil n/k \rceil, \overset{k}{\dots}, \lceil n/k \rceil; t_1, \dots, t_k\right)\right).$$

The problem on 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 [4] in 1954. This was generalized to arbitrary complete bipartite 2-graphs by C. Hyltén-Cavallius [3] in 1958. The result is stated and proved here for completeness.

Theorem 2.33. *Let $0 < s \le u$ and $0 < t \le w$ be integers. Then*

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

Proof. Suppose that we have a bipartite graph $G = (U, W; E)$ with $|U| = u$, $|W| = w$ and $|E| = z$ exceeding the bound stated above. Let us consider the set

$$P = \left\{ (x, Y) \in U \times \binom{W}{t} \middle| x \in N_G(Y) \right\}.$$

Counting on the first coordinate, we get

$$|P| = \sum_{x \in U} \binom{d_G(x)}{t} = \sum_{x \in U} \varphi(d_G(x)) \geq u \sum_{x \in U} \varphi(z/u) = u\binom{z/u}{t}, \tag{9}$$

where we define

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

The function φ is convex, so we get the inequality in (9) 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 $\{a, b\} \in E$. This would mean that the inclusion map from $S \cup T$ to $U \cup W$ is an embedding of $K(s, t)$ in G, as described in Definition 2.31. Supposing that this is not the case, by the averaging we get

$$|P| \leq (s-1)\binom{w}{t}. \tag{10}$$

Putting inequalities (9) and (10) together, we get

$$u\binom{z/u}{t} \leq (s-1)\binom{w}{t}. \tag{11}$$

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

$$\frac{(z/u-(t-1))^t}{\binom{z/u}{t}} \leq \frac{(w-(t-1))^t}{\binom{w}{t}}, \tag{12}$$

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.

$$g(x) = \prod_{i=0}^{t-1} (x-(t-1)) \frac{i+1}{x-i} = t! \prod_{i=0}^{t-1} \frac{x-(t-1)}{x-i}. \tag{13}$$

Every factor in the product on the right side of (13) is increasing in x for $x \geq t-1 \geq i$, proving the claim. Multiplying inequalities (11) and (12) 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 with our assumption. $\square$

Remark 2.34. Following Remark 2.32, we can use this bound to get an upper bound on the extremal 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 [2] generalized this result to arbitrary complete partite k-graphs in the following theorem.

Theorem 2.35. *For $k \geq 2$ and $2 \leq t \leq \frac{n}{k}$, $\operatorname{ex}\left(n, K\left(t, \overset{k}{\ldots}, t\right)\right) = \mathcal{O}\left(n^{k-\frac{1}{t^{k-1}}}\right)$.*

Proof. By Remark 2.32, it suffices to show that

$$z = z\left(w, \overset{k}{\ldots}, w; t, \overset{k}{\ldots}, t\right) = \mathcal{O}\left(w^{k-\frac{1}{t^{k-1}}}\right)$$

as $w \to \infty$. We prove this by induction on k. For $k = 2$, this is obtained by setting $u = w$ and $s = t$ in Theorem 2.33. For $k > 2$, suppose by way of contradiction that the theorem is false. For all $w_0 \in \mathbb{N}$, $K \in \mathbb{R}^+$, there exists a k-partite k-graph $G = (W_1, \ldots, W_k; E)$ with part sizes $|W_i| = w \geq w_0$ and $|E| \geq Kw^{k-\frac{1}{t^{k-1}}}$ such that no embedding of $K\left(t, \overset{k}{\ldots}, t\right)$ in it exists. Consider, for each set $T \in \binom{W_k}{t}$, the associated $(k-1)$-link $L_G(T; k-1)$. We claim that it does not contain $K\left(t, \overset{k-1}{\ldots}, t\right)$ as a subgraph. If it did (say, $T_1 \times \cdots \times T_{k-1} \in E(L_G(T; k-1))$), then $T_1 \times \cdots \times T_{k-1} \times T \in E$ would contradict the assumption that G does not contain $K\left(t, \overset{k}{\ldots}, t\right)$ as a subgraph. This means that

$$L_G(T; k-1) \text{ has at most } z' \text{ edges for all } T \in \binom{W_k}{t}, \tag{14}$$

where

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

Now, consider the bipartite graph $G' = (U, W; E')$, where

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

Condition (14) is equivalent to saying that there is no embedding of $K(z'+1, t)$ onto G' respecting the respective partitions. Furthermore, G' has the same number of edges as G. Finally, we invoke Theorem 2.33 with $u = |U| = w^{k-1}$ and $s = z' + 1$ to get

$$Kw^{k-\frac{1}{t^{k-1}}} \leq |E| = |E'| \leq (z')^{1/t}(w - t + 1)w^{(k-1)(1-1/t)} + (t-1)w^{k-1}. \tag{15}$$

By the inductive hypothesis, for w_0 and K' large enough, we can bound

$$z' \leq K'w^{(k-1)-\frac{1}{t^{k-2}}}. \tag{16}$$

Substituting inequality (16) into (15) and approximating yields

$$Kw^{k-\frac{1}{t^{k-1}}} \leq (K')^{1/t}w^{k-\frac{1}{t^{k-1}}} + (t-1)w^{k-1}.$$

Combining like terms and picking $K > 2(K')^{1/t}$ gives

$$\frac{1}{2}Kw^{k-\frac{1}{t^{k-1}}} < (t-1)w^{k-1},$$

which we can rewrite as

$$\frac{1}{2}Kw^{1-\frac{1}{t^{k-1}}} < (t-1). \tag{17}$$

This is a contradiction, because the right side of inequality (17) is constant in w, while the left side grows to infinity as w increases. □

This approach can be generalized to give a lower bound on the number of copies of $K(t_1, \dots, t_k)$ in a k-partite k-graph G with different part sizes [1], therefore upper bounding all generalized Zarankiewicz numbers.

Degenerate Turán problems are also of interest as tools to solve non-degenerate ones. A classic example is the following.

Proof of the Erdős-Stone-Simonovits Theorem (Theorem 2.29). TODO □

3. Our Algorithm

Let $G = (V, E)$ 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 G with all part sizes at least

$$t = t(n, d, k) = \left\lfloor \left(\frac{\log(n/2^{k-1})}{\log(3/d)} \right)^{\frac{1}{k-1}} \right\rfloor, \tag{18}$$

where

$$d = m/n^k \tag{19}$$

is the "un-normalized" density of G, which 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 G). More precisely, we show the following:

Theorem 3.1. *There is an algorithm that, given a k-graph G satisfying the conditions above, finds a complete balanced k-partite k-graph in G with part sizes $t = t(n, d, k)$. That is, the algorithm returns a tuple of sets $(V_1, \dots, V_k) \subset \binom{V}{t}^k$ such that $V_1 \times \cdots \times V_k \subset E$. Furthermore, the algorithm's runtime is polynomial in n.*

Remark 3.2. The stated condition implies that the sets $V_1, \dots, V_k$ 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 $(v, v, v_3, \dots, v_k) \in V_1 \times \cdots \times V_k$ has size $k - 1$ as an unordered set so it cannot be an edge in G. This means that the inclusion map from $K(V_1, \dots, V_k)$ to V defines an embedding, as desired.

This gives a constructive proof of Theorem 2.35. Indeed, suppose we have a fixed value for t. For n large enough, we may choose d such that $t(n, d, k) \geq t$. By our definition of t, we only need that

$$d \geq 3 \left(\frac{2^{k-1}}{n} \right)^{\frac{1}{t^{k-1}}},$$

which is satisfied for

$$m = \left\lceil 3n^k \left(\frac{2^{k-1}}{n} \right)^{\frac{1}{t^{k-1}}} \right\rceil = \mathcal{O}\left(n^{k - \frac{1}{t^{k-1}}} \right).$$

The construction in Theorem 3.1 then proves

$$\mathrm{ex}\left(n, K\left(t, \overset{k}{\dots}, t\right)\right) < m = \mathcal{O}\left(n^{k - \frac{1}{t^{k-1}}} \right).$$

For $k = 2$, this problem was already solved by an algorithm of Mubayi and Turán [5], 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 $V(G)$ 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 $G = (V, E)$ 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 G
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 is required because if $d = o\left(n^{-1/2}\right)$ then there may not even be a $K(2, 2)$ in G. 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 2.33 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

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

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 Theorem 2.35. This is the algorithm mentioned in Theorem 3.1, and the main contribution of this work.

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

$$|W| = w = w(n, d, k) = \left\lceil \frac{2t(n, d, k)}{d} \right\rceil . \tag{20}$$

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. For a specific T, the set S satisfies

$$|S| \geq s = s(n, d, k) = \left\lceil d^{t(n,d,k)} n^{k-1} \right\rceil . \tag{21}$$

We define a new $(k-1)$-graph G' with vertex set $V \setminus W$ and edge set S. As it turns out, S is large enough (21) that applying the algorithm recursively to G' yields a $K\left(t', \overset{k-1}{\dots}, t'\right)$ in G' 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 k-tuple $(X_1, \dots, X_{k-1}, T)$, of t*−sets* of V which by the definition of S satisfies $X_1 \times \cdots \times X_{k-1} \times T \subset E = E(G)$ so it forms a $K\left(t, \overset{k}{\ldots}, t\right)$ in G.

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

1: **function** $\mathrm{FIND\_PARTITE}(G, k)$
2: **assert** G is a k-graph
3: **if** $k = 1$ **then**
4: **return** $(\{x : \{x\} \in E(G)\})$
5: **end if**
6: $V \leftarrow V(G), E \leftarrow E(G), n \leftarrow |V|,\ m \leftarrow |E|,\ d \leftarrow m/n^k$
7: $t \leftarrow t(n, d, k),\ w \leftarrow w(n, d, k),\ s \leftarrow s(n, d, k)$
8: **assert** $t \geq 2$
9: $W \leftarrow$ a set of w vertices with highest degree in G
10: $U \leftarrow \binom{V \setminus W}{k-1}$
11: **for all** $T \in \binom{W}{t}$ **do**
12: $S \leftarrow \{\, y \in U \colon \{x\} \cup y \in E$ for all $x \in T \,\}$
13: **if** $|S| \geq s$ **then**
14: $G' \leftarrow (V \setminus W, S)$
15: $(V_1, \dots, V_{k-1}) \leftarrow \mathrm{FIND\_PARTITE}(G', k-1)$
16: **return** $(V_1, \dots, V_{k-1}, T)$
17: **end if**
18: **end for**
19: **end function**

The implementation of the algorithm 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 Algorithm 2 the size reduction of the sets obtained from the recursive call. The algorithm as stated 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 this algorithm 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 3.1. 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 3.3. The requirement $t \geq 2$ is met whenever

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

However, d satisfies

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

so we get a minimum value of n:

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

This also lets us prove the bound

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

for all $k \geq 2$. We have already seen that this is the case for $k = 2$. 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 $n < 4$.

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 3.4. *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

$$\frac{n}{2} \leq w = \left\lceil \frac{2t}{d} \right\rceil \leq 1 + \frac{2t}{d} < 1 + \frac{2\log(n/2)\sqrt{n}}{3} = 1 + \frac{n}{4}.$$

This implies that $n < 4$, in contradiction to Remark 3.3. It is also clear from Remark 3.3 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 (21) 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 $w < \frac{n}{2}$,

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

which implies

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

In the last inequality, we have used that $t \geq 2$ and that $d \leq \frac{1}{k!}$. Since $k!^2 \geq (2k)^{k-1}$ 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 3.5. *Given $W \subset V$ as defined in line 9 of Algorithm 2, There are at least $\frac{3}{2}dwn^{k-1}$ edges of G 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 $2 \leq j \leq n$, 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 $\binom{w}{j}\binom{n-w}{k-j} \leq \frac{w^j n^{k-j}}{j!} \leq \frac{w^j n^{k-j}}{j}$ of them. Thus, the total contribution of these edges is at most $w^j n^{k-j} \leq w^2 n^{k-2}$. The number of edges with only one vertex in W is then at least

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

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

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

which implies

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

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

$$nd^2 < 6t.$$

We now closely follow the steps of Mubayi and Turán [5].

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

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

which is a contradiction.

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

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 3.6. *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 2.33 to the 2-partite 2-graph

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

where F is defined as

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

By Lemma 3.5, $\mathcal{P}$ 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 2.33 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 $w > 0$:

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

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

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

which is a contradiction. $\square$

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 G' of part sizes at least t. For this, we bound the density d' of G':

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

and ensure that the associated part size

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

satisfies $t' \geq t$.

Lemma 3.7. *For all $k \geq 3$, $t' \geq t$.*

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

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

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

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

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 $d \leq 1$):

$$t' \geq \left\lfloor \left(\frac{\log(n/2^{k-1})}{\log 3 - \left(\frac{\log(n/2^{k-1})}{\log(3/d)}\right)^{\frac{1}{k-1}} \log d} \right)^{\frac{1}{k-2}} \right\rfloor = \left\lfloor \left(\frac{\log(n/2^{k-1})^{1-\frac{1}{k-1}}}{\frac{\log 3}{\log(n/2^{k-1})^{\frac{1}{k-1}}} - \frac{\log d}{\log(3/d)^{\frac{1}{k-1}}}} \right)^{\frac{1}{k-2}} \right\rfloor. \tag{22}$$

Now we argue that $n/2^{k-1} \geq 3/d$. Otherwise, by Remark 3.3, we would have

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

which implies

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

so that

$$n < k!^2,$$

against the minimum value of n in Remark 3.3.

This allows us to find a common denominator on the right side of (22):

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

This means that, assuming that the algorithm finds a $K\left(t', \overset{k-1}{\ldots}, t'\right)$ in G' in the recursive call, it finds a $K\left(t, \overset{k}{\ldots}, t\right)$ in G. 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 G', 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 3.8. *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

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

which is false for all $t \geq 0$. □

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

Theorem 3.9. *Algorithm 2 finds a balanced partite k-graph in a k-graph G 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 3.6 and 3.8. 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 $V_1 \times T \subset E(G)$.

If $k \geq 3$, Lemma 3.6 tells us that the algorithm reaches line 14 at some iteration of the loop. Furthermore, Lemma 3.7 tells us 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 $(X_1, X_2, \ldots, X_{k-1}) \in \mathcal{P}(V)^{k-1}$ such that $|X_i| \geq t(n-w, d', k-1) \geq t$ for all i and $X_1 \times \cdots \times X_{k-1} \subset E(G')$. However, by construction of G',

$$E(G') = S = \left\{ x \in \binom{V \setminus W}{k-1} : \{x\} \cup y \in E \text{ for all } y \in T \right\}.$$

That is, the tuple $(X_1, \ldots, X_{k-1}, T)$ returned in line 16 satisfies $X_1 \times \cdots \times X_{k-1} \times T \subset E = E(G)$, 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 [5], but we reproduce the argument here for completeness: As seen in [6], the t-sets of W can be enumerated in $O\left(\binom{w}{t}\right)$ steps. 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.

□

4. Bibliography

References

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

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

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

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

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

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

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