The maximal length of a gap between $r$-graph Turán densities

Oleg Pikhurko*
Mathematics Institute and DIMAP
University of Warwick
Coventry CV4 7AL, UK
0.Pikhurko@warwick.ac.uk

Submitted: Apr 7, 2015; Accepted: Oct 5, 2015; Published: Oct 30, 2015
Mathematics Subject Classifications: 05D05

Abstract

The Turán density $\pi(F)$ of a family $F$ of $r$-graphs is the limit as $n \to \infty$ of the maximum edge density of an $F$-free $r$-graph on $n$ vertices. Erdős [Israel J. Math 2 (1964):183–190] proved that no Turán density can lie in the open interval $(0,r!/r^r)$. Here we show that any other open subinterval of $[0,1]$ avoiding Turán densities has strictly smaller length. In particular, this implies a conjecture of Grosu [arXiv:1403.4653, 2014].

1 Introduction

Let $F$ be a (possibly infinite) family of $r$-graphs (that is, $r$-uniform set systems). We call elements of $F$ forbidden. An $r$-graph $G$ is $F$-free if no member $F \in F$ is a subgraph of $G$, that is, we cannot obtain $F$ by deleting some vertices and edges from $G$. The Turán function $\text{ex}(n,F)$ is the maximum number of edges that an $F$-free $r$-graph on $n$ vertices can have. This is one of the central questions of extremal combinatorics that goes back to the fundamental paper of Turán [16]. We refer the reader to the surveys of the Turán function by Füredi [8], Keevash [12], and Sidorenko [15].

As was observed by Katona, Nemetz, and Simonovits [11], the limit

$$\pi(F) := \lim_{n \to \infty} \frac{\text{ex}(n,F)}{\binom{n}{r}}$$

exists. It is called the Turán density of $F$. Let $\Pi_{\infty}^{(r)}$ consist of all possible Turán densities of $r$-graph families and let $\Pi_{\text{fin}}^{(r)}$ be the set of all possible Turán densities when finitely
many $r$-graphs are forbidden. It is convenient to allow empty forbidden families, so that 1 is also a Turán density. Clearly, $\Pi_\text{fin}^{(r)} \subseteq \Pi_\infty^{(r)}$. A result of Brown and Simonovits [3, Theorem 1] implies that the topological closure $\text{cl}(\Pi_\text{fin}^{(r)})$ of $\Pi_\text{fin}^{(r)}$ contains $\Pi_\infty^{(r)}$ while the converse inclusion was established in [14, Proposition 1]; thus

$$\Pi_\infty^{(r)} = \text{cl}(\Pi_\text{fin}^{(r)}), \quad \text{for every integer } r \geq 2. \quad (1)$$

For $r = 2$, the celebrated Erdős-Stone-Simonovits Theorem [5, 6] determines the Turán density for every family $\mathcal{F}$. In particular, we have

$$\Pi_\text{fin}^{(2)} = \Pi_\infty^{(2)} = \left\{ \frac{m-1}{m} : m = 1, 2, 3, \ldots, \infty \right\}. \quad (2)$$

Unfortunately, the Turán function for hypergraphs (that is, $r$-graphs with $r \geq 3$) is much more difficult to analyse and many problems (even rather basic ones) are wide open.

Fix some $r \geq 2$. A gap is an open interval $(a, b) \subseteq (0, 1)$ that is disjoint from $\Pi_\infty^{(r)}$ (which, by (1), is equivalent to being disjoint from $\Pi_\text{fin}^{(r)}$). Here we consider $g_r$, the maximal possible length of a gap. In other words, $g_r$ is the maximal $g$ such that there is a real $a$ with $(a, a + g) \subseteq (0, 1) \setminus \Pi_\infty^{(r)}$. For example, (2) implies that $g_2 = 1/2$. Erdős [4] proved that $(0, r!/r^r)$ is a gap; in particular, $g_r \geq r!/r^r$. Here we show that this is equality and every other gap has strictly smaller length.

**Theorem 1.** For every $r \geq 3$, we have that $g_r = r!/r^r$ and, furthermore, $(0, r!/r^r)$ is the only gap of length $r!/r^r$ for $r$-graphs.

In particular we obtain the following result that was conjectured by Grosu [9, Conjecture 10].

**Corollary 2.** The union of $r$-graph Turán densities over all $r \geq 2$ is dense in $[0, 1]$, that is, $\text{cl}(\cup_{r=2}^\infty \Pi_\infty^{(r)}) = [0, 1]$. \hfill $\square$

The question whether the set $\Pi_\text{fin}^{(r)}$ is a well-ordered subset of $([0, 1], \prec)$ for $r \geq 3$ was a famous $1000$ problem of Erdős that was answered in the negative by Frankl and Rödl [7]. Despite a number of results that followed [7], very little is known about other gaps in $\Pi_\infty^{(r)}$ for $r \geq 3$. For example, let $g'_r$ be the second largest gap length, that is, the maximum $g \geq 0$ such that $(a, a + g) \subseteq (r!/r^r, 1) \setminus \Pi_\text{fin}^{(r)}$ for some $a$. The computer-generated proof of Barber and Talbot [2] implies that $g'_3 \geq 0.0017$. Klas Markström and Fei Song [13] conjectured that $2/7, 8/27$ is the (unique) second largest gap for $3$-graphs (and, in particular, $g'_3 = 2/189$). However, not for a single $r \geq 4$ is it known, for example, whether $g'_r$ is zero (i.e. whether $\Pi_\text{fin}^{(r)}$ is dense in $[r!/r^r, 1]$).

This paper is organised as follows. In Section 2 we give some definitions and auxiliary results. Theorem 1 is proved in Section 3. We give another proof of Corollary 2 in Section 4. Although the latter proof is not strong enough to prove Theorem 1, its advantage is that it produces explicit elements of $\Pi_\text{fin}^{(r)}$ (as opposed to the implicit values of certain maximisation problems returned by the proof in Section 3). So we include both proofs here, even though the second one is longer.
2 Preliminaries

For $n \in \mathbb{N}$, define $[n] := \{1, \ldots, n\}$. For reals $a \leq b$, let $(a, b)$ and $[a, b]$ be respectively the open and closed interval of reals with endpoints $a$ and $b$. The standard $(m-1)$-dimensional simplex is

$$S_m := \{ x \in \mathbb{R}^m : x_1 + \cdots + x_m = 1, \ \forall i \in [m] \ x_i \geq 0 \}.$$

An $r$-pattern is a collection $P$ of $r$-multisets on $[m]$, for some $m \in \mathbb{N}$. (By an $r$-multiset we mean an unordered collection of $r$ elements with repetitions allowed.) Let $V_1, \ldots, V_m$ be disjoint sets and let $V = V_1 \cup \cdots \cup V_m$. The profile of an $r$-set $X \subseteq V$ (with respect to $V_1, \ldots, V_m$) is the $r$-multiset on $[m]$ that contains $i \in [m]$ with multiplicity $|X \cap V_i|$. For an $r$-multiset $Y$ on $[m]$, let $Y(|V_1, \ldots, V_m|)$ consist of all $r$-subsets of $V$ whose profile is $Y$. We call this $r$-graph the blow-up of $Y$ (with respect to $V_1, \ldots, V_m$) and the $r$-graph

$$P(|V_1, \ldots, V_m|) := \bigcup_{Y \in P} Y(|V_1, \ldots, V_m|)$$

is called the blow-up of $P$. Let the Lagrange polynomial of $P$ be

$$\lambda_P(x_1, \ldots, x_m) := r! \sum_{D \in P} \prod_{i=1}^m \frac{x_i^{D(i)}}{D(i)!} \in \mathbb{R}[x_1, \ldots, x_m],$$

where $D(i)$ denotes the multiplicity of $i$ in $D$. This definition is motivated by the fact that, for every partition $[n] = V_1 \cup \cdots \cup V_m$, we have that

$$|P(|V_1, \ldots, V_m|)| = \lambda_P\left(\frac{|V_1|}{n}, \ldots, \frac{|V_m|}{n}\right) \times \binom{n}{r} + O(n^{r-1}), \quad \text{as } n \rightarrow \infty.$$

For example, if $r = 3$, $m = 3$, and $P$ consists of multisets $\{1,1,2\}$ and $\{1,2,3\}$, then $P(|V_1, \ldots, V_m|)$ contains all triples that have two vertices in $V_1$ and one vertex in $V_2$ plus all triples with exactly one vertex in each part; here $\lambda_P(x_1, x_2, x_3) = 3x_1^2x_2 + 6x_1x_2x_3$.

Let the Lagrangian of $P$ be $\Lambda_P := \max \{ \lambda_P(x) : x \in S_m \}$, the maximum value of the polynomial $\lambda_P$ on the compact set $S_m$. One obvious connection of this parameter to $r$-graph Turán densities is that, if each blow-up of $P$ is $\mathcal{F}$-free, then $\pi(\mathcal{F}) \geq \Lambda_P$. Also, it is not hard to show that $\Lambda_P = \pi(\mathcal{F})$, where $\mathcal{F}$ consists of all $r$-graphs $F$ such that every blow-up of $P$ is $F$-free; thus $\Lambda_P \in \Pi^{(r)}_\infty$. As shown in [14, Theorem 3], we have in fact that

$$\Lambda_P \in \Pi^{(r)}_\infty, \quad \text{for every } r \text{-pattern } P. \quad (3)$$

We will use the special case of Muirhead’s inequality (see e.g. [10, Theorem 45]) which states that, for any $0 \leq i < j \leq k$, we have

$$x^{k+i}y^{k-i} + x^{k-i}y^{k+i} \leq x^{k+j}y^{k-j} + x^{k-j}y^{k+j}, \quad \text{for } x, y \geq 0. \quad (4)$$
3 Proof of Theorem 1

Let \( r \geq 3 \). Fix a sufficiently large integer \( m = m(r) \) so that \( r!\binom{m}{r}/m^r > 1 - r!/r^r \). Consider \( r \)-graphs \( G_0, \ldots, G_{\binom{m}{r}} \) on \([m]\) such that \( G_0 \) has no edges and, for \( i = 1, \ldots, \binom{m}{r} \), the \( r \)-graph \( G_i \) is obtained from \( G_{i-1} \) by adding a new edge. In other words, we enumerate all \( r \)-subsets of \([m]\) as \( R_1, \ldots, R_{\binom{m}{r}} \) and let \( G_i := \{R_1, \ldots, R_i\} \). Let

\[
\lambda_i(x) := \lambda_{G_i}(x) = r! \sum_{D \subseteq G_i} \prod_{j \in D} x_j,
\]

be the Lagrange polynomial of \( G_i \) and \( A_i := A_{G_i} \) be its Lagrangian, where we view \( G_i \) as an \( r \)-pattern. Since \( G_{i-1} \subseteq G_i \), we have that \( A_{i-1} \subseteq A_i \).

We claim that for every \( i \in \left[ \binom{m}{r} \right] \)

\[
A_i - A_{i-1} \leq r!/r^r. \tag{5}
\]

Indeed, pick \( x \in S_m \) with \( A_i = \lambda_i(x) \). Let \( R_i = \{u_1, \ldots, u_r\} \). When we remove the term \( r!x_{u_1} \cdots x_{u_r} \) from \( \lambda_i(x) \), we get the evaluation of \( \lambda_{i-1} \) on \( x \in S_m \). By definition, \( A_{i-1} \geq \lambda_{i-1}(x) \). Also, since \( x_{u_1} + \cdots + x_{u_r} \leq 1 \), we have \( x_{u_1} \cdots x_{u_r} \leq r^{-r} \) by the Geometric-Arithmetic Mean Inequality. Thus we obtain the stated bound:

\[
A_i = \lambda_i(x) = \lambda_{i-1}(x) + r!x_{u_1} \cdots x_{u_r} \leq A_{i-1} + r!/r^r.
\]

Also, we have \( A_{\binom{m}{r}} \geq \lambda_{\binom{m}{r}}(\frac{1}{m}, \ldots, \frac{1}{m}) = r!\binom{m}{r}/m^r > 1 - r!/r^r \). This and (3) imply that \( g_r \leq r!/r^r \) (while the above-mentioned result of Erdős [4] gives the converse inequality). Also, if we have equality in (5), then necessarily \( x_{u_1} = \cdots = x_{u_r} = 1/r \), each other \( x_j \) is zero, and \( A_{i-1} = \lambda_{i-1}(x) = 0 \), implying the uniqueness part of Theorem 1.

4 Alternative proof of Corollary 2

For integers \( r, s \geq 2 \), let \( P_{r,s} \) consist of ordered \( s \)-tuples \((r_1, \ldots, r_s)\) of non-negative integers such that \( r_1 \geq \cdots \geq r_s \) and \( r_1 + \cdots + r_s = r \). This set admits a partial order, where \( x \succ y \) if \( \sum_{i=1}^k x_i \geq \sum_{i=1}^k y_i \) for every \( k \in [s-1] \). For example, the (unique) maximal element is \((r, 0, \ldots, 0)\) and the (unique) minimal element is \((\lfloor r/s \rfloor, \ldots, \lfloor r/s \rfloor)\).

Let \( A \subseteq P_{r,s} \). The set \( A \) is called down-closed if \( y \in A \) whenever \( x \in A \) and \( x \succ y \). Let \( G_A \) consist of all \( r \)-multisets \( X \) on \([s]\) such that the multiplicities of \( X \) satisfy \( \langle X(1), \ldots, X(s) \rangle \in A \), where \( \langle x \rangle \) denotes the non-increasing ordering of a vector \( x \). Also, we use the shorthand \( \lambda_A := \lambda_{G_A} \) and \( A_s := A_{G_A} \).

Lemma 3. Let \( r, s \geq 2 \). If \( A \subseteq P_{r,s} \) is down-closed, then \( A_s = \lambda_A(\frac{1}{s}, \ldots, \frac{1}{s}) \).

Proof. We use induction on \( s \).

First, we prove the base case \( s = 2 \). Let \( k := r/2 \). For \( h \geq 0 \), let \( I_h \) consist of all integer translates of \( k \) whose absolute value is at most \( h \), that is, \( I_h := (\mathbb{Z} + k) \cap [-h, h] \).
Also, let $I^+_h := I_h \cap [0,h]$. (These definitions will allow us to deal with the cases of even and odd $r$ uniformly.) For example, $P_{r,2} = \{(k+i,k-i): i \in I^+_h\}$.

Take a down-closed set $A \subseteq P_{r,2}$. It consists of pairs $(k+i,k-i)$ with $i \in I^+_h$ for some $h \leq k$. Then $G_A$ consists of all $2k$-multisets on $\{1,2\}$ that contain 1 with multiplicity $k+i$ for $i \in I_h$. By the homogeneity of the polynomials involved, the required inequality can be rewritten as

$$
\sum_{i \in I_h} \binom{2k}{k+i} \left( \frac{x+y}{2} \right)^{2k} - \sum_{i \in I_h} \binom{2k}{k+i} x^{k+i} y^{k-i} \geq 0, \quad \text{for } x,y \geq 0. \tag{6}
$$

We will apply the so-called bunching method where we try to write the desired inequality as a positive linear combination of Muirhead’s inequalities (4). If $j \in I_h$, then the coefficient in front of $x^{k+j} y^{k-j}$ in (6) is

$$
2^{-2k} \binom{2k}{k+j} \sum_{i \in I_h} \binom{2k}{k+i} - \binom{2k}{k+j} \leq 0.
$$

If $j \in I_k \setminus I_h$, then the coefficient is $2^{-2k} \binom{2k}{k+j} \sum_{i \in I_h} \binom{2k}{k+i}$ $\geq 0$. Thus, if we group the left-hand side of (6) into terms $x^{k+j} y^{k-j} + x^{k-j} y^{k+j}$, then we get non-positive coefficients for $0 \leq j \leq h$ followed by non-negative coefficients for $j > h$. Also, the total sum of coefficients is zero because (6) becomes equality for $x = y = 1$. Thus we can “bunch” $I_h$-terms with $(I_k \setminus I_h)$-terms and use (4) to derive the desired inequality (6). This proves the case $s = 2$.

Now, let $s \geq 3$ and suppose that we have proved the lemma for $s-1$ (and all $r$). The function $\lambda_A$ is a continuous function on the compact set $S_s$. Let it attain its maximum on some $\mathbf{x} \in S_s$. If there is more than one choice, then choose $\mathbf{x}$ so that $\Delta := \sum_{i \neq j} |x_i - x_j|$ is minimised. Suppose that $\Delta \neq 0$, say $x_1 \neq x_2$. Note that $\lambda_A$ is a homogeneous polynomial of degree $r$, and the coefficient at $x_1^r \cdots x_s^r$ is $\langle r \rangle_{\{r_1,\ldots,r_s\}}$ if the ordering $\langle r \rangle$ of $r$ is in $A$ and 0 otherwise.

Fix $j \in \{0,\ldots,r\}$. If we collect all terms in front of $x_i^j$, we get

$$
\sum_{\langle r,j \rangle \in A \setminus j, r_1,\ldots,r_s-1} \binom{r}{r_1,\ldots,r_s-1} \prod_{i=1}^{s-1} x_i^{r_i} = \binom{r}{j} \lambda_A \setminus j (x_1,\ldots,x_{s-1}),
$$

where $\langle y,j \rangle$ is obtained from $y$ by appending $j$ and ordering the obtained sequence, while $A \setminus j$ consists of those $y \in P_{r-j,s-1}$ such that $\langle y,j \rangle \in A$.

Let us show that $A \setminus j \subseteq P_{r-j,s-1}$ is down-closed. Take arbitrary $\mathbf{z} \in A \setminus j$ and $\mathbf{y} \preceq \mathbf{z}$. We have to show that $\mathbf{y} \in A \setminus j$. Since $A \supseteq \langle z,j \rangle$ is down-closed, it is enough to show that $\langle z,j \rangle \succ \langle y,j \rangle$. We have to compare the sums of the first $i$ terms of $\langle z,j \rangle$ and of $\langle y,j \rangle$ for each $i \in [s-1]$. A problem could arise only if the new entry $j$ was included into these terms for $\langle y,j \rangle$, say as the term number $h \leq i$, but not for $\langle z,j \rangle$. Since $\mathbf{z} \succeq \mathbf{y}$, we have that $\sum_{f=1}^{h-1} z_f \geq \sum_{f=1}^{h-1} y_f$ (and these are also the initial sums for $\langle z,j \rangle$ and $\langle y,j \rangle$).
Furthermore, each of the subsequent $i - (h - 1)$ entries is at least $j$ for $\langle z, j \rangle$ and at most $j$ for $\langle y, j \rangle$. It follows that $\langle z, j \rangle \geq \langle y, j \rangle$. Thus $A \setminus j$ is down-closed, as claimed.

By the induction assumption (and since $\lambda_{A \setminus j}$ is a homogeneous polynomial), we have that $\lambda_{A \setminus j}(x_1, \ldots, x_{s-1}) \leq \lambda_{A \setminus j}(1 - x_s, \ldots, 1 - x_s)$. Thus

$$A_A = \lambda_A(x) = \sum_{j=0}^{r} {r \choose j} \lambda_{A \setminus j}(x_1, \ldots, x_{s-1}) x^j \leq \lambda_A\left(\frac{1 - x_s}{s - 1}, \ldots, \frac{1 - x_s}{s - 1}, x_s\right).$$

Clearly, the sum $\sum_{i=1}^{s-1} |x_s - x_i|$ does not increase if we replace each of $x_1, \ldots, x_{s-1}$ by their arithmetic mean $(1 - x_s)/(s - 1)$. Since $x_1 \neq x_2$, we have found another optimal element of $S$, with strictly smaller $\Delta$, a contradiction. The lemma is proved. \hfill \bbox

Fix some enumeration $\mathcal{P}_{r,r} = \{R_1, \ldots, R_t\}$ such that if $R_i \geq R_j$ then $i \geq j$. For $j \in \{0, \ldots, t\}$, let $A_j := \{R_i : i \in [j]\}$. Thus, for example, $A_0 = \emptyset$ and $A_t = \mathcal{P}_{r,r}$. By (3), $\Pi_{\text{fin}}^{(r)}$ contains all of the following numbers:

$$0 = A_{A_0} \leq A_{A_1} \leq \cdots \leq A_{A_t} = 1.$$ 

Let us show that $\max\{A_{A_i} - A_{A_{i-1}} : i \in [t]\} = o(1)$ as $r \to \infty$. By definition, each $A_j \subseteq \mathcal{P}_{r,r}$ is down-closed. Thus, by Lemma 3 the difference $A_{A_i} - A_{A_{i-1}}$ is the probability that, when $r$ balls are uniformly and independently distributed into $r$ cells, the ordered ball distribution is given by $R_i$. Expose the first $r - m$ balls, where, for example, $m := \lceil \log r \rceil$. Let $k$ be the number of empty cells. Its expected value is $r(1 - 1/r)^{r-m} = (e^{-1} + o(1)) r$. By Azuma’s inequality (see e.g. [1, Theorem 7.2.1]), we have whp (i.e. with probability $1 - o(1)$ as $r \to \infty$) that $k$ is in $I := [r/4, 3r/4]$. Assume that $k \in I$ and expose the remaining $m$ balls. Let $J$ be the number of balls that land inside the $k$ cells that were empty after the first round. The probability that $J = j$ for any particular integer $j \in [m/8, 7m/8]$ is

$$\left(\frac{m}{j}\right) \left(\frac{k}{r}\right)^j \left(\frac{r-k}{r}\right)^{m-j} = (1 + o(1)) \sqrt{\frac{m}{2\pi j(m-j)}} \left(\frac{mk}{jr}\right)^j \left(\frac{(m-k)(m-j)}{(m-j)r}\right)^{m-j} \leq (1 + o(1)) \sqrt{\frac{m}{2\pi j(m-j)}} = o(1),$$

where we used Stirling’s formula and the Arithmetic-Geometric Mean Inequality. On the other hand, we have whp that $m/8 \leq J \leq 7m/8$ (by Azuma’s inequality and our assumption $k \in I$) and that the last $m$ balls all go into different cells (since $m^2 = o(r)$). Once the first $r - m$ balls are exposed and we condition on the event that the last $m$ balls all land into distinct cells, there is at most one value of $J$ for which the final ball distribution is $R_i$. Thus the probability of getting $R_i$ is $o(1)$ uniformly in $i$, as desired.

This finishes the second proof of Corollary 2.

**Acknowledgements**

The author would like to thank Codrut Grosu for helpful comments.
References