

An exact result for hypergraphs and upper bounds for the Turán density of K_{r+1}^r

Linyuan Lu *

Department of Mathematics
University of South Carolina
Columbia, SC 29208

Yi Zhao †

Department of Mathematics and Statistics
Georgia State University
Atlanta, GA 30303

February 21, 2009

Abstract

We first answer a question of de Caen [2]: given $r \geq 3$, if G is an r -uniform hypergraph on n vertices such that every $r + 1$ vertices span 1 or $r + 1$ edges, then $G = K_n^r$ or K_{n-1}^r , assuming that $n > (p - 1)r$, where p is the smallest prime factor of $r - 1$. We then show that the Turán density

$$\pi(K_{r+1}^r) \leq 1 - \frac{1}{r} - \left(1 - \frac{1}{r^{p-1}}\right) \frac{(r-1)^2}{2r^p \left(\binom{r+p}{p-1} + \binom{r+1}{2}\right)},$$

for all even $r \geq 4$, improving a well-known bound $1 - \frac{1}{r}$ of de Caen [1] and Sidorenko [15].

1. Introduction

Given a positive integer n and an r -uniform hypergraph H (or r -graph, for short), the Turán number $\text{ex}(n, H)$ is the maximum number of edges in an r -graph on n vertices that does not contain H as a subgraph (called H -free). It is easy to see that $f(n, H) = \text{ex}(n, H) / \binom{n}{r}$ is an decreasing function of n (Katona et al. [10]). The limit $\pi(H) = \lim_{n \rightarrow \infty} f(n, H)$, which always exists, is called the Turán density of H . Let K_k^r denote the complete r -graph on k vertices. Turán determined $\text{ex}(n, K_k^2)$ which implies that $\pi(K_k^2) = 1 - \frac{1}{k-1}$ for all $k \geq 3$. However, no Turán density $\pi(K_k^r)$ is known for any $k > r \geq 3$. The most well-known case $k = 4$ and $r = 3$ is a conjecture of Turán [20], claiming that $\pi(K_4^3) = 5/9$. Erdős [5] offered \$500 for determining

*Research supported in part by NSF grant DMS-0701111.

†Research supported in part by NSA grants H98230-06-1-0140, H98230-07-1-0019.

any $\pi(K_k^r)$ with $k > r \geq 3$ and \$1000 for answering it for all k and r . The best (general) known upper bound is due to de Caen [1]

$$\pi(K_k^r) \leq 1 - \frac{1}{\binom{k-1}{r-1}}, \quad (1)$$

and the special case

$$\pi(K_{r+1}^r) \leq 1 - \frac{1}{r} \quad (2)$$

was also given by Sidorenko [15]. For the lower bound, Sidorenko [18] showed that $\pi(K_{r+1}^r) \geq 1 - \frac{\ln r}{2r}(1 + o(1))$ for large r . See survey papers of de Caen [2] and Sidorenko [17] for other bounds.

For odd $r \geq 3$, Chung and Lu [4] improved (2) to

$$\pi(K_{r+1}^r) \leq 1 - \frac{5r + 12 - \sqrt{9r^2 + 24r}}{2r(r+3)} = 1 - \frac{1}{r} - \frac{1}{r(r+3)} + O\left(\frac{1}{r^3}\right). \quad (3)$$

When $r = 3$, this gives the best known upper bound $\frac{3+\sqrt{17}}{12} \approx 0.5936$ for $\pi(K_4^3)$.

However, it seems more difficult to improve (2) for even $r \geq 4$. For odd r , the edge density $f(n, K_{r+1}^r)$ of extremal hypergraphs decreases quickly for small n , for example, $f(r+2, K_{r+1}^r) < f(r+1, K_{r+1}^r)$. It is no longer the case for even $r \geq 4$. It is easy to see that

$$f(8, K_5^4) = f(7, K_5^4) = f(6, K_5^4) = f(5, K_5^4) = \frac{4}{5}.$$

Let G_8^4 be the 4-graph on $[8] = \{1, 2, \dots, 8\}$ with *non-edges*

$$\begin{aligned} &\{1, 2, 3, 4\}, \{1, 2, 5, 6\}, \{1, 2, 7, 8\}, \{1, 3, 5, 7\}, \{1, 3, 6, 8\}, \{1, 4, 5, 8\}, \{1, 4, 6, 7\}, \\ &\{2, 3, 5, 8\}, \{2, 3, 6, 7\}, \{2, 4, 5, 7\}, \{2, 4, 6, 8\}, \{3, 4, 5, 6\}, \{3, 4, 7, 8\}, \{5, 6, 7, 8\}. \end{aligned}$$

It is easy to check that G_8^4 is K_5^4 -free with $\frac{4}{5}\binom{8}{4}$ edges, and every 5 vertices on $[8]$ contain exactly 4 edges and one non-edge.

While studying Turán numbers, De Caen [2] asked the following question: *Describe all r -graphs such that every $r+1$ vertices span 1 or $r+1$ edges.* Frankl and Füredi [7] solved a similar problem earlier: they described all 3-graphs such that every 4 vertices contain 0 or 2 edges (the general case when every $r+1$ vertices span 0 or 2 edges is still open). Theorem 1 below answers the question of de Caen. In fact, our underlying hypergraph G is the complement of the r -graph in de Caen's question:

$$\text{every } r+1 \text{ vertices contain 0 or } r \text{ edges of } G. \quad (4)$$

Fix an n -vertex set V . The empty r -graph is the one with no edge; a complete star is an r -graph whose edge set consists of all r -sets containing some fixed vertex $x \in V$. Clearly the empty graph and complete stars satisfy (4), and their complements are K_n^r and K_{n-1}^r .

Theorem 1. *Let $r \geq 2$. Suppose that G is an r -graph on n vertices satisfying (4).*

1. *If $r = 2$, then G is a complete bipartite graph.*
2. *If $r \geq 3$ and $n > r(p - 1)$, then G is either the empty graph or a complete star, where p is the smallest prime factor of $r - 1$.*

The proof for the $r = 2$ case is rather straightforward; the $r \geq 3$ case requires more work, *e.g.*, we use the upper bound for $\text{ex}(n, K_{r+1}^r)$ corresponding to (2). Note that our construction G_8^4 suggests that a lower bound for n is necessary when $r \geq 3$.

Let K_4^{3-} denote the unique 3-graph with 4 vertices and 3 edges. It is not hard to prove that $\text{ex}(n, K_4^{3-}) \leq \frac{1}{3}$ (*e.g.*, de Caen [1]). Mubayi [13] used the result of Frankl and Füredi [7] and supersaturation to obtain that $\text{ex}(n, K_4^{3-}) < \frac{1}{3} - 10^{-6}$. Talbot [19] recently improved the bound to $\frac{1}{3} - \frac{1}{280}$ by considering related *Chromatic Turán* problems. Motivated by these works, we apply Lemma 4, the key step in the proof of Theorem 1, to slightly improve (2) for all $r \geq 3$. (We can also improve (2) by Theorem 1 and supersaturation though the result turns out to be weaker than Theorem 2. See the last section for details.)

Theorem 2. *For $r \geq 3$, let p be the smallest prime factor of $r - 1$.*

$$\pi(K_{r+1}^r) \leq \begin{cases} 1 - \frac{1}{r} - \frac{8}{9r^2} + O\left(\frac{1}{r^3}\right), & p = 2, \\ 1 - \frac{1}{r} - \left(1 - \frac{1}{r^{p-1}}\right) \frac{(r-1)^2}{2r^p \binom{r+p}{p-1} \binom{r+1}{2}}, & p \geq 3. \end{cases}$$

When r is odd (thus $p = 2$), Theorem 2 is slightly weaker than (3). For even r , Theorem 2 gives a new upper bound for $\pi(K_{r+1}^r)$. For example, when $r = 6k + 4$ (thus $p = 3$), it gives

$$\pi(K_{r+1}^r) \leq 1 - \frac{1}{r} - \frac{1}{2r^3} + O\left(\frac{1}{r^4}\right).$$

The rest of the paper is organized as follows. After recalling the proof of (2) in Section 2, we prove Theorem 1 in Section 3 and Theorem 2 in Section 4. In the last section we give some concluding remarks.

2. Preliminaries

In this section we recall the proof of (2), $\pi(K_{r+1}^r) \leq 1 - \frac{1}{r}$, which indicates the source of our improvement. Let G be a K_{r+1}^r -free r -graph. For $0 \leq i \leq r$, let Δ_i denote the family of $(r + 1)$ -sets that contain exactly i edges of G and $\delta_i = |\Delta_i|$.

Lemma 3. *For all $n \geq r$,*

$$\text{ex}(n, K_{r+1}^r) \leq \left(1 - \frac{1}{r} + \frac{1}{r(n - r + 1)}\right) \binom{n}{r}. \quad (5)$$

Moreover, if every K_{r+1}^r -free r -graph G on n vertices satisfies $\sum_{i=1}^{r-1} \delta_i \geq y \binom{n}{r+1}$, then $x = \pi(K_{r+1}^r)$ satisfies

$$x^2 - x \left(1 - \frac{1}{r}\right) + \frac{(r-1)y}{r(r+1)} \leq 0 \quad (6)$$

Proof. Let G be a K_{r+1}^r -free r -graph G on n vertices and e edges. For an $(r-1)$ -set $S \subset V(G)$, let $d_S = |\{v : S \cup \{v\} \in E(G)\}|$ be the degree of S . We have

$$\sum_S d_S = er \quad (7)$$

$$\sum_S \binom{d_S}{2} = \sum_{i=1}^r \binom{i}{2} \delta_i \quad (8)$$

$$\sum_{i=1}^r i \delta_i = e(n-r). \quad (9)$$

We derive that

$$\begin{aligned} \sum_S d_S^2 &= 2 \sum_S \binom{d_S}{2} + \sum_S d_S \\ &= 2 \sum_{i=1}^r \binom{i}{2} \delta_i + er \quad \text{using (7) and (8)} \\ &= \sum_i i^2 \delta_i - \sum_i i \delta_i + er \\ &= \sum_i i^2 \delta_i - e(n-2r) \quad \text{using (9)} \\ &\leq r \sum_{i=1}^r i \delta_i - e(n-2r) \\ &= re(n-r) - e(n-2r). \end{aligned}$$

Using the Cauchy-Schwarz inequality

$$\frac{(er)^2}{\binom{n}{r-1}} = \frac{(\sum_S d_S)^2}{\binom{n}{r-1}} \leq \sum_S d_S^2$$

and (7), we conclude that

$$\begin{aligned} \frac{(er)^2}{\binom{n}{r-1}} &\leq re(n-r) - e(n-2r) \\ \frac{e}{\binom{n}{r}} &\leq \left(1 - \frac{1}{r}\right) + \frac{1}{r(n-r+1)} \end{aligned}$$

with equality holds if all d_S are the same and $\sum_{i=1}^{r-1} \delta_i = 0$. This proves (5).

When $\sum_{i=1}^{r-1} \delta_i \geq y \binom{n}{r+1}$, we can refine the estimate of $\sum d_S^2$ as

$$\begin{aligned}
\sum_S d_S^2 &= \sum_{i=1}^r i^2 \delta_i - e(n-2r) \\
&= r \sum_{i=1}^r i \delta_i - \sum_{i=1}^{r-1} i(r-i) \delta_i - e(n-2r) \\
&\leq r \sum_{i=1}^r i \delta_i - (r-1) \sum_{i=1}^{r-1} \delta_i - e(n-2r) \\
&\leq re(n-r) - (r-1)y \binom{n}{r+1} - e(n-2r). \tag{10}
\end{aligned}$$

Consequently $\frac{(er)^2}{\binom{n}{r-1}} \leq re(n-r) - (r-1)y \binom{n}{r+1} - e(n-2r)$, and (6) follows by letting $e = x \binom{n}{r}$. \square

3. Proof of Theorem 1

In this section we prove Theorem 1. We need some notations. Let $G = (V, E)$ be an r -graph and T be an $(r-1)$ -subset of V . The neighborhood N_T is the set of $x \in V$ such that $\{x\} \cup T \in E$; and let $\bar{N}_T = V - T - N_T$. We define the *degree* $d_T = |N_T|$ and the *non-degree* $\bar{d}_T = |\bar{N}_T|$. The following lemma provides information on the structures of all r -graphs satisfying (4).

Lemma 4. *Let G be an r -graph with $r \geq 2$ such that every $r+1$ vertices contain 0 or r edges. The following are true for every $(r-1)$ -vertex set T .*

1. $T \cup \bar{N}_T$ contains no edge of G .
2. For every vertex $x \in N_T$ and every $(r-1)$ -subset $D \subseteq T \cup \bar{N}_T$, $\{x\} \cup D$ is an edge of G .
3. If $r \geq 3$, then either $d_T < 2$ or $\bar{d}_T < p-1$.

Proof. To prove Part 1, we need to show that for $i = 1, \dots, r$, every i -subset of \bar{N}_T and every $(r-i)$ -subset of T together form a non-edge. We prove this claim by induction on i . The definition of \bar{N}_T justifies the claim for $i = 1$. Suppose the claim holds for some $i \geq 1$. Consider $A \subseteq \bar{N}_T$ and $B \subseteq T$ with $|A| = i+1$ and $|B| = r-i-1$. Pick an $(r-i)$ -set C such that $B \subset C \subseteq T$. We know that the $r+1$ set $A \cup C$ contains either 0 or r edges, and by induction hypothesis, none of the $i+1 (\geq 2)$ i -subsets of A together with C form an edge. Therefore $A \cup C$ contains no edge, in particular, $A \cup B$ is not an edge.

To prove Part 2, we need to show that for $i = 0, \dots, r-1$, every i -subset of \bar{N}_T , every $(r-1-i)$ -subset of T and x together form an edge. We again apply induction on i . The fact that $x \in N_T$ justifies the $i = 0$ case. Suppose that $A \subseteq \bar{N}_T$ and $B \subseteq T$ with $|A| = i+1$ and $|B| = r-i-2$ for some $i \geq 0$. Pick a set C such that $B \subset C \subseteq T$ with $|C| = r-i-1$. We know that the

$(r + 1)$ -set $A \cup C \cup \{x\}$ contains either 0 or r edges. The induction hypothesis says that all the i -subsets of A form edges together with $C \cup \{x\}$. On the other hand, we know from Part 1 that $A \cup C$ is not an edge. Therefore $A \cup C$ is the unique non-edge in $A \cup C \cup \{x\}$, which implies that $A \cup B \cup \{x\}$ is an edge.

To prove Part 3, we assume instead that some $(r - 1)$ -set T satisfies $d_T \geq 2$ and $\bar{d}_T \geq p - 1$. Let x, y be two vertices in N_T , and S be an $(r + p - 2)$ -subset of $T \cup \bar{N}_T$. By Part 2, every $(r - 1)$ -subset of S together with x or y form an edge. This forces that every $(r - 1)$ -subset of S together with x, y contain precisely r edges. Define an auxiliary $(r - 2)$ -graph H on S in which an $(r - 2)$ -subset R of S is an edge if and only if $R \cup \{x, y\}$ is a non-edge of G (here we need $r \geq 3$). For $n \geq k > t$, we call a t -graph on n vertices an $[n, k, t]$ -system if every k vertices contain exactly one edge. Hence H is an $[r + p - 2, r - 1, r - 2]$ -system. The number of edges of H , $\binom{r+p-2}{r-1}/p$ is therefore an integer. This implies that p divides $\binom{r+p-2}{r-1} = \binom{r+p-2}{p-1}$, or $p!$ divides $(r + p - 2) \cdots r$. Recall that p divides $r - 1$ or $r \equiv 1 \pmod{p}$. We thus obtain a contradiction because $(r + p - 2) \cdots r \equiv (p - 1)! \not\equiv 0 \pmod{p}$. \square

Proof of Theorem 1. When $r = 2$, we apply Lemma 4 Parts 1 and 2 to any $T = \{x\}$ and learn that $A = \{x\} \cup \bar{N}_{\{x\}}$ is an independent set and $ab \in E(G)$ for any $a \in A$, $b \notin A$. We thus know that $B = V - A$ is independent from the assumption (4). Therefore G is a complete bipartite graph with partition sets A and B .

Now let $r \geq 3$. By Lemma 4 Part 3, every $(r - 1)$ -subset $T \subset V(G)$ satisfies $d_T \leq 1$ or $\bar{d}_T \leq p - 2$. We claim that if $n > (p - 1)r$, then there always exists an $(r - 1)$ -vertex set T_0 with $d_{T_0} \leq 1$. In fact, suppose that $\bar{d}_T \leq p - 2$ for all T . Then the number of non-edges is at most

$$\frac{\binom{n}{r-1}(p-2)}{r} = \frac{p-2}{n-r+1} \binom{n}{r} < \left(\frac{1}{r} - \frac{1}{r(n-r+1)} \right) \binom{n}{r}$$

because $n > (p - 1)r$. In other words, G contains more than $\left(1 - \frac{1}{r} + \frac{1}{r(n-r+1)}\right) \binom{n}{r}$ edges. By Lemma 3, there exists an $(r + 1)$ -set which contains $r + 1$ edges, contradicting (4).

Let $U = T_0 \cup \bar{N}_{T_0}$. Then $|U| \geq n - 1$. Lemma 4 Part 1 says that U contains no edges of G . If $|U| = n$, G is the empty graph. Otherwise, $|U| = n - 1$. Let $\{x\} = V(G) - U$. Lemma 4 Part 2 implies that G is the complete star with center x . \square

4. Proof of Theorem 2

Suppose that $G = (V, E)$ is a K_{r+1}^r -free r -graph. We call an $(r + 1)$ -set $S \subseteq V$ **bad** if S contains i edges of G for some $1 \leq i \leq r - 1$. Let \mathcal{B} be the family of all bad sets. Let \mathcal{Q} consist of all $(r + p)$ -subsets $Q \subseteq V$ such that some $(r - 1)$ -subset S of Q has exactly two neighbors in Q (i.e., $|N_S \cap Q| = 2$). Then in every induced sub-hypergraph $G[Q]$, $Q \in \mathcal{Q}$, we have $d_S = 2$ and $\bar{d}_S = p - 1$. Lemma 4 Part 3 thus implies that every $Q \in \mathcal{Q}$ contains some $B \in \mathcal{B}$. For an arbitrary $(r + p)$ -set $Q \subseteq V$, we define $f(Q)$ as the number of $(r - 1)$ -sets $S \subset Q$ such that

$|N_S \cap Q| = 2$, and $w(Q)$ as the number of bad sets in Q . Then $w(Q) \geq 1$ for all $Q \in \mathcal{Q}$. Let $R(G) = \max_{Q \in \mathcal{Q}} f(Q)/w(Q)$. The heart of the proof of Theorem 2 is the lemma below.

Lemma 5. *For $r \geq 3$, let p be the smallest prime factor of $r - 1$ and*

$$R \geq \max\{r - 1, R(G) : G \text{ is a } K_{r+1}^r\text{-free } r\text{-graph}\}. \quad (11)$$

Then the Turán density $x = \pi(K_{r+1}^r)$ satisfies

$$x + \frac{r-1}{2R}x(1-x)^{p-1} \leq 1 - \frac{1}{r}. \quad (12)$$

The proof of Lemma 3 used the Cauchy-Schwarz inequality, or the convexity of x^2 . In order to prove Lemma 5, we instead use the convexity of $f(x) = x^2 + cx^2(1-x)^{p-1}$ with $0 < c \leq 1/2$. Since $f(x) > x^2$, this convexity helps us to get a tighter bound for $\pi(K_{r+1}^r)$.

Proposition 6. *Let c, p be constants satisfying $0 \leq c \leq 1/2$ and $p \geq 2$. Then the function $f(x) = x^2 + cx^2(1-x)^{p-1}$ is convex on $[0, 1]$.*

Proof. First let $p = 2$. From $f(x) = x^2 + cx^2(1-x)$, we derive that $f''(x) = 2 + 2c - 6cx$. Since $1 - 3x \geq -2$ and $0 \leq c \leq 1/2$, we have $f''(x) \geq 0$ on $[0, 1]$.

Now assume that $p \geq 3$. We have

$$f'(x) = 2x + 2cx(1-x)^{p-1} - c(p-1)x^2(1-x)^{p-2}.$$

and

$$\begin{aligned} f''(x) &= 2 + 2c(1-x)^{p-1} - 4c(p-1)x(1-x)^{p-2} + c(p-1)(p-2)x^2(1-x)^{p-3} \\ &\geq 2 + 2c(1-x)^{p-1} - 4c(p-1)x(1-x)^{p-2} \\ &= 2 + 2cg(x), \end{aligned}$$

where $g(x) = (1-x)^{p-1} - 2(p-1)x(1-x)^{p-2} = (1-x)^{p-2}(1 - (2p-1)x)$. It is not hard to see that $g(x) \geq -2\left(\frac{2p-4}{2p-1}\right)^{p-2} > -2$ on $[0, 1]$. Thus $f''(x) > 2 + 2c(-2) \geq 0$ and consequently $f(x)$ is convex on $[0, 1]$. \square

Proof of Lemma 5. Let $G = G(n)$ be a K_{r+1}^r -free r -graph on n vertices with $x(n) \binom{n}{r} = \text{ex}(n, K_{r+1}^r)$ edges. The fact that $w(Q) \geq 1$ for all $Q \in \mathcal{Q}$ and the definition of R give

$$\begin{aligned} \sum_{|S|=r-1} \binom{d_S}{2} \binom{\bar{d}_S}{p-1} &= \sum_S \sum_{Q \supset S: |Q|=r+p, |N_S \cap Q|=2} 1 = \sum_{S, Q} \sum_{B \in \mathcal{B}: B \subset Q} \frac{1}{w(Q)} \\ &= \sum_{B \in \mathcal{B}} \sum_{Q \in \mathcal{Q}: B \subset Q} \frac{f(Q)}{w(Q)} \leq R |\mathcal{B}| \binom{n-r+1}{p-1}. \end{aligned}$$

We thus obtain a bound for the number of bad sets:

$$|\mathcal{B}| = y \binom{n}{r+1} \geq \frac{\sum_S \binom{d_S}{2} \binom{\bar{d}_S}{p-1}}{\binom{n-r+1}{p-1} R}. \quad (13)$$

We define a function $f(x) = x^2 + cx^2(1-x)^{p-1}$ with $c = \frac{r-1}{2R}$. The definition of R forces $c \leq \frac{1}{2}$. By Proposition 6, $f(x)$ is convex on $[0, 1]$. For every $(r-1)$ -set S , let $x_S = \frac{d_S}{n-r+1}$. The convexity of f gives

$$f\left(\frac{\sum_S x_S}{\binom{n}{r-1}}\right) \leq \frac{\sum_S f(x_S)}{\binom{n}{r-1}}.$$

Since $\frac{\sum_S x_S}{\binom{n}{r-1}} = \frac{\sum_S d_S}{(n-r+1)\binom{n}{r-1}} = \frac{e}{\binom{n}{r}} = x(n)$, we have $f\left(\frac{\sum_S x_S}{\binom{n}{r-1}}\right) = f(x(n))$.

On the other hand,

$$\frac{\sum_S f(x_S)}{\binom{n}{r-1}} = \frac{\sum_S (x_S^2 + cx_S^2(1-x_S)^{p-1})}{\binom{n}{r-1}} = \frac{\sum_S d_S^2}{(n-r+1)^2 \binom{n}{r-1}} + c \frac{\sum_S d_S^2 \bar{d}_S^{p-1}}{(n-r+1)^{p+1} \binom{n}{r-1}}.$$

Using (10),

$$\begin{aligned} \frac{\sum_S d_S^2}{(n-r+1)^2 \binom{n}{r-1}} &\leq \frac{re(n-r) - (r-1)y \binom{n}{r+1} - e(n-2r)}{(n-r+1)^2 \binom{n}{r-1}} \\ &= \frac{re(n-r) - e(n-2r)}{(n-r+1)^2 \binom{n}{r-1}} - \frac{r-1}{(r+1)r} y + O\left(\frac{1}{n}\right). \end{aligned}$$

Since $d_S, \bar{d}_S \leq n$ for all S , we have $\sum_S d_S^2 \bar{d}_S^{p-1} = O(n^{p+r-1}) + \sum_S 2(p-1)! \binom{d_S}{2} \binom{\bar{d}_S}{p-1}$ and

$$\begin{aligned} c \frac{\sum_S d_S^2 \bar{d}_S^{p-1}}{(n-r+1)^{p+1} \binom{n}{r-1}} &= \frac{O(n^{p+r-1}) + \sum_S 2c(p-1)! \binom{d_S}{2} \binom{\bar{d}_S}{p-1}}{(n-r+1)^{p+1} \binom{n}{r-1}} \\ &\leq 2c(p-1)! \frac{R \binom{n-r+1}{p-1} y \binom{n}{r+1}}{(n-r+1)^{p+1} \binom{n}{r-1}} + O\left(\frac{1}{n}\right) \\ &= 2c \frac{R}{(r+1)r} y + O\left(\frac{1}{n}\right) \\ &= \frac{r-1}{(r+1)r} y + O\left(\frac{1}{n}\right). \end{aligned}$$

Putting these together, we have

$$\begin{aligned} f(x(n)) &\leq \frac{\sum_S f(x_S)}{\binom{n}{r-1}} \leq \frac{re(n-r) - e(n-2r)}{(n-r+1)^2 \binom{n}{r-1}} + O\left(\frac{1}{n}\right) \\ &= \left(1 - \frac{1}{r}\right) x(n) + O\left(\frac{1}{n}\right). \end{aligned}$$

The claim (12) follows by letting $n \rightarrow \infty$ and substituting $f(x) = x^2 + \frac{r-1}{2R} x^2(1-x)^{p-1}$. \square

Trivially we can choose $R = \binom{r+p}{r-1}$ – applying it in (12) already yields a better bound for $\pi(K_{r+1}^r)$ than (2). The following technical lemma refines our bound for R .

Lemma 7. For any K_{r+1}^r -free r -graph G with at least $r + p$ vertices,

$$R(G) \leq \begin{cases} \left(\frac{3r+2}{4}\right)^2, & p = 2, \\ \binom{r+1}{2} + \binom{r+p}{p-1}, & p \geq 3. \end{cases}$$

Proof. Fix $Q \in \mathcal{Q}$. Let $q = r + p = |Q|$. We consider a p -graph H on Q whose edges are the complements of edges of G , namely,

$$E(H) = \{Q - e : e \in E(G), e \subset Q\}.$$

Therefore a $(p - 1)$ -subset $S \subset Q$ has degree i in H if and only if the $(r + 1)$ -set $Q \setminus S$ contains exactly i edges of G . Let W be the family of $(p - 1)$ -sets with degree i (in H) for some $1 \leq i \leq r - 1$ (their complementary sets in Q are bad sets for G), and F be the set of $(p + 1)$ -sets containing exactly 2 edges of H (their complementary sets in Q have degree 2 in $G[Q]$). Following the definition of $w(Q)$ and $f(Q)$ in the beginning of the section, we have $|W| = w(Q)$ and $|F| = f(Q)$. Note that the degree of any $(p - 1)$ -set in H is at most r since G is K_{r+1}^r -free. We also know that $|W| = w(Q) \geq 1$ since Q contains at least one bad set. From now on we consider H as our underlying hypergraph (instead of G). Our goal is to obtain upper bounds for $|F|/|W|$.

The $p = 2$ case: Partition the vertex set Q into $W = \{x : 1 \leq d_{\{x\}} \leq r - 1\}$, $U = \{x : d_{\{x\}} = r\}$, and $Z = \{x : d_{\{x\}} = 0\}$. Since each vertex in U misses at most one vertex of Q , non-edges inside U form a matching M . Let $U_1 = U - V(M)$ consist of vertices in U not covered by M . The family F consists of all triples $T = abc$ such that $ab, ac \in E(H)$ and $bc \notin E(H)$. Clearly $T \subset W \cup U$ for all $T \in F$.

We partition $F = F_0 \cup F_1 \cup F_2$ such that F_0 consists of all $T \in F$ with $|T \cap W| \geq 2$, F_1 consists of those $T = abc$ with $b, c \in U$, and F_2 consists of those abc with exactly one of b, c in W . We have $|F_0| \leq \binom{|W|}{2}(q - 2)$, and $|F_1| \leq |M|(q - 2)$. Since $|M| \leq (q - |U_1| - |W|)/2$ and $|W| \geq 1$,

$$\begin{aligned} |F_0| + |F_1| &\leq \binom{|W|}{2}(q - 2) + \frac{1}{2}(q - |U_1| - |W|)(q - 2) \\ &\leq \frac{1}{2}|W|(q - 2) \left(\frac{1}{2}(|W| - 1) + (q - |U_1| - |W|)\right) \\ &= \frac{1}{2}|W|(q - 2)(q - |U_1| - 1). \end{aligned} \tag{14}$$

To estimate F_2 , let us consider a vertex $x \in W$. By definition of M , x is adjacent to all the vertices in $V(M)$. If x has t non-neighbors in U , then $t \leq |U_1|$ and the number of triples of F_2 containing x is at most

$$t(|U| - t) \leq \begin{cases} |U_1|(|U| - |U_1|) & \text{if } |U_1| \leq |U|/2, \\ |U|^2/4 & \text{if } |U_1| > |U|/2. \end{cases}$$

First assume that $|U_1| \leq |U|/2$. Since $|U| \leq q - 1$, we have

$$|F_2| \leq |W||U_1|(|U| - |U_1|) \leq |W||U_1|(p - 1 - |U_1|).$$

Together with (14) and using $|W| \geq 1$, we derive that

$$\begin{aligned} |F| &\leq |W| \left(\frac{1}{2}(q-2)(q-|U_1|-1) + |U_1|(q-1-|U_1|) \right) \\ &\leq |W| \left(\frac{3q}{4} - 1 \right)^2, \end{aligned}$$

where equality holds when $|U_1| = q/4$. In contrast, when $|U_1| > |U|/2$, we have

$$\begin{aligned} |F| &\leq |W| \left(\frac{1}{2}(q-2)(q-|U_1|-1) + \frac{1}{4}|U|^2 \right) \\ &\leq |W| \left(\frac{1}{2}(q-2)(q-\frac{1}{2}|U|-1) + \frac{1}{4}|U|^2 \right) \\ &\leq \frac{1}{4}|W|(q-1)(2q-3), \end{aligned}$$

where equality holds when $|U| = q-1$ and $|U_1| = |U|/2$. It is easy to check that $(\frac{3q}{4}-1)^2 \geq \frac{1}{4}(q-1)(2q-3)$ for any q . We therefore conclude that $|F|/|W| \leq \left(\frac{3q}{4}-1\right)^2 = \left(\frac{3r+2}{4}\right)^2$.

The $p \geq 3$ case: Partition F into $F_1 = \{S \in F : S \supset T \text{ for some } T \in W\}$ and $F_2 = F - F_1$. Clearly $|F_1| \leq |W| \binom{q-p+1}{2}$. We prove below that $|F_2| \leq \sqrt{|W|} \binom{q}{p-1}$.

Fix a $(p+1)$ -set $S \in F_2$. By the definition of F_2 , all $(p-1)$ -subsets of S either have degree 0 or degree r in H (equivalently, non-degree 1), and S contains exactly two edges (p -sets) of H . We denote these two edges by $T \cup \{x\}$ and $T \cup \{y\}$, where $T \subset S$ is a $(p-1)$ -set. We have two observations.

Observation 1. *For any $a \in T$, the $(p-1)$ -set $T \cup \{x\} - \{a\}$ must have degree r since $T \cup \{x\} \in E(H)$ and $S - \{a\} \notin E(H)$. This implies that $T \cup \{x, u\} - \{a\} \in E(H)$ for any vertex $u \notin S$.*

Observation 2. *For two distinct vertices $a, b \in T$, the $(p-1)$ -set $S - \{a, b\}$ must have degree 0 since two p -sets $S - \{a\}, S - \{b\}$ are not edges. Note that we need $|T| \geq 2$ or $p \geq 3$ here.*

We define a function $g : F_2 \rightarrow \binom{Q}{p-1}$ such that $g(S) = T$. If g is a one-to-one function, then $|F_2| \leq \binom{q}{p-1} \leq \sqrt{|W|} \binom{q}{p-1}$. Otherwise, let T be a $(p-1)$ -set with the maximum size of $g^{-1}(T) = \{S \in F_2 : g(S) = T\}$. Let M consist of all 2-sets $S - T$ such that $S \in g^{-1}(T)$. Then $|M| = |g^{-1}(T)| \geq 2$.

We first claim that M is a matching, namely, any two pairs in M are disjoint. Suppose instead $xy, xz \in M$. For any $a \in T$, the $(p-1)$ -set $\{x\} \cup T - \{a\}$ is contained in two non-edges $\{x, y\} \cup T - \{a\}$ and $\{x, z\} \cup T - \{a\}$, contradicting Observation 1. We next claim that $\{x, y\} \cup T - \{a, b\} \in W$ for two pairs $P, Q \in M$, any vertices $x \in P, y \in Q$, and any two vertices $a, b \in T$. In fact, Observation 2 says that $T_P = P \cup T - \{a, b\}$ has degree 0, and consequently $\{y\} \cup T_P$ is a non-edge. Similarly $\{x\} \cup T_Q$ is also a non-edge with $T_Q = Q \cup T - \{a, b\}$. On the other hand, we know that $\{x, y\} \cup T - \{a\}$ is an edge from Observation 1. Hence $\{x, y\} \cup T - \{a, b\}$ has degree at least 1 and non-degree at least 2 and thus it belongs to W . This implies that $|W| \geq 2^2 \binom{|M|}{2} \binom{p-1}{2} \geq |M|^2$ (since $|M| \geq 2$). We thus conclude that $\sum_{T \in \mathcal{R}_2} |g^{-1}(T)| \leq \sqrt{|W|} \binom{q}{p-1}$.

Consequently $|F| \leq |W| \binom{q-p+1}{2} + \sqrt{|W|} \binom{q}{p-1}$ and $|F|/|W| \leq \binom{q-p+1}{2} + \binom{q}{p-1}$. \square

Proof of Theorem 2: Let R be the upper bound for $R(G)$ in Lemma 7:

$$R = \begin{cases} \left(\frac{3r+2}{4}\right)^2, & p = 2, \\ \binom{r+1}{2} + \binom{r+p}{p-1}, & p \geq 3. \end{cases}$$

It is easy to check that $R \geq r - 1$ and consequently $c = \frac{r-1}{2R} \leq \frac{1}{2}$. Let the Turán density $x = \pi(n, K_{r+1}^r) = 1 - \frac{1}{r} - \varepsilon$. The inequality (12) from Lemma 5 gives

$$\varepsilon \geq c \left(1 - \frac{1}{r} - \varepsilon\right) \left(\frac{1}{r} + \varepsilon\right)^{p-1} > c \left(1 - \frac{1}{r} - \varepsilon\right) \left(\frac{1}{r}\right)^{p-1},$$

which implies that

$$\begin{aligned} \varepsilon &> \left(1 - \frac{1}{r}\right) \frac{c}{c + r^{p-1}} \\ &> \left(1 - \frac{1}{r}\right) \frac{c}{r^{p-1}} \left(1 - \frac{c}{r^{p-1}}\right) \end{aligned}$$

We now apply the definition of R to $c = \frac{r-1}{2R}$. When $p = 2$, this gives $\varepsilon > \frac{8}{9r^2} + O(1/r^3)$. When $p \geq 3$, we have

$$\varepsilon > \left(1 - \frac{1}{r}\right) \frac{r-1}{2r^{p-1} \left(\binom{r+p}{p-1} + \binom{r+1}{2}\right)} \left(1 - \frac{c}{r^{p-1}}\right) > \frac{(r-1)^2}{2r^p \left(\binom{r+p}{p-1} + \binom{r+1}{2}\right)} \left(1 - \frac{1}{r^{p-1}}\right). \quad \square$$

5. Concluding Remarks

- Following the approach of Mubayi [13], we can improve (2) by Theorem 1 and supersaturation as follows. By Theorem 1, every K_{r+1}^r -free r -graph on $(p-1)r+1$ vertices either contains a bad set or has at most $\binom{(p-1)r}{r-1}$ edges. In fact, from a simple removing argument and the structure of the r -graphs that satisfy (4), we can show that every K_{r+1}^r -free r -graph on $n_0 = (p-1)(r+1)$ vertices either contains $p-1$ bad sets or has at most $\binom{n_0-1}{r-1}$ edges. By averaging and (5), every K_{r+1}^r -free r -graph with $n \geq n_0$ vertices and $x \binom{n}{r}$ edges contains at least

$$\frac{x - \frac{r}{n_0}}{1 - \frac{1}{r} + \frac{1}{r(n_0-r+1)} - \frac{r}{n_0} \frac{\binom{n_0}{r+1}}{\binom{n_0-1}{r+1}}} \frac{p-1}{r} \binom{n}{r+1}$$

bad sets – this technique is usually called *supersaturation* [6]. By (6) in Lemma 3, as $r \rightarrow \infty$,

$$\pi(K_{r+1}^r) \leq 1 - \frac{1}{r} - (1 + o(1)) \frac{p-2}{\binom{(p-1)(r+1)}{r+1} r(p-1)},$$

which is weaker than Theorem 2. This suggests that (13) gives a better estimate on the number of bad sets than the one from supersaturation.

- The bound $n > r(p - 1)$ in Theorem 1 is tight for odd r and $r = 4$ but we do not know if it is tight for any even number $r > 4$. The bound for R in Lemma 7 is near tight for $p = 2$, tight up to a constant factor for $p = 3$, but we do not know if it is tight for $p > 3$.
- Let us consider K_5^4 . The following bounds are due to Giraud [8, 9].

$$\frac{11}{16} \leq \pi(K_5^4) \leq \frac{3}{4}, \quad (15)$$

Sidorenko showed [15] that $\pi(K_5^4) \leq 0.74912$ and conjectured [17] that the lower bound above gives the correct value. Applying (12) with $R = 13$ (the proof of $R = 13$ is not very hard but not simple either), we derive that $\pi(K_5^4) \leq 0.74439$. Markström [12] recently obtained a much better bound $\pi(K_5^4) \leq 0.73655$ by determining $\text{ex}(n, K_5^4)$ for $n \leq 16$ and applying the simple bound $\pi(K_5^4) \leq \text{ex}(n, K_5^4) / \binom{n}{4}$ ($\text{ex}(n, K_5^4)$ was only known [3] for $n \leq 10$ before).

- For any r -graph F with f edges, Sidorenko [16] used an analytic approach to obtain

$$\pi(F) \leq 1 - \frac{1}{f-1}. \quad (16)$$

When $F = K_{r+1}^r$, $f = r + 1$ and this gives (2). An r -graph is called a forest if there is an ordering of its edges E_1, \dots, E_t such that for every $i = 2, \dots, t$, there exists $j_i < i$ such that $(E_i \cap \cup_{\ell=1}^{i-1} E_\ell) \subseteq E_{j_i}$. Keevash [11] found that the method of Sidorenko actually implies that $\pi = \pi(F)$ satisfies

$$\pi^{t-1} + (f-t)(\pi-1) \leq 0, \quad (17)$$

where t is the size of a largest forest contained in F . When $t = 2$, (17) reduces to (16); when $t > 2$, (17) improves (16). In our case (17) does not improve the $1 - 1/r$ bound for $\pi(K_{r+1}^r)$ because the largest forest in K_{r+1}^r has only 2 edges.

Acknowledgments

The authors thanks referees for their comments that improved the presentation.

References

- [1] D. de Caen, Extension of a theorem of Moon and Moser on complete subgraphs. *Ars Combin.* **16** (1983), 5–10.
- [2] D. de Caen, The current status of Turán’s problem on hypergraphs. *Extremal problems for finite sets* (Visegrád, 1991), 187–197, Bolyai Soc. Math. Stud., 3, János Bolyai Math. Soc., Budapest, 1994.

- [3] D. de Caen, D. L. Kreher, S.P. Radziszowski, and W.H. Mills, On the coverings t -sets with $t + 1$ -sets: $C(9, 5, 4)$ and $C(10, 6, 5)$. *Discrete Math.* **92** (1991), 65–77.
- [4] F. Chung, L. Lu, An upper bound for the Turán number $t_3(n, 4)$. *J. Combin. Theory Ser. A* **87** (1999), no. 2, 381–389.
- [5] P. Erdős, On the combinatorial problems which I would like to see solved, *Combinatorica* **1** (1981), 25–42.
- [6] P. Erdős, M. Simonovits, Supersaturated graphs and hypergraphs. *Combinatorica* **3** (1983), no. 2, 181–192.
- [7] P. Frankl, Z. Füredi, An exact result for 3-graphs, *Discrete Math.* **50** (1984), 323–328.
- [8] G. Giraud, Majoration du nombre de Ramsey ternaire-bicolore en $(4, 4)$. *C. R. Acad. Sci. Paris Sér. A-B* **269** (1969) A620–A622.
- [9] G. Giraud, Remarques sur deux problèmes extrémaux. (French) [Remarks on two extremal problems] *Discrete Math.* **84** (1990), no. 3, 319–321.
- [10] Gy Katona, T. Nemetz, M. Simonovits, On a problem of Turán in the theory of graphs. *Mat. Lapok* **15** (1964) 228–238.
- [11] P. Keevash, The Turán problem for hypergraphs on fixed size. *Electron. J. Combin.* **12** (2005), Note 11, 6 pp. (electronic).
- [12] K. Markström, Extremal graphs and bounds for the Turan density of the 4-uniform \mathbf{K}_5 , manuscript.
- [13] D. Mubayi, On hypergraphs with every four points spanning at most two triples, *Electron. J. Combin.* **10** (2003), no. 1, Note 10, 4 pp. (electronic).
- [14] A. Sidorenko, On Turán numbers $T(n, 5, 4)$ and numbers of monochromatic 4-cliques in 2-colored 3-graphs (in Russian). *Voprosi Kibernetiki* **64** (1980), 117–124.
- [15] A. Sidorenko, The method of quadratic forms in a combinatorial problem of Turán. (in Russian) *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* **76** (1982), no. 1, 3–6.
- [16] A. Sidorenko, An analytic approach to extremal problems for graphs and hypergraphs. *Extremal problems for finite sets* (Visegrád, 1991), 423–455, Bolyai Soc. Math. Stud., 3, János Bolyai Math. Soc., Budapest, 1994.
- [17] A. Sidorenko, What we know and what we do not know about Turán numbers. *Graphs Combin.* **11** (1995), no. 2, 179–199.
- [18] A. Sidorenko, Upper bounds for Turán numbers, *J. Combin. Theory Ser. A* **77** (1997), no. 1, 134–147.

- [19] J. Talbot, Chromatic Turán problems and a new upper bound for the Turán density of \mathcal{K}_4^- , *Europ. J. Combin.* **28** (2007), no. 8, 2125–2142.
- [20] P. Turán, On an extremal problem in graph theory. *Mat. Fiz. Lapok* **48** (1941), 436–452.