Non-Uniform Turán-Type problems

Dhruv Mubayi * **Yi Zhao** †
Department of Mathematics, Statistics, and Computer Science
University of Illinois at Chicago
Chicago, IL 60607

November 24, 2004

Abstract

Given positive integers n, k, t, with $2 \leq k \leq n$, and $t < 2^k$, let $m(n, k, t)$ be the minimum size of a family $\mathcal{F}$ of (nonempty distinct) subsets of $[n]$ such that every k-subset of $[n]$ contains at least t members of $\mathcal{F}$, and every $(k-1)$-subset of $[n]$ contains at most $t-1$ members of $\mathcal{F}$. For fixed k and t, we determine the order of magnitude of $m(n, k, t)$. We also consider related Turán numbers $T_{\geq r}(n, k, t)$ and $T_r(n, k, t)$, where $T_{\geq r}(n, k, t)$ ($T_r(n, k, t)$) denotes the minimum size of a family $\mathcal{F} \subset \binom{[n]}{\geq r}$ ($\mathcal{F} \subset \binom{[n]}{r}$) such that every k-subset of $[n]$ contains at least t members of $\mathcal{F}$. We prove that $T_{\geq r}(n, k, t) = (1 + o(1))T_r(n, k, t)$ for fixed r, k, t with $t \leq \binom{k}{r}$ and $n \to \infty$.

1 Introduction

Given positive integers n, k, t, with $2 \leq k \leq n$ and $t < 2^k$. We call a family $\mathcal{F} \subset 2^{[n]} \setminus \emptyset$ a (k, t)-system if every k-subset of $[n]$ contains at least t sets from $\mathcal{F}$, and every $(k-1)$-subset of $[n]$ contains at most $t-1$ sets from $\mathcal{F}$. Analogously, given integers n, k, t, r, with $1 \leq r \leq k \leq n$ and $0 \leq t < 2^k$, a *Turán-*$_{\geq r}(n, k, t)$*-system* (Turán-$_r(n, k, t)$-system) is a family $\mathcal{F} \subset \binom{[n]}{\geq r}$ ($\mathcal{F} \subset \binom{[n]}{r}$) so that every k-subset of $[n]$ contains at least t members of $\mathcal{F}$. We denote by $m(n, k, t)$ the minimum size of a (k, t)-system, and by $T_{\geq r}(n, k, t)$ ($T_{\geq r}(n, k, t)$) the minimum size of a Turán-$_{\geq r}(n, k, t)$-system (Turán-$_r(n, k, t)$-system).

Computer scientists introduced and studied $m(n, k, t)$ (see [4, 5, 7] for its history and applications). [7] proves that $m(n, k, t) = \Theta(n^{k-1})$ for $1 < t < k$ and $m(n, 3, 2) = \binom{n-1}{2} + 1$, and [4] proves that for fixed k, $m(n, k, 2) = (1 + o(1))T_{k-1}(n, k, 2)$. $T_r(n, k, t)$ (especially $T_r(n, k, 1)$) is well-known and sometimes called the generalized Turán number, though its nonuniform version $T_{\geq r}(n, k, t)$ appears not to have been studied before. Note that when $T_r(n, k, t) = \Omega(n^r)$ (for fixed $t, r < k$),

*Research supported in part by NSF grants DMS-9970325 and DMS-0400812.

†Research supported in part by NSF grant DMS-9983703, a VIGRE Postdoctoral Fellowship at University of Illinois at Chicago.

the asymptotics of $T_r(n,k,t)$ are not known for any $r \geq 3$ and $t \geq 1$ (see [1] for an introduction to this problem, and [3, 6] for surveys in the case $t = 1$).

In this note, we first study $m(n,k,t)$ for all $1 \leq t < 2^k$, determining its order of magnitude for fixed k, t.

Theorem 1. *Let $2 \leq k \leq n$, $2 \leq t < 2^k$, and $\binom{k}{\leq j-1} < t \leq \binom{k}{\leq j}$. Then there exists a constant c, depending only on k and t, such that*

$$c\binom{n}{k-j} \leq m(n,k,t) \leq \binom{n}{k-j} \quad \text{for} \quad t \leq \binom{k-1}{j} \qquad (*)$$

$$\frac{1}{\binom{k}{j}}\binom{n}{j} \leq m(n,k,t) \leq \binom{n}{\leq j} \quad \text{for} \quad t > \binom{k-1}{j}.$$

Remark. When $j = 1$, $(*)$ yields $m(n,k,t) = \Theta(n^{k-1})$ for $2 \leq t \leq k-1$, a result from [7].

We can obtain the exact value of $m(n,k,t)$ for some choices of t. For $t = 1, k$, and $2^k - 1$, it is trivial to see that $m(n,k,t)$ is equal to $\binom{n}{k}$, n, and $\binom{n}{\leq k}$, respectively. We claim that $m(n,k,2^k-2) = \binom{n}{\leq k-1}$. To see this, we call a (k,t)-system $\mathcal{H}$ *minimal* if $\sum_{S\in\mathcal{H}}|S| \leq \sum_{S\in\mathcal{H}'}|S|$ for every (k,t)-system $\mathcal{H}'$. If $\mathcal{F}$ is a minimal (k,t)-system for $t = 2^k - 2 \geq 2^{k-1}$, and $A \in \mathcal{F}$, then $2^A \setminus \emptyset \subset \mathcal{F}$, since replacing A by $B \subset A$ for some $B \notin \mathcal{F}$ creates another (k,t)-system that contradicts the minimality of $\mathcal{F}$. Consequently $\mathcal{F}_k = \emptyset$, because $S \in \mathcal{F}_k$ now implies that $\mathcal{F}_k \setminus S$ is a (k,t)-system. Since $t = 2^k - 2$, we must have $\mathcal{F} = \binom{[n]}{\leq k-1}$.

Before proceeding our upcoming Theorem 3 which relates $T_{\geq r}(n,k,t)$ and $T_r(n,k,t)$, we make the following observation.

Observation 2. *Let $1 \leq r \leq k$ and $0 \leq t < \sum_{i=r}^{k}\binom{k}{i}$. Let j be the unique integer satisfying $\sum_{i=r}^{j-1}\binom{k}{i} \leq t < \sum_{i=r}^{j}\binom{k}{i}$ and let $t_0 = t - \sum_{i=r}^{j-1}\binom{k}{i} \geq 0$. If $\mathcal{F}$ is a Turán-$_j(n,k,t_0)$-system, then $\mathcal{F}' = \bigcup_{i=r}^{j-1}\binom{[n]}{i} \cup \mathcal{F}$ is a Turán-$_{\geq r}(n,k,t)$-system. This implies that $T_{\geq r}(n,k,t) \leq \sum_{i=r}^{j-1}\binom{n}{i} + T_j(n,k,t_0)$.*

Theorem 3. *Let r, k, t, j, t_0 be fixed as in Observation 2.*

1. *If $t_0 = 0$, then $T_{\geq r}(n,k,t) = \sum_{i=r}^{j-1}\binom{n}{i}$.*

2. *If $t_0 \geq 1$, then $T_{\geq r}(n,k,t) = (1+o(1))\left(\sum_{i=r}^{j-1}\binom{n}{i} + T_j(n,k,t_0)\right)$.*

Conjecture 4. *Given r, k, t, j, t_0 as in Observation 2, $T_{\geq r}(n,k,t) = \sum_{i=r}^{j-1}\binom{n}{i} + T_j(n,k,t_0)$.*

Most of our notations are standard: Given a set X and an integer a, let $\binom{X}{a} = \{S \subset X : |S| = a\}$, $\binom{X}{\leq a} = \{S \subset X : 1 \leq |S| \leq a\}$, $\binom{X}{\geq a} = \{S \subset X : |S| \geq a\}$, and $2^X = \{S : S \subset X\}$. For $\mathcal{F} \subset 2^{[n]}$, let $\mathcal{F}_t = \mathcal{F} \cap \binom{[n]}{t}$ and $\overline{\mathcal{F}_t} = \binom{[n]}{t} \setminus \mathcal{F}_t$. Let $\mathcal{F}_{\leq t} = \cup_{i\leq t}\mathcal{F}_i$ and $\mathcal{F}_{\geq t} = \cup_{i\geq t}\mathcal{F}_i$. Write $\mathcal{F}(X)$ for $\mathcal{F} \cap 2^X$. An r-graph on X is a (hyper)graph $\mathcal{F} \subset \binom{X}{r}$.

2 Proofs

Proof of Theorem 1. The Theorem follows easily from the following four statements.

(1) If $\binom{k-1}{j-1} < t \le \binom{k}{j}$, then $m(n,k,t) \le \binom{n}{k-j}$.

(2) If $\binom{k-1}{j} < t \le \binom{k}{\le j}$, then $m(n,k,t) \le \binom{n}{\le j}$.

(3) If $t > \binom{k}{\le j-1}$, then $m(n,k,t) \ge \binom{n}{j}/\binom{k}{j}$.

(4) If $\binom{k}{\le j-1} < t \le \binom{k-1}{j}$, then $m(n,k,t) \ge c\binom{n}{k-j}$, where c depends only on k and j.

The proofs of **(1)** and **(3)** are straightforward, so we only prove **(2)** and **(4)**.

Proof of (2). Consider the smallest $i' \in [1,j]$ and the largest $i \in [1,j]$ such that $1+\sum_{\ell=i'}^{j}\binom{k-1}{\ell} \le t \le \sum_{\ell=i}^{j}\binom{k}{\ell}$. Such i,i' exist since $\binom{k-1}{j} < t \le \binom{k}{\le j}$. We first show that $i' \le i$. This is trivial for $i=j$, so assume that $i<j$. The choice of i implies that

$$t > \sum_{\ell=i+1}^{j}\binom{k}{\ell} = \binom{k}{i+1} + \sum_{\ell=i+2}^{j}\binom{k}{\ell} \ge \left[\binom{k-1}{i}+\binom{k-1}{i+1}\right] + \sum_{\ell=i+2}^{j}\binom{k-1}{\ell}+1.$$

Since this is equal to $\sum_{\ell=i}^{j}\binom{k-1}{\ell}+1$, the choice of i' implies that $i' \le i$. Now let $\mathcal{F} = \cup_{\ell=i}^{j}\binom{[n]}{\ell}$. Every k-set of $[n]$ has $\sum_{\ell=i}^{j}\binom{k}{\ell} \ge t$ members of $\mathcal{F}$; every $(k-1)$-set of $[n]$ has $\sum_{\ell=i}^{j}\binom{k-1}{\ell} \le \sum_{\ell=i'}^{j}\binom{k-1}{\ell} \le t-1$ members of $\mathcal{F}$. Consequently, $m(n,k,t) \le |\mathcal{F}| \le \binom{n}{\le j}$. □

Proof of (4). First, the assumption $\binom{k}{\le j-1} < \binom{k-1}{j}$ implies that $j < k-j$. Let $\mathcal{F}$ be a (k,t)-system. Let $K_{k-1}^{(i)}$ denote the complete i-graph of order $k-1$. Then $K_{k-1}^{(i)} \not\subset \mathcal{F}$ for all $i \in [j, k-1-j]$, otherwise we obtain a $(k-1)$-set which contains $\binom{k-1}{i} \ge \binom{k-1}{j} \ge t$ members of $\mathcal{F}$, a contradiction. Recall that the Ramsey number $R^{(i)}(s,t)$ is the smallest N such that every i-graph on N vertices contains a copy of either $K_s^{(i)}$ or $\overline{K_t^{(i)}}$. By Ramsey's theorem, $R^{(i)}(s,t)$ is finite. Define $m_{k-2j+1}=k$, and $m_\ell = R^{(k-j-\ell)}(k-1, m_{\ell+1})$ recursively for $\ell = k-2j, k-2j-1, \ldots, 2, 1$.

We claim that every m_1-set of $[n]$ contains at least one member of $\mathcal{F}_{\ge k-j}$. Indeed, consider an m_1-set S_1. Because $K_{k-1}^{(k-j-1)} \not\subset \mathcal{F}$, the definition of m_1 implies that there exists a m_2-subset $S_2 \subseteq S_1$ with all of its $(k-j-1)$-subsets absent from $\mathcal{F}$. Repeating this analysis, we find a sequence of subsets $S_3 \supseteq \cdots \supseteq S_{k-2j+1} = S$ of sizes $m_3 > \cdots > m_{k-2j+1} = k$, respectively. The k-set S thus contains no members of $\mathcal{F}$ of size $k-j-1, \ldots, j$. On the other hand, the k-set S must contain at least $t > \binom{k}{\le j-1}$ members of $\mathcal{F}$, thus at least one member of $\mathcal{F}_{\ge j}$. Hence S contains a member of $\mathcal{F}_{\ge k-j}$. By an easy averaging argument, we obtain $|\mathcal{F}| \ge \binom{n}{k-j}/\binom{m_1}{k-j} = c\binom{n}{k-j}$. □

Proof of Theorem 3 Part 1. Let $\mathcal{F} \subset \binom{[n]}{\ge r}$ be a minimal Turán-$_{\ge r}(n,k,t)$-system. We are to show that $|\mathcal{F}| \ge \sum_{i=r}^{j-1}\binom{n}{i}$. Consider $\overline{\mathcal{F}_{<j}} = \bigcup_{i=r}^{j-1}\binom{[n]}{i} \setminus \mathcal{F}$. For every k-set S of $[n]$,

$$\sum_{i=r}^{j-1}\binom{k}{i} = t \le |\mathcal{F}(S)| = |\mathcal{F}_{<j}(S)| + |\mathcal{F}_{\ge j}(S)| = \sum_{i=r}^{j-1}\binom{k}{i} - |\overline{\mathcal{F}_{<j}}(S)| + |\mathcal{F}_{\ge j}(S)|.$$

Therefore $|\overline{\mathcal{F}_{<j}}(S)| \le |\mathcal{F}_{\ge j}(S)|$. Consequently (using $\binom{n-x}{k-x}$ is decreasing in x for $0 \le x \le k$), $|\overline{\mathcal{F}_{<j}}|\binom{n-j}{k-j} < \sum_{S\in\binom{[n]}{k}}|\overline{\mathcal{F}_{<j}}(S)| \le \sum_{S\in\binom{[n]}{k}}|\mathcal{F}_{\ge j}(S)| \le |\mathcal{F}_{\ge j}|\binom{n-j}{k-j}$. Thus $|\overline{\mathcal{F}_{<j}}| \le |\mathcal{F}_{\ge j}|$, and therefore $|\mathcal{F}| = |\mathcal{F}_{<j}| + |\mathcal{F}_{\ge j}| \ge |\mathcal{F}_{<j}| + |\overline{\mathcal{F}_{<j}}| = \sum_{i=r}^{j-1}\binom{n}{i}$. □

The main tool to prove the second part of Theorem 3 is the following well-known fact. For a family $\mathcal{G}$ of r-graphs, the extremal function $\mathrm{ex}(n,\mathcal{G})$ is the maximum number of edges in an r-graph on n vertices that contains no copy of any member of $\mathcal{G}$.

Theorem 5 (Erdős-Simonovits [2]). *For every* $\varepsilon > 0$ *and every family of* r*-graphs* $\mathcal{G}$*, each of whose members has* k *vertices, there exists* $\delta > 0$*, such that every* r*-graph on* n *vertices with at least* $\mathrm{ex}(n, \mathcal{G}) + \varepsilon\binom{n}{r}$ *edges contains at least* $\delta\binom{n}{k}$ *copies of members of* $\mathcal{G}$*.*

Proof of Theorem 3 Part 2. It suffices to show that for every $\varepsilon > 0$, there exists $n_0 = n_0(\varepsilon, k, t) > 0$, such that for all $n \ge n_0$, $T_{\ge r}(n, k, t) \ge (1-\varepsilon)\left(\sum_{i=r}^{j-1}\binom{n}{i} + T_j(n, k, t_0)\right)$. In fact, this follows from the following claims (taking $n_0 = \max\{n_1, n_2\}$):

(a) $T_{\ge r}(n, k, t) \ge T_{\ge j}(n, k, t_0)$,

(b) $T_{\ge j}(n, k, t_0) > (1-\varepsilon/2)T_j(n, k, t_0)$ for $n > n_1$,

(c) $T_j(n, k, t_0) \ge T_j(n, k, 1) \ge \binom{n}{j}/\binom{k}{j} > \frac{2(1-\varepsilon)}{\varepsilon}\sum_{i=r}^{j-1}\binom{n}{i}$, for $n > n_2$.

Since (a) and (c) are easy to see, we only prove (b). Suppose that $\mathcal{F}$ is a Turán-$_{\ge j}(n, k, t_0)$-system. Let $\mathcal{G}$ be the family of all j-graphs on k vertices with more than $\binom{k}{j} - t_0$ edges. Let δ be the output of Theorem 5 for inputs $\varepsilon/[2\binom{k}{j}]$ and $\mathcal{G}$, and choose n_1 so that $\delta\binom{n}{k} > n^{k-1}$ for all $n > n_1$ (note that $n_1 = n_1(\varepsilon, k, t)$). We will show that $|\mathcal{F}_j| > (1-\varepsilon/2)T_j(n, k, t_0)$ for $n > n_1$. Suppose, for contradiction, that $|\mathcal{F}_j| \le (1-\varepsilon/2)T_j(n, k, t_0)$. Since $\mathrm{ex}(n, \mathcal{G}) = \binom{n}{j} - T_j(n, k, t_0)$ and $T_j(n, k, t_0) \ge \binom{n}{j}/\binom{k}{j}$,

$$|\overline{\mathcal{F}_j}| \ge \binom{n}{j} - \left(1 - \frac{\varepsilon}{2}\right)T_j(n, k, t_0) = \mathrm{ex}(n, \mathcal{G}) + \frac{\varepsilon}{2}T_j(n, k, t_0) \ge \mathrm{ex}(n, \mathcal{G}) + \frac{\varepsilon}{2\binom{k}{j}}\binom{n}{j}.$$

By Theorem 5 applied with input $\varepsilon/[2\binom{k}{j}]$, the j-graph with vertex set $[n]$ and edge set $\overline{\mathcal{F}_j}$ contains at least $\delta\binom{n}{k}$ copies of (not necessarily the same) members of $\mathcal{G}$. In other words, there are at least $\delta\binom{n}{k}$ k-sets of $[n]$ that contain fewer than t_0 members of $\mathcal{F}_j$.

Now consider the family of k-sets of $[n]$ which contains at least one member of $\mathcal{F}_i$ for some $i > j$. Denote this by $\mathcal{K}_i$ and let $\mathcal{K} = \cup_{j<i\le k}\mathcal{K}_j$. Since $|\mathcal{K}_i| \le |\mathcal{F}_i|\binom{n-i}{k-i}$ and $|\mathcal{F}| \le \binom{n}{j}$,

$$|\mathcal{K}| = \sum_{j<i\le k}|\mathcal{K}_i| \le \sum_{j<i\le k}|\mathcal{F}_i|\binom{n-i}{k-i} \le \binom{n-j-1}{k-j-1}|\mathcal{F}| \le \binom{n-j-1}{k-j-1}\binom{n}{j} < n^{k-1}.$$

Since $\delta\binom{n}{k} > n^{k-1} > |\mathcal{K}|$ for $n > n_1$, at least one k-set S of $[n]$ contains fewer than t_0 members of $\mathcal{F}_j$ and no member of $\mathcal{F}_i$ for $i > j$. Consequently S contains fewer than t_0 members of $\mathcal{F}$. This contradicts the assumption that $\mathcal{F}$ is a Turán-$_{\ge j}(n, k, t_0)$-system. □

3 Acknowledgments

The authors thank Gy. Turán for introducing them to the parameter $m(n, k, t)$, and the referees for suggestions that improved the presentation.

References

[1] W. G. Brown, P. Erdős, V. T. Sós, Some extremal problems on r-graphs. New directions in the theory of graphs (Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich, 1971), pp. 53–63. Academic Press, New York, 1973

[2] P. Erdős, M. Simonovits, Supersaturated graphs and hypergraphs. Combinatorica 3 (1983), no. 2, 181–192.

[3] Z. Füredi, Turán type problems. Surveys in combinatorics, 1991 (Guildford, 1991), 253–300, London Math. Soc. Lecture Note Ser., 166, Cambridge Univ. Press, Cambridge, 1991.

[4] Z. Füredi, R. H. Sloan, K. Takata, Gy. Turán, On set systems with a threshold property, submitted.

[5] S. Jukna, Computing threshold functions by depth-3 threshold circuits with smaller thresholds of their gates, Inf. Proc. Lett., 56 (1995), 147–150.

[6] A. Sidorenko, What we know and what we do not know about Turán numbers. (English. English summary) Graphs Combin. 11 (1995), no. 2, 179–199.

[7] R. H. Sloan, K. Takata, Gy. Turán, On frequent sets of Boolean matrices, Annals of Mathematics and Artificial Intelligence, 24 (1998) 193–209.