

Turán numbers of Berge trees

Ervin Győri^{1,2}

Nika Salia^{1,2}

Casey Tompkins^{3,4}

Oscar Zamora^{2,5}

¹Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences.

`gyori.ervin@renyi.mta.hu`, `nika@renyi.hu`

²Central European University, Budapest.

`oscar.zamoraluna@ucr.ac.cr`

³Discrete Mathematics Group, Institute for Basic Science (IBS), Daejeon, Republic of Korea.

`ctompkins496@gmail.com`

⁴Karlsruhe Institute of Technology, Germany.

⁵Universidad de Costa Rica, San José.

April 16, 2020

Abstract

A classical conjecture of Erdős and Sós asks to determine the Turán number of a tree. We consider variants of this problem in the settings of hypergraphs and multi-hypergraphs. In particular, for all k and r , with $r \geq k(k-2)$, we show that any r -uniform hypergraph $\mathcal{H}$ with more than $\frac{n(k-1)}{r+1}$ hyperedges contains a Berge copy of any tree with k edges different from the k -edge star. This bound is sharp when $r+1$ divides n and for such values of n we determine the extremal hypergraphs.

1 Background

We recall a classic theorem of Erdős and Gallai [3].

Theorem 1 (Erdős, Gallai [3]). *Let n, k be positive integers and let G be an n -vertex graph containing no path of k edges, then*

$$e(G) \leq \frac{(k-1)n}{2}.$$

Equality is obtained if and only if k divides n and G is the graph consisting of n/k disjoint complete graphs of size k .

Erdős and Sós [2] conjectured that the same bound would hold for any tree with k edges. A proof of this conjecture for sufficiently large k was announced in the 90's by Ajtai, Komlós, Simonovits and Szemerédi. We will consider a variant of this problem in the setting of hypergraphs and multi-hypergraphs. We obtain exact results for the case of large uniformity.

Given a hypergraph $\mathcal{H}$, we denote the vertex and edge sets of $\mathcal{H}$ by $V(\mathcal{H})$ and $E(\mathcal{H})$, respectively. We denote the number of vertices and hyperedges by $v(\mathcal{H}) = |V(\mathcal{H})|$ and $e(\mathcal{H}) = |E(\mathcal{H})|$. A

hypergraph is said to be r -uniform if all of its hyperedges have size r . We now provide some definitions which we will need.

Definition 1. For a given uniformity r and a fixed graph G , an r -uniform multi-hypergraph $\mathcal{H}$ is a Berge copy of G , if there exists an injection $f_1 : V(G) \rightarrow V(\mathcal{H})$ and a bijection $f_2 : E(G) \rightarrow E(\mathcal{H})$, such that if $e = \{v_1, v_2\} \in E(G)$, then $\{f_1(v_1), f_1(v_2)\} \subseteq f_2(e)$. The set of Berge copies of G is denoted by $\mathcal{BG}$. The sets $f_1(V(G))$ and $f_2(E(G))$ are called the defining vertices and hyperedges, respectively.

We recall the classical definition of the Turán number of a family of hypergraphs.

Definition 2. The Turán number of a family of r -uniform hypergraphs $\mathcal{F}$, denoted $\text{ex}_r(n, \mathcal{F})$, is the maximum number of hyperedges in an n -vertex, r -uniform, simple-hypergraph which does not contain an isomorphic copy of $\mathcal{H}$, for all $\mathcal{H} \in \mathcal{F}$, as a sub-hypergraph.

The same question may be asked for multi-hypergraphs, we denote the Turán number for multi-hypergraphs by $\text{ex}_r^{\text{multi}}(n, \mathcal{F})$.

Remark 1. If every hypergraph in $\mathcal{F}$ has at least $r+1$ vertices, then $\text{ex}_r^{\text{multi}}(n, \mathcal{F})$ is infinite, since a hypergraph on r vertices and multiple copies of the same hyperedge is $\mathcal{F}$ -free.

The classical theorem of Erdős and Gallai was extended to Berge paths in r -uniform hypergraphs by Győri, Katona and Lemons [7].

Theorem 2 (Győri, Katona, Lemons [7]). Let n, k, r be positive integers and let $\mathcal{H}$ be an r -uniform hypergraph with no Berge path of length k . If $k > r+1 > 3$, we have

$$e(\mathcal{H}) \leq \frac{n}{k} \binom{k}{r}.$$

If $r \geq k > 2$, we have

$$e(\mathcal{H}) \leq \frac{n(k-1)}{r+1}.$$

The remaining case when $k = r+1$ was settled later by Davoodi, Győri, Methuku and Tompkins [1], the Turán number matches the upper bound of Theorem 2 in the $k > r+1$ case.

We now turn our attention to the case of trees in hypergraphs. The Turán number of certain kinds of trees in r -uniform hypergraphs has long been a major topic of research. For example, there is a notoriously difficult conjecture of Kalai [4] which is more general than the Erdős-Sós conjecture. The trees which Kalai considers are generalizations of the notion of tight paths in hypergraphs. In another direction, Füredi [6] investigated linear trees, constructed by adding $r-2$ new vertices to every edge in a (graph) tree. In this setting, he proved asymptotic results for all uniformities at least 4. Whereas, the articles above considered classes of trees containing tight and linear paths, respectively, we will consider the setting of Berge trees.

In the range when $k > r$, a number of results on forbidding Berge trees were obtained by Gerbner, Methuku and Palmer in [5]. In particular they proved that if we assume the Erdős-Sós conjecture holds for a tree T with k edges and all of its sub-trees and also that $k > r+1$, we have $\text{ex}_r(n, \mathcal{BT}) \leq \frac{n}{k} \binom{k}{r}$ (a construction matching this bound when k divides n is given by n/k disjoint copies of the complete r -uniform hypergraph on k vertices). In the present paper, we will consider the range $r > k$, where we prove some exact results.

2 Main Results

Considering multi-hypergraphs, we prove the following.

Theorem 3. *Let n, k, r be positive integers and let T be a k -edge tree, then for all $r \geq (k-1)(k-2)$,*

$$\text{ex}_r^{\text{multi}}(n, \mathcal{BT}) \leq \frac{n(k-1)}{r}.$$

If $r > (k-1)(k-2)$ and T is not a star, equality holds if and only if r divides n and the extremal multi-hypergraph is $\frac{n}{r}$ disjoint hyperedges, each with multiplicity $k-1$. If T is a star equality holds only for all $(k-1)$ -regular multi-hypergraphs.

We conjecture that Theorem 3 holds for the following wider set of parameters.

Conjecture 1. *Let n, k, r be positive integers and let T be a k -edge tree, then for all $r \geq k+1$,*

$$\text{ex}_r^{\text{multi}}(n, \mathcal{BT}) \leq \frac{n(k-1)}{r}.$$

For all trees T , where T is not a star, equality holds if and only if r divides n and the extremal multi-hypergraph is $\frac{n}{r}$ disjoint hyperedges each with multiplicity $k-1$.

The special case of Conjecture 1, when the forbidden tree is a path, was settled by Györi, Lemons, Salia and Zamora [8] (see the first corollary).

We now define a class of hypergraphs which we will need when we classify the extremal examples in our main result about simple hypergraphs, Theorem 4.

Definition 3. *An r -uniform hypergraph $\mathcal{H}$ is two-sided if $V(\mathcal{H})$ can be partitioned into a set X and pairwise disjoint sets A_i , $i = 1, 2, \dots, t$ (also disjoint from X) of size $r-1$, such that every hyperedge is of the form $\{x\} \cup A_i$ for some $x \in X$. We say that a two-sided r -uniform hypergraph is (a, b) -regular if every vertex of X has degree a and every vertex of $\bigcup_{i=1}^t A_i$ has degree b .*

Remark 2. *A two-sided r -uniform hypergraph can also be viewed as a graph obtained by taking a bipartite graph G with bipartite classes X and Y , and “blowing up” each vertex of Y to a set of size $r-1$, and replacing each edge $\{x, y\}$ by the r -hyperedge containing x together with the blown up set for y .*

Theorem 4. *Let n, k, r be positive integers and let T be a k -edge tree which is not a star, then for all $r \geq k(k-2)$,*

$$\text{ex}_r(n, \mathcal{BT}) \leq \frac{n(k-1)}{r+1}.$$

Equality holds if and only if $r+1$ divides n , and the extremal hypergraph is obtained from $\frac{n}{r+1}$ disjoint sets of size $r+1$, each containing $k-1$ hyperedges. Unless k is odd, and T is the balanced double star, where the balanced double star is the tree obtain from an edge by adding $\frac{k-1}{2}$ incident edges to each of the ends of the edge, in which case equality holds if and only if $r+1$ divides n and $\mathcal{H}$ is obtained from the disjoint union of sets of size $r+1$ containing $k-1$ hyperedges each and possibly a $(k-1, \frac{k-1}{2})$ -regular two-sided r -uniform hypergraph (see Figure 1).

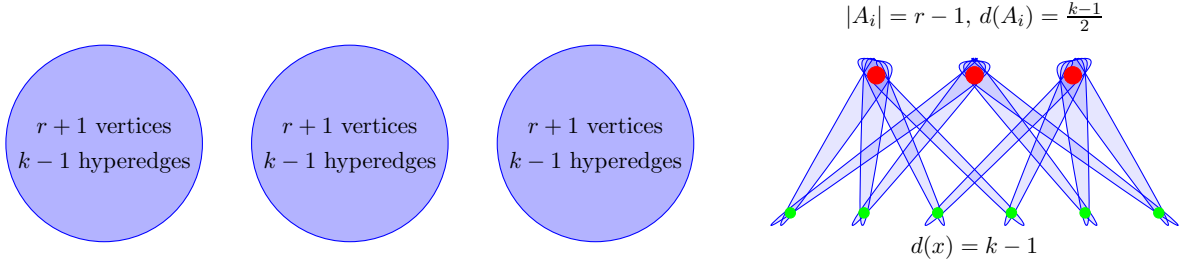


Figure 1: An extremal graph for Theorem 4 is pictured. Any such graph can be obtained from disjoint copies of a sets of $r + 1$ vertices with $k - 1$ hyperedges and if T is the balanced double star, possibly a $(k - 1, \frac{k-1}{2})$ -regular two-sided r -uniform hypergraph.

3 Proofs of the main results

We start with some results about graphs.

Definition 4. For a Graph G , we denote by $d(G)$ the average degree of G , that is $d(G) = \frac{2e(G)}{v(G)}$.

Lemma 1. Any non-empty graph G contains a subgraph G' with minimum degree greater than $d(G)/2$.

The previous lemma is a well-known result in graph theory, which can be proved using the following lemma.

Lemma 2. Let G be a graph and $V' \subseteq V$, if V' is incident with at most $\frac{d(G)}{2} |V'|$ edges, then $d(G[V \setminus V']) \geq d(G)$.

Proof. Note that if m is the number of edges incident with V' , then we have that

$$2e(G[V \setminus V']) = 2e(G) - 2m \geq d(G)v(G) - d(G)|V'| = d(G)(|V| - |V'|) = d(G)v(G[V \setminus V']). \quad \square$$

We are going to use the following fact about trees, before proving the next bound on the degrees of the vertices in clusters.

Claim 1. If T is k -edge tree which is not a star, then there exists a vertex of T which is not a leaf and has degree at most $\frac{k+1}{2}$.

Proof. Let T' be the tree obtained by T by removing every leaf of T , since T is not a star, T' has at least two vertices, take any v, w which are leaves in T' , and note that for each, every neighbor but one is a leaf, and also, since at most one the k edges of T is incident with both u and v , we have that $d_T(u) + d_T(v) \leq k + 1$. And so, one of these vertices have the desired properties. $\square$

Now we introduce two more definitions which we will need in the proofs.

Definition 5. Let $\mathcal{H}$ be a (multi-)hypergraph. A $(k - 1)$ -cluster is a set of $k - 1$ hyperedges of $\mathcal{H}$ that intersect in at least $k - 1$ vertices. The intersection of the $k - 1$ hyperedges is called the core of the $(k - 1)$ -cluster. The union of the $k - 1$ hyperedges is called the span of the $(k - 1)$ -cluster.

Definition 6. Let $\mathcal{H} = (V, E)$ be a multi-hypergraph. A multi-hypergraph $\mathcal{H}' = (V', E')$ is called a reduced sub-hypergraph of $\mathcal{H}$ if $V' \subseteq V$ and there exists an injection $f : E' \rightarrow E$ such that $h \subseteq f(h)$ for all $h \in E'$. For an edge $h \in E'$ we call $f(h) \in E$ its correspondent edge in $\mathcal{H}$.

In the following claims, we bound the degrees of the vertices in a $(k-1)$ -cluster for a hypergraph which does not contain a copy of a Berge tree.

Claim 2. Let n, k, r be positive integers, with $r \geq k+1$, and let T be a k -edge tree. If $\mathcal{H}$ is an r -uniform multi-hypergraph containing no Berge copy of T and S is a $(k-1)$ -cluster in $\mathcal{H}$, then the vertices in the core of S have degree exactly $k-1$. In particular, the core vertices of S are only incident with the hyperedges of S .

Proof. Let C be the set of vertices in the core of S . Suppose, by contradiction, there is a vertex v in C with degree at least k , and let T' be a tree obtained from T by removing any two leaves x, y . Suppose that the neighbors of these leaves are x' and y' respectively (it is possible that $x' = y'$). Since C has at least $k-1$ vertices and there are $k-1$ hyperedges containing all the vertices in C , we can greedily embed T' in C in such a way that v takes the role of x' . Suppose the vertex u takes the role of y' in this greedy embedding. We can complete the embedding of T by using the last hyperedge of S and an unused vertex in it (one exists since $r \geq k+1$) to embed y . Then since the degree of v is at least k , we have a hyperedge available to embed x as a unused vertex of this hyperedge. Thus we have found a Berge copy of T in $\mathcal{H}$, a contradiction. $\square$

Claim 3. Let n, k, r be positive integers, with $r \geq k+1$, and let T be a k -edge tree which is not a star. If $\mathcal{H}$ is an r -uniform multi-hypergraph containing no Berge copy of T and S is a $(k-1)$ -cluster of $\mathcal{H}$, then any vertex in the span of S that is incident with a hyperedge not from S , has degree at most $\lfloor \frac{k+1}{2} \rfloor$.

Proof. Since T is not a star, by Claim 1, there is a vertex $x \in V(T)$ which is not a leaf and has degree s , $s \leq \lfloor \frac{k+1}{2} \rfloor$, such that all but one of its neighbors is a leaf, let y be the neighbor of x which is not a leaf. Suppose, by contradiction, there is a vertex v in the span of S which is incident with a hyperedge that is not in S and v has degree at least $\lfloor \frac{k+1}{2} \rfloor$. Let C be the set of vertices in the core of S . From Claim 2 we know that v cannot be in C . Pick s hyperedges $h_1, h_2, \dots, h_s$ incident to v in such a way that h_1 is not in S and h_2 is in S . Choose a vertex $w \in h_1$ not in C (in fact, every vertex in h_1 is outside C by Claim 2) and $u \in h_2$ in C . Choose further distinct vertices $v_3, v_4, \dots, v_s$ from the hyperedges $h_3, h_4, \dots, h_s$. The vertex v will be assigned to the vertex x in the tree, and the vertex u will be assigned to the vertex y ($v_3, v_4, \dots, v_s$ will be assigned to the leaves adjacent to x). Thus, using the hyperedges $h_1, h_2, \dots, h_s$ we can embed the vertex x and all its neighbors in T using at most $s-1$ hyperedges from S and at most $s-1$ vertices from C (v and w are not in C).

There are at least $(k-1) - (s-1) = k-s$ remaining vertices in C . Each of these is contained in at least $k-s$ unused hyperedges of S . Thus, the remaining $k-s$ vertices of the tree can be mapped to distinct vertices from C , and the remaining edges of the tree may be assigned to distinct unused hyperedges of S . $\square$

Remark 3. Note that by Claim 2 and Claim 3, if $\mathcal{H}$ is a multi-hypergraph with uniformity $r \geq k+1$ that does not contain a Berge copy of a tree on k edges which is not a star, then $(k-1)$ -clusters of $\mathcal{H}$ are edge-disjoint.

Lemma 3. *Let k be a positive integer and let T be a k -edge tree which is not a star. Let $\mathcal{H}$ be a multi-hypergraph not necessarily uniform, not containing a Berge copy of T , and assume that each hyperedge in $\mathcal{H}$ has size at least $k + 1$. If there exists a reduced sub-hypergraph $\mathcal{H}' = (V', E')$ of $\mathcal{H}$ such that $d_{\mathcal{H}'}(v) \geq k - 1$ for each $v \in V'$ and $|h| \geq k - 1$ for each $h \in E'$, then $\mathcal{H}'$ contains a $(k - 1)$ -cluster. Note that if S is a $(k - 1)$ -cluster in $\mathcal{H}'$, then the correspondent edges of S in $\mathcal{H}$ are a $(k - 1)$ -cluster.*

Proof. Let $h_2 \in E'$. We will show that every vertex in h_2 is contained in the same set of hyperedges in E' . Let $v_1, v_2 \in h_2$, and suppose by contradiction that there exists a hyperedge h_3 incident to v_2 and not to v_1 . Enumerate the vertices of T by $x_0, x_1, \dots, x_k$ in such a way that the graph induced by the vertices $x_0, x_1, \dots, x_i$ is connected for all i , x_0 is a leaf of T and x_0, x_1, x_2, x_3 is a path of length 3 (such a path exists since T is not a star). For each $i = 1, 2, \dots, k$, the vertex x_i is adjacent to exactly one vertex of smaller index, call the edge using x_i and the vertex of smaller index e_i .

We can embed T into $\mathcal{H}$ in the following way. First assign v_1 to x_1 , h_2 to $\{x_1, x_2\}$, v_2 to x_2 , h_3 to $\{x_2, x_3\}$ and any vertex in $v_3 \in h_3 \setminus \{v_1, v_2\}$ to x_3 . For $i = 4, \dots, k$, suppose $e_i = \{x_i, x_{j_i}\}$. Pick any hyperedge $h_i \in E'$ incident to v_{j_i} and distinct from $h_2, h_3, \dots, h_{i-1}$ (such hyperedges exist since $d_{\mathcal{H}'}(v_{j_i}) \geq k - 1$) and assign it to e_i . If $i \leq k - 1$, pick any $v_i \in h_i \setminus \{v_1, v_2, \dots, v_{i-1}\}$, and if $i = k$, then let $\tilde{h}_k$ be the correspondent hyperedge of h_k in $\mathcal{H}$. As $\tilde{h}_k$ has size bigger than k , let v_k be any vertex in $\tilde{h}_k \setminus \{v_1, v_2, \dots, v_{k-1}\}$. This vertex v_k is assigned to x_k . Finally, since v_1 is incident with at least $k - 1$ hyperedges distinct to h_3 , there is a hyperedge h_1 incident to v_1 and distinct from the already chosen hyperedges. Let $\tilde{h}_1$ be the correspondent hyperedge of h_1 . Take any vertex in $\tilde{h}_1$ which has not been assigned yet and assign it to x_0 . Thus, by replacing the edge h_i with their correspondent hyperedges, we have found a Berge copy of T in $\mathcal{H}$, a contradiction.

It follows that for any $v_1, v_2 \in h_2$, we have that v_1 and v_2 must be incident with the same set of hyperedges in $\mathcal{H}'$ (by assumption at least $k - 1$), and so $\mathcal{H}'$ contains a $(k - 1)$ -cluster. $\square$

Lemma 3 says that if $\mathcal{H}$ does not contain a Berge copy of a tree and we are able to find a large enough reduced sub-hypergraph, then $\mathcal{H}$ must have at least one $(k - 1)$ -cluster. The main idea of the proofs of the main results is to show that if $\mathcal{H}$ has too many hyperedges and no Berge copy of a tree, then after removing all $(k - 1)$ -clusters, we would still be able to find a large enough reduced sub-hypergraph. This would imply that there is still another $(k - 1)$ -cluster in $\mathcal{H}$, a contradiction.

Proof of Theorem 3. Let T be a k -edge tree, which is not a star. Suppose that $\mathcal{H}$ is an n -vertex r -uniform hypergraph with at least $\frac{n(k-1)}{r}$ hyperedges such that $\mathcal{H}$ does not contain a Berge copy T , and let G be the incidence bipartite graph of $\mathcal{H}$, i.e., the bipartite graph with color classes $V(\mathcal{H})$ and $E(\mathcal{H})$ where $v \in V(\mathcal{H})$ is adjacent to $h \in E(\mathcal{H})$ if and only if $v \in h$.

Since $e(\mathcal{H}) \geq \frac{n(k-1)}{r}$, we have $\frac{e(G)}{v(G)} = \frac{re(\mathcal{H})}{n + e(\mathcal{H})} = \frac{r}{\frac{n}{e(\mathcal{H})} + 1} \geq \frac{r}{\frac{r}{k-1} + 1} = \frac{r(k-1)}{r + k - 1}$, and

note that

$$\frac{r(k-1)}{r + k - 1} \geq k - 2 \Leftrightarrow r(k-1) \geq (k-2)(r + k - 1) = r(k-1) + (k-2)(k-1) - r \Leftrightarrow r \geq (k-2)(k-1).$$

Hence $d(G) = \frac{2e(G)}{v(G)} \geq 2 \left(\frac{r(k-1)}{r + k - 1} \right) \geq k - 2$, since $r \geq (k-2)(k-1)$. Suppose $\mathcal{H}$ has t distinct $(k - 1)$ -clusters $S_1, S_2, \dots, S_t$ (recall that by Remark 3 $(k - 1)$ -clusters are edge-disjoint). For each S_i , let X_i be the set of vertices which are incident only with hyperedges of S_i , let $X = \bigcup_{i=1}^t X_i$ and let Y be the set of vertices that are not in X but are incident with at least one of the $(k - 1)$ -clusters.

Let G_1 be the induced subgraph of G obtained by removing X , Y and all $(k-1)$ -cluster hyperedges from the vertex set of G . We will show that $d(G_1) \geq d(G)$ (provided G_1 is not the empty graph).

The number of edges removed in G is $\sum_{v \in X} d_{\mathcal{H}}(v) + \sum_{v \in Y} d_{\mathcal{H}}(v)$. Since the degree of each $v \in X$ is at most $k-1$, we have that $\left(\sum_{v \in X} d_{\mathcal{H}}(v)\right) \leq |X|(k-1)$. Also X is only incident with the $(k-1)$ -cluster hyperedges, thus we also have the bound $\left(\sum_{v \in X} d_{\mathcal{H}}(v)\right) \leq tr(k-1)$, and since the degree of each $v \in Y$ is at most $\frac{k-1}{2}$ (Claim 3), we have that $\left(\sum_{v \in Y} d_{\mathcal{H}}(v)\right) \leq \frac{(k-1)|Y|}{2}$. Therefore

$$\begin{aligned} & \left(\sum_{v \in X} d_{\mathcal{H}}(v) + \sum_{v \in Y} d_{\mathcal{H}}(v)\right) (r+k-1) \\ &= \left(\sum_{v \in X} d_{\mathcal{H}}(v)\right) r + \left(\sum_{v \in X} d_{\mathcal{H}}(v)\right) (k-1) + \left(\sum_{v \in Y} d_{\mathcal{H}}(v)\right) (r+k-1) \\ &\leq |X| r(k-1) + tr(k-1)^2 + \frac{(k-1)|Y|}{2} (r+k-1) \leq r(k-1)(|X| + t(k-1) + |Y|), \end{aligned}$$

where in the last inequality we used $\frac{r+k-1}{2} < r$. Thus, equality can hold only if $Y = \emptyset$.

Rearranging we have

$$\left(\sum_{v \in X} d_{\mathcal{H}}(v) + \sum_{v \in Y} d_{\mathcal{H}}(v)\right) \leq \frac{r(k-1)}{r+k-1} (|X| + t(k-1) + |Y|). \quad (1)$$

The left-hand side of (1) is the number of removed edges, and the right-hand side is $d(G)/2$ times the number of removed vertices. Therefore, by Lemma 2, if G_1 is non-empty, we have that

$$d(G_1) \geq d(G) \geq 2(k-2).$$

Hence, by Lemma 1 there is a subgraph G_2 of G_1 with minimum degree at least $k-1$. Suppose that G_2 has bipartite classes $A \subseteq V(\mathcal{H})$ and $B \subseteq E(\mathcal{H})$, and define $\mathcal{H}'$ by taking the vertex set $V' = A$ and $E' = \{h \cap V' : h \in B\}$. The condition on the minimum degree of G_2 implies that every vertex of $\mathcal{H}'$ has degree at least $k-1$ and every hyperedge of $\mathcal{H}'$ has size at least $k-1$. Then by Lemma 3, $\mathcal{H}'$ contains a $(k-1)$ -cluster, but this $(k-1)$ -cluster corresponds to a $(k-1)$ -cluster in $\mathcal{H}$ contradicting the fact that we removed every $(k-1)$ -cluster from $\mathcal{H}$. So $\mathcal{H}$ must contain a Berge copy of T , unless G_1 is empty.

Note that, for G_1 to be empty it is necessary that $d(G) = 2\frac{r(k-1)}{r+k-1}$ and that equality holds in the inequality (1). This can be possible only if $Y = \emptyset$ and

$$|X| = \frac{1}{k-1} \sum_{v \in X} d_{\mathcal{H}}(v) = tr.$$

Since every $(k-1)$ -cluster contains at least r vertices, we have $|X_i| \geq r$, and so each X_i must have size exactly r , hence $\mathcal{H}$ is the disjoint union of t hyperedges each with multiplicity $k-1$. Therefore the number of vertices would be a multiple of r and $e(\mathcal{H}) = \frac{n(k-1)}{r}$. Hence if $e(\mathcal{H}) \geq \frac{n(k-1)}{r}$, then $\mathcal{H}$ must contain a Berge copy of T , or $r|n$ and $\mathcal{H}$ is the disjoint union of $\frac{n}{r}$ hyperedges each with multiplicity $k-1$. $\square$

Remark 4. For $r = (k-2)(k-1)$, the proof above also shows that if $e(\mathcal{H}) > \frac{n(k-1)}{r}$, then $\mathcal{H}$ must contain a Berge copy of T . However, the extremal construction does not follow from that proof.

Proof of Theorem 4. Let T be a k -edge tree which is not a star. We may assume $k > 3$, since otherwise T is a path, and we already know the result for paths. Let $\mathcal{H}$ be an n -vertex hypergraph with at least $\frac{n(k-1)}{r+1}$ hyperedges and $r \geq k(k-2)$. We will proceed by induction on the number of vertices n ; the base cases $n \leq r+1$ are trivial.

If there is a set U of size $r+1$ which is incident with at most $k-1$ hyperedges, put $V' = V \setminus U$ and let $n' = |V'| = n - r - 1$. By induction, $\mathcal{H}'$ the hypergraph induced by V' , has at most $\frac{n'(k-1)}{r+1}$ hyperedges and equality holds if $r+1|n'$ and $\mathcal{H}'$ is the disjoint union of cliques, unless T is the balanced double star, then it may contain a $(k-1, \frac{k-1}{2})$ -regular two-sided hypergraph as described in the statement of the theorem. Note that if one of the hyperedges incident with U is incident with a vertex v , $v \in V'$, then v has degree at least $\lfloor \frac{k+1}{2} \rfloor$, and v is in a $(k-1)$ -cluster of $\mathcal{H}'$, thus we have a Berge copy of T from Claim 3. Hence, the $k-1$ hyperedges incident with U are contained in the vertex set U and $\mathcal{H}$ has the desired structure.

Similarly to the proof of Theorem 3, we have that

$$\frac{e(G)}{v(G)} = \frac{re(\mathcal{H})}{n + r(\mathcal{H})} = \frac{r}{\frac{n}{e(\mathcal{H})} + 1} \geq \frac{r}{\frac{r+1}{k-1} + 1} = \frac{r(k-1)}{r+k},$$

and note that

$$\frac{r(k-1)}{r+k} \geq k-2 \Leftrightarrow r(k-1) \geq (k-2)(r+k) = r(k-1) + (k-2)k - r \Leftrightarrow r \geq k(k-2).$$

Hence $d(G) = \frac{2e(G)}{v(G)} \geq 2 \left(\frac{r(k-1)}{r+k-1} \right) \geq k-2$, since $r \geq (k-2)(k-1)$. Suppose that $\mathcal{H}$ has t distinct $(k-1)$ -clusters $S_1, S_2, \dots, S_t$. Define the sets $X_1, \dots, X_t, X$ and Y as in the proof of Theorem 3. We are going to remove all vertices and hyperedges of these $(k-1)$ -clusters as in the previous proof, and we will denote the incidence bipartite graph of $\mathcal{H}$ by G . By G_1 we will denote the incidence bipartite graph of the hypergraph $\mathcal{H}'$, obtained from $\mathcal{H}$ after removing the $(k-1)$ -clusters.

If $|X_i| \geq r+1$ for some i , then by taking $U \subseteq X_i$ of size $r+1$, we would have that U is incident with at most $k-1$ hyperedges, and we would be done by induction. Hence we assume that $|X_i| \leq r$.

For each i , with $|X_i| = r$, we have

$$\sum_{v \in X_i} d_{\mathcal{H}}(v) \leq (r-1)(k-1) + 1 = |X_i|(k-1) - (k-2),$$

since any hyperedge is incident with at most $r-1$ vertices from X_i , with the possible exception of at most one hyperedge (X_i , if $X_i \in E(\mathcal{H})$).

For each i , with $|X_i| \leq r-1$, we have

$$\sum_{v \in X_i} d_{\mathcal{H}}(v) \leq |X_i|(k-1) \leq (r-1)(k-1).$$

Let a be the number of X_i , $1 \leq i \leq t$, with the size r . Then we have the following inequalities

$$\sum_{v \in X} d_{\mathcal{H}}(v) = \sum_{\substack{|X_i|=r \\ v \in X_i}} d_{\mathcal{H}}(v) + \sum_{\substack{|X_i| < r \\ v \in X_i}} d_{\mathcal{H}}(v) \leq t(r-1)(k-1) + a, \quad (2)$$

and

$$\sum_{v \in X} d_{\mathcal{H}}(v) \leq \sum_{\substack{|X_i|=r \\ v \in X_i}} (|X_i|)(k-1) - (k-2)) + \sum_{\substack{|X_i| \leq r \\ v \in X_i}} |X_i| (k-1) = |X| (k-1) - a(k-2). \quad (3)$$

We also have

$$tr(k-1) \leq \sum_{v \in X} d_{\mathcal{H}}(v) + \sum_{v \in Y} d_{\mathcal{H}}(v) \leq t(r-1)(k-1) + a + \frac{k-1}{2} |Y|, \quad (4)$$

where in the first inequality follows from the fact the set the edges of $t(k-1)$ hyperedges of the t cluster are incident only with the set $X \cup Y$, and the second inequality follows directly from Claim 3 together with the fact that $d_{\mathcal{H}}(v) \leq k-1$ by definition.

Rearranging (4) yields

$$t(k-1) \leq a + \frac{|Y|(k-1)}{2}. \quad (5)$$

The following three bounds come from multiplying inequality (3) and (2) by r and k , respectively, and the bound from Claim 3 by $k+r$.

$$\left(\sum_{v \in X} d_{\mathcal{H}}(v) \right) r \leq |X| r(k-1) - ar(k-2). \quad (6)$$

$$\left(\sum_{v \in X} d_{\mathcal{H}}(v) \right) k \leq t(r-1)k(k-1) + ak. \quad (7)$$

$$\left(\sum_{v \in Y} d_{\mathcal{H}}(v) \right) (k+r) \leq \frac{|Y|(k-1)}{2} (k+r). \quad (8)$$

Now we bound the number of deleted hyperedges times $r+k$. From (6), (7), (8) and then (5), it follows that

$$\begin{aligned} & \left(\sum_{v \in X} d_{\mathcal{H}}(v) + \sum_{v \in Y} d_{\mathcal{H}}(v) \right) (k+r) \leq |X| r(k-1) - ar(k-2) + t(r-1)k(k-1) + ak + \frac{|Y|(k-1)}{2} (k+r) \\ & = |X| r(k-1) - ar(k-2) + tr(k-1)^2 + t(k-1)(r-k) + ak + \frac{|Y|(k-1)}{2} (k+r) \\ & \leq |X| r(k-1) - ar(k-2) + tr(k-1)^2 + (r-k) \left(a + \frac{|Y|(k-1)}{2} \right) + ak + \frac{|Y|(k-1)}{2} (k+r) \\ & = |X| r(k-1) - ar(k-3) + tr(k-1)^2 + |Y|(k-1)r = r(k-1)(|X| + |Y| + t(k-1)) - ar(k-3) \\ & \leq r(k-1)(|X| + |Y| + t(k-1)). \end{aligned}$$

Rearranging we have

$$\left(\sum_{v \in X} d_{\mathcal{H}}(v) + \sum_{v \in Y} d_{\mathcal{H}}(v) \right) \leq \frac{r(k-1)}{r+k} (|X| + t(k-1) + |Y|). \quad (9)$$

The left-hand side of (9) is the number of removed edges, and the right-hand side of (9) is $d(G)/2$ times the number of removed vertices.

Hence, by Lemma 2 if G_1 is nonempty, we have that

$$d(G_1) \geq d(G) \geq 2(k-2).$$

Thus, by Lemma 1 we can find a subgraph G_2 of G_1 with minimum degree at least $k-1$. Suppose that G_2 has bipartite classes $A \subseteq V$ and $B \subseteq E(\mathcal{H})$, define $\mathcal{H}'$ by taking the vertex set $V' = A$ and hyperedge set $E' = \{h \cap V' : h \in B\}$. The condition on the minimum degree of G_2 implies that every vertex of $\mathcal{H}$ has minimum degree at least $k-1$, and every hyperedge of $\mathcal{H}'$ has size at least $k-1$. Then by Lemma 3, $\mathcal{H}'$ contains a $(k-1)$ -cluster, which contradicts that we have removed all $(k-1)$ -clusters in $\mathcal{H}$.

For G_1 to be empty it is necessary that $d(G) = 2\frac{r(k-1)}{r+k}$, and for (9) to hold with equality, we must have that $e(\mathcal{H}) = \frac{n(k-1)}{r+1}$. To obtain equality in (9), it is necessary that $a = 0$ (since $k > 3$) and that every hyperedge contains one of the X_i . It then follows that $|X| = t(r-1)$, and by (5), $|Y| = 2t$. By (8), for every $v \in Y$, we have $d_{\mathcal{H}}(v) = \frac{k-1}{2}$, so $n = t(r+1)$. Then $\mathcal{H}$ is a disjoint union of sets of $r+1$ vertices with $k-1$ hyperedges, and a hypergraph constructed from the classes $A = \{X_1, X_2, \dots, X_t\}$ and $B = Y$, where $\{y, X_i\}$ is an edge if $X_i \cup \{y\}$ is a hyperedge of $\mathcal{H}$. Note that $2t = 2|A| = |B|$, the degree of every vertex in B is $\frac{k-1}{2}$ and every vertex of A has degree $k-1$; that is, $\mathcal{H}$ is a $(k-1, \frac{k-1}{2})$ -regular two-sided hypergraph.

However, this is only possible if k is odd, and it is simple to check that this construction contains a Berge copy of every k -edge tree which is not a balanced k -edge double star or the k -edge star. $\square$

Acknowledgments

The research of the first three authors was partially supported by the National Research, Development and Innovation Office NKFIH, grants K116769, K117879 and K126853. The research of the second author is partially supported by Shota Rustaveli National Science Foundation of Georgia SRNSFG, grant number DI-18-118. The research of the third author is supported by the Institute for Basic Science (IBS-R029-C1).

References

- [1] A. Davoodi, E. Győri, A. Methuku and C. Tompkins. An Erdős-Gallai type theorem for uniform hypergraphs. *European Journal of Combinatorics* 69 (2018): 159–162.
- [2] , P. Erdős. On some problems in graph theory, combinatorial analysis and combinatorial number theory, *Graph Theory and Combinatorics (Cambridge, 1983)*, Academic Press, London (1984):1–17.
- [3] P. Erdős and T. Gallai. On maximal paths and circuits of graphs. *Acta Mathematica Hungarica* 10(3–4): (1959): 337–356.
- [4] P. Frankl, Z. Füredi. Exact solution of some Turán-type problems. *Journal of Combinatorial Theory Series A* 45 (1987): 226–262.

- [5] D. Gerbner, A. Methuku and C. Palmer. General lemmas for Berge-Turán hypergraph problems. *arXiv preprint* arXiv:1808.10842 (2018).
- [6] Z. Füredi. Linear trees in uniform hypergraphs. *European Journal of Combinatorics* 35 (2014): 264–272.
- [7] E. Győri, Gy. Katona and N. Lemons. Hypergraph extensions of the Erdős-Gallai theorem. *European Journal of Combinatorics* 58 (2016): 238–246.
- [8] E. Győri, N. Lemons, N. Salia and O. Zamora. The structure of hypergraphs without long Berge cycles. *arXiv preprint* arXiv:1812.10737 (2018).