On the maximum size of connected hypergraphs without a path of given length

Ervin Győri ∗  Abhishek Methuku †  Nika Salia ‡  Casey Tompkins §  Máté Vizer ¶

October 24, 2017

Abstract

In this note we asymptotically determine the maximum number of hyperedges possible in an \( r \)-uniform, connected \( n \)-vertex hypergraph without a Berge path of length \( k \), as \( n \) and \( k \) tend to infinity. We show that, unlike in the graph case, the multiplicative constant is smaller with the assumption of connectivity.

1 Introduction

Let \( P_k \) denote a path consisting of \( k \) edges in a graph \( G \). There are several notions of paths in hypergraphs the most basic of which is due to Berge. A Berge path of length \( k \) is a set of \( k + 1 \) distinct vertices \( v_1, v_2, \ldots, v_{k+1} \) and \( k \) distinct hyperedges \( h_1, h_2, \ldots, h_k \) such that for \( 1 \leq i \leq k \), \( v_i, v_{i+1} \in h_i \). A Berge path is also denoted simply as \( P_k \), and the vertices \( v_i \) are called basic vertices. If \( v_1 = v \) and \( v_{k+1} = w \), then we call the Berge path a Berge \( v-w \)-path.

A hypergraph \( \mathcal{H} \) is called connected if for any \( v \in V(\mathcal{H}) \) and \( w \in V(\mathcal{H}) \) there is a Berge \( v-w \)-path. Let \( N_s(G) \) denote the number of \( s \)-vertex cliques in the graph \( G \).

A classical result of Erdős and Gallai [3] asserts that

\[ |E(G)| \leq \frac{(k-1)n}{2}. \]

\[ \text{Theorem 1 (Erdős-Gallai). Let } G \text{ be a graph on } n \text{ vertices not containing } P_k \text{ as a subgraph, then} \]

\[ |E(G)| \leq \frac{(k-1)n}{2}. \]
In fact, Erdős and Gallai deduced this result as a corollary of the following stronger result about cycles,

**Theorem 2** *(Erdős-Gallai)*. Let $G$ be a graph on $n$ vertices with no cycle of length at least $k$, then

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

Kopylov [5] and later Balister, Győri, Lehel and Schelp [1] determined the maximum number of edges possible in a connected $P_k$-free graph.

**Theorem 3.** Let $G$ be a connected $n$-vertex graph with no $P_k$, $n > k \geq 3$. Then $|E(G)|$ is bounded above by

$$\max\left\{\left(\frac{k-1}{2}\right) + n - k + 1, \left\lceil \frac{k+1}{2} \right\rceil + \left\lfloor \frac{k-1}{2} \right\rfloor (n - \left\lceil \frac{k+1}{2} \right\rceil)\right\}.$$

Observe that, although the upper bound is lower in the connected case, it is nonetheless the same asymptotically. Balister, Győri, Lehel and Schelp also determined the extremal cases.

**Definition 1.** The graph $H_{n,k,a}$ consists of 3 disjoint vertex sets $A, B, C$ with $|A| = a$, $|B| = n - k + a$ and $|C| = k - 2a$. $H_{n,k,a}$ contains all edges in $A \cup C$ and all edges between $A$ and $B$. $B$ is taken to be an independent set. The number of $s$-cliques in this graph is

$$f_s(n, k, a) = \left(\begin{array}{c} k-a \\ s \end{array}\right) + (n - k + a)\left(\begin{array}{c} a \\ s-1 \end{array}\right).$$

The upper bound of Theorem 3 is attained for the graph $H_{n,k,1}$ or $H_{n,k,\left\lfloor \frac{k+1}{2} \right\rfloor}$.

We now mention some recent results of Luo [6] which will be essential in our proof.

**Theorem 4** *(Luo)*. Let $n - 1 \geq k \geq 4$. Let $G$ be a connected $n$-vertex graph with no $P_k$, then the number of $s$-cliques in $G$ is at most

$$\max\{f_s(n, k, \lfloor (k-1)/2 \rfloor), f_s(n, k, 1)\}.$$

As a corollary, she also showed

**Corollary 1** *(Luo)*. Let $n \geq k \geq 3$. Assume that $G$ is an $n$-vertex graph with no cycle of length $k$ or more, then

$$N_s(G) \leq \frac{n-1}{k-2}\left(\begin{array}{c} k-1 \\ s \end{array}\right).$$

Győri, Katona and Lemons [3] initiated the study of Berge $P_k$-free hypergraphs. They proved
Theorem 5 (Győri-Katona-Lemons). Let $\mathcal{H}$ be an $r$-uniform hypergraph with no Berge path of length $k$. If $k > r + 1 > 3$, we have

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

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

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

The case when $k = r + 1$ was settled later [2]:

Theorem 6 (Davoodi-Győri-Methuku-Tompkins). Let $\mathcal{H}$ be an $n$-vertex $r$-uniform hypergraph. If $|E(\mathcal{H})| > n$, then $\mathcal{H}$ contains a Berge path of length at least $r + 1$.

Our main result is the asymptotic upper bound for the connected version of Theorem 5, as $n$ and $k$ tend to infinity.

Theorem 7. Let $\mathcal{H}_{n,k}$ be a largest $r$-uniform connected $n$-vertex hypergraph with no Berge path of length $k$, then

$$\lim_{k \to \infty} \lim_{n \to \infty} \frac{|E(\mathcal{H}_{n,k})|}{k^{r-1}n} = \frac{1}{2^{r-1}(r-1)!}.$$  

A construction yielding the bound in Theorem 7 is given by partitioning an $n$-vertex set into two classes $A$, of size $\left\lfloor \frac{k-1}{2} \right\rfloor$, and $B$, of size $n - \left\lfloor \frac{k-1}{2} \right\rfloor$ and taking $X \cup \{y\}$ as a hyperedge for every $(r-1)$-element subset $X$ of $A$ and every element $y \in B$. This hypergraph has no Berge $P_k$ as we could have at most $\left\lfloor \frac{k-1}{2} \right\rfloor$ basic vertices in $A$ and $\left\lfloor \frac{k-1}{2} \right\rfloor + 1$ basic vertices in $B$, thus yielding less than the required $k + 1$ basic vertices.

Observe that in Theorem 5 the corresponding limiting value of the constant factor is $\frac{1}{r!}$ which is $\frac{2^{r-1}}{r!}$ times larger than in the connected case. Note that the ideas of the proof of Theorem 7 can be used to prove that the limiting value of the constant factor in Theorem 5 is $\frac{1}{r!}$.

2 Proof of Theorem 7

We will use the following simple corollary of Theorem 4

Corollary 2. Let $G$ be a connected graph on $n$ vertices with no $P_k$, then $G$ has at most

$$\frac{k^{r-1}n}{2^{r-1}(r-1)!}$$  

$r$-cliques if $n \geq c_{k,r}$ for some constant $c_{k,r}$ depending only on $k$ and $r$.

Proof. From Theorem 4 it follows that for large enough $n$, the number of $r$-cliques is at most

$$\left(n - \left\lfloor \frac{k-1}{2} \right\rfloor \right) \binom{k-1}{2} r - 1) + \binom{k-1}{r-2} < n \binom{k}{2} (r - 1).$$  

3
Given an $r$-uniform hypergraph $\mathcal{H}$ we define the shadow graph of $\mathcal{H}$, denoted $\partial \mathcal{H}$ to be the graph on the same vertex set with edge set:

$$E(\partial \mathcal{H}) := \{\{x, y\} : \{x, y\} \subseteq E(\mathcal{H})\}.$$ 

**Definition 2.** If $r = 3$, then we call an edge $e \in E(\partial \mathcal{H})$ fat if there are at least 2 distinct hyperedges $h_1, h_2$ with $e \subseteq h_1, h_2$. If $r > 3$, then we call an edge $e \in E(\partial \mathcal{H})$ fat if there are at least $k$ distinct hyperedges $h_1, h_2, \ldots, h_k$ in $\mathcal{H}$ with $e \subseteq h_i$ for $1 \leq i \leq k$.

We call an edge $e \in E(\partial \mathcal{H})$ thin if it is not fat.

Thus, the set $E(\partial \mathcal{H})$ decomposes into the set of fat edges and the set of thin edges. We will refer to the graph whose edges consist of all fat edges in $\partial \mathcal{H}$ as the fat graph and denote it by $F$.

**Lemma 1.** There is no $P_k$ in the fat graph $F$ of the hypergraph $\mathcal{H}$.

**Proof.** Suppose we have such a $P_k$ with edges $e_1, e_2, \ldots, e_k$. For $r = 3$, if a hyperedge contains two edges from the path, then it must contain consecutive edges $e_i, e_{i+1}$. Select hyperedges $h_1, h_2, \ldots, h_k$ where $e_i \subseteq h_i$ in such a way that $h_{i+1}$ is different from $h_i$ for all $1 \leq i \leq k-1$, and these edges yield the required Berge path.

Suppose now that $r > 3$, we will find a Berge path of length $k$ in $\mathcal{H}$, greedily. For $e_1$, select an arbitrary hyperedge $h_1$ containing it. Suppose we have found a distinct hyperedge $h_i$ containing the fat edge $e_i$ for all $1 \leq i < i^*$. Since the edge $e_i$ is fat, there are at least $k$ different hyperedges $h^1_i, h^2_i, \ldots, h^k_i$ containing it. Select one of them, say $h^1_i$, which is not equal to any of $h_1, h_2, \ldots, h_{i-1}$. Thus, we may find distinct hyperedges $h_1, h_2, \ldots, h_k$ where $e_i \subseteq h_i$ for $1 \leq i \leq k$, and thus, we have a Berge path of length $k$.

We call a hyperedge $h \in E(\mathcal{H})$ fat if $h$ contains no thin edge. Let $\mathcal{F}$ denote the hypergraph on the same set of vertices as $\mathcal{H}$ consisting of the fat hyperedges, then

**Lemma 2.** If $r = 3$, then

$$|E(\mathcal{H} \setminus \mathcal{F})| \leq \frac{(k-1)n}{2}.$$ 

If $r > 3$, then

$$|E(\mathcal{H} \setminus \mathcal{F})| \leq \frac{(k-1)^2n}{2}.$$ 

**Proof.** Arbitrarily select a thin edge from each $h \in \mathcal{H} \setminus \mathcal{F}$. Let $G$ be the graph consisting of the selected thin edges. We know that each edge in $G$ was selected at most once if $r = 3$ and at most $k-1$ times in the $r > 3$. Thus, we have that $|\mathcal{H} \setminus \mathcal{F}| \leq |E(G)|$ for $r = 3$ and $|\mathcal{H} \setminus \mathcal{F}| \leq (k-1)E(G)$ for $r > 3$. Moreover, $G$ is $P_k$-free since a $P_k$ in $G$ would imply a Berge $P_k$ in $\mathcal{H}$ by considering any hyperedge from which each edge was selected. It follows by Theorem [1] that $|E(G)| \leq \frac{(k-1)n}{2}$, so $|\mathcal{H} \setminus \mathcal{F}| \leq \frac{(k-1)n}{2}$ if $r = 3$, and $|\mathcal{H} \setminus \mathcal{F}| \leq \frac{(k-1)^2n}{2}$ if $r > 3$. 

[1]
Any hyperedge of $F$ contains only fat edges, so it corresponds to a unique $r$-clique in $F$. This implies the following.

**Observation 1.** *The number of hyperedges in $E(F)$ is at most the number of $r$-cliques in the fat graph $F$.***

To this end we will upper bound the number of $r$-cliques in $F$, by making use of the following important lemma.

**Lemma 3.** *There are no two disjoint cycles of length at least $k/2 + 1$ in the fat graph $F$.***

*Proof.* Let $C$ and $D$ be two such cycles. By connectivity, there are vertices $v \in V(C)$ and $w \in V(D)$ and a Berge path from $v$ to $w$ in $H$ containing no additional vertices of $C$ or $D$ as defining vertices. This path can be extended using the hyperedges containing the edges of $C$ and $D$ to produce a Berge path of length $k$ in $H$ (note that here we used that the edges of $C$ and $D$ are fat), a contradiction. 

Assume that $F$ has connected components $C_1, C_2, \ldots, C_t$. Trivially,

$$N_r(F) = \sum_{i=1}^{t} N_r(C_i). \quad (1)$$

If $|V(C_i)| \leq k/2$, then trivially

$$N_r(C_i) \leq \left( \frac{|V(C_i)|}{r} \right) \leq \frac{|V(C_i)|^r}{r!} \leq \frac{k^{r-1} |V(C_i)|}{2^{r-1}(r-1)!}.$$ 

So we can assume $|V(C_i)| \geq k/2$. By Lemma 3 we have that for all but at most one $i$, $C_i$ does not contain a cycle of length at least $k/2 + 1$. So by Corollary 1, for all but at most one $i$, say $i_0$, we have

$$N_r(C_i) \leq \frac{|V(C_i)| - 1}{k/2 - 2} \left( \frac{k/2 - 1}{r} \right) \leq \frac{k^{r-1} |V(C_i)|}{2^{r-1}(r-1)!} + O(k^{r-2}).$$

If $|V(C_{i_0})| \geq c_{k,r}$, then by Lemma 1 and by Corollary 2 we have

$$N_r(C_{i_0}) \leq \frac{k^{r-1} |V(C_{i_0})|}{2^{r-1}(r-1)!}.$$ 

Otherwise, $N_r(C_{i_0}) \leq \left( \frac{|V(C_{i_0})|}{r} \right) = o(n)$. Therefore, by (1), we have

$$N_r(F) = \sum_{i=1}^{t} N_r(C_i) \leq \sum_{i=1}^{t} \left( \frac{k^{r-1} |V(C_i)|}{2^{r-1}(r-1)!} + O(k^{r-2}) \right) + o(n) \leq \frac{k^{r-1} n}{2^{r-1}(r-1)!} + O(k^{r-2}) n + o(n).$$
Therefore, by Observation 1,

\[ |E(F)| \leq N_r(F) \leq \frac{k^{r-1}n}{2^{r-1}(r - 1)!} + O(k^{r-2})n + o(n). \tag{2} \]

Since \( |E(H)| = |E(H \setminus F)| + |E(F)| \), adding up the upper bounds in (2) and Lemma 2, we obtain the desired upper bound on \( |E(H)| \).

\qed

Acknowledgements

The authors Győri, Methuku, Salia and Tompkins were supported by the National Research, Development and Innovation Office – NKFIH under the grant K116769. The author Vizer was supported by the Hungarian National Research, Development and Innovation Office – NKFIH under the grant SNN 116095.

References

[1] P. N. Balister, E. Győri, J. Lehel, and R. H. Schelp. Connected graphs without long paths. Discrete Mathematics, 308(19):4487–4494, 2008.

[2] A. Davoodi, E. Győri, A. Methuku, and C. Tompkins. An Erdős-Gallai type theorem for hypergraphs. arXiv preprint arXiv:1608.03241, 2016.

[3] P. Erdős and T. Gallai. On maximal paths and circuits of graphs. Acta Mathematica Hungarica, 10(3-4):337–356, 1959.

[4] E. Győri, Gy. Y. Katona, and N. Lemons. Hypergraph extensions of the Erdős-Gallai theorem. European Journal of Combinatorics, 58:238–246, 2016.

[5] G. N. Kopylov. Maximal paths and cycles in a graph. DOKLADY AKADEMII NAUK SSSR, 234(1):19–21, 1977.

[6] R. Luo. The maximum number of cliques in graphs without long cycles. arXiv preprint arXiv:1701.07472, 2017.