Combinatorics, Probability and Computing (1999) 8, 483–492. Printed in the United Kingdom

The Minimum Size of Saturated Hypergraphs

OLEG PIKHURKO†

Department of Pure Mathematics and Mathematical Statistics,
Cambridge University, Cambridge CB2 1SB, England
(e-mail: O.Pikhurko@dpmms.cam.ac.uk)

Received 9 July 1997; revised 24 May 1999

Let $\mathscr{F}$ be a family of forbidden k-hypergraphs (k-uniform set systems). An $\mathscr{F}$-saturated hypergraph is a maximal k-uniform set system not containing any member of $\mathscr{F}$. As the main result we prove that, for any finite family $\mathscr{F}$, the minimum number of edges of an $\mathscr{F}$-saturated hypergraph is $O(n^{k-1})$. In particular, this implies a conjecture of Tuza. Some other related results are presented.

1. Introduction

A *k-hypergraph* H is, as usual, a pair $(V(H), E(H))$ (*vertices* and *edges*) where

$$E(H) \subset (V(H))^{(k)} = \{A \subset V(H) : |A| = k\}.$$

We sometimes call H a *k-graph* or even simply a *graph* when k is understood. The *size* of H is $e(H) = |E(H)|$ and its *order* is $|H| = |V(H)|$.

Given a family $\mathscr{F}$ of forbidden k-graphs, we say that a k-graph H is *$\mathscr{F}$-admissible* if no $F \in \mathscr{F}$ is a subgraph of H. Next, H is *$\mathscr{F}$-saturated* if it is $\mathscr{F}$-admissible and the addition of any extra edge to H violates this property. In other words, an $\mathscr{F}$-saturated graph is a maximal $\mathscr{F}$-admissible graph. Let

$$\mathrm{SAT}(n, \mathscr{F}) = \{H : H \text{ is } \mathscr{F}\text{-saturated}, \ |H| = n\}.$$

A typical extremal forbidden subgraph problem or 'Turán-type' problem asks about

$$\mathrm{ex}(n, \mathscr{F}) = \max\{e(H) : H \in \mathrm{SAT}(n, \mathscr{F})\}.$$

On the other hand, we can ask about the opposite extremum: how large is

$$\mathrm{sat}(n, \mathscr{F}) = \min\{e(H) : H \in \mathrm{SAT}(n, \mathscr{F})\}? \tag{1.1}$$

† Supported by an External Research Studentship, Trinity College, Cambridge, England.

We are interested only in the latter problem. If $\mathscr{F}$ has only a single member F, we write $\mathrm{sat}(n, F)$ instead of $\mathrm{sat}(n, \{F\})$, *etc.*

The sat-function was considered as early as the late 1940s by Zykov [13], but the theory seems to be less developed than the theory of Turán-type problems. In part, this might be the case because minimum saturated graphs are harder to handle. For example, as demonstrated in [8], the sat-function lacks many of the natural regularity properties of the ex-function (such as monotonicity).

Although a number of results have been obtained for special families $\mathscr{F}$ (see, for instance, [1, 5, 8]), little is known about $\mathrm{sat}(n, \mathscr{F})$ for a general family $\mathscr{F}$. Tuza in [11, 12] conjectured (also an unpublished conjecture of Bollobás of the late 1960s) that, for any k-hypergraph F,

$$\mathrm{sat}(n, F) = O(n^{k-1}). \tag{1.2}$$

The conjecture was proved by Kászonyi and Tuza [8] in the case $k = 2$; indeed they showed that $\mathrm{sat}(n, \mathscr{F}) = O(n)$ for all families $\mathscr{F}$ of 2-graphs, including infinite families.

Extending some results of Erdős, Hajnal and Moon [6], Bollobás [1] proved that

$$\mathrm{sat}(n, K_m^k) = \binom{n}{k} - \binom{n-m+k}{k}, \tag{1.3}$$

where K_m^k is the complete k-uniform hypergraph of order m. As the right-hand side of (1.3) is of order n^{k-1}, the estimate (1.2) is essentially best possible.

The main result of this paper is to prove Tuza's conjecture; in fact we will show that

$$\mathrm{sat}(n, \mathscr{F}) = O(n^{k-1}) \tag{1.4}$$

for all families of k-graphs for which the independence numbers are bounded by a constant. Also, we consider the question whether the limit $\lim_{n\to\infty} \mathrm{sat}(n, \mathscr{F})/n^{k-1}$ exists.

Then, we try to evaluate $\mathrm{sat}(n, F)$ for one special hypergraph. As we will see, even for rather simple hypergraphs the exact evaluation may turn out to be a hard task, since it may involve questions of design theory.

Finally, we investigate whether estimate (1.4) is true if reformulated for some classes of directed hypergraphs. Although it is not true for the class of all directed hypergraphs we demonstrate that for some natural subclasses a form of (1.4) holds.

There are many related results: we refer the reader to Bollobás's survey [3] where there is a section on minimum saturated graphs.

2. The size of saturated hypergraphs

2.1. The main theorem

Here we present the following estimate of $\mathrm{sat}(n, \mathscr{F})$ which implies Tuza's conjecture.

For a k-graph H we say that $A \subset V(H)$ is *independent* if it does not span an edge of H and the *independence number* $\alpha(H)$ is the size of a largest independent set in H. Formally,

$$\alpha(H) = \max\{|A| : A \subset V(H),\ A^{(k)} \cap E(H) = \emptyset\}.$$

Theorem 2.1. *Let $\mathscr{F}$ be a family of k-graphs such that $s = \max\{\alpha(F) : F \in \mathscr{F}\}$ exists, and let $s' = \min\{|F| : F \in \mathscr{F}\}$. Then, for any n,*

$$\operatorname{sat}(n, \mathscr{F}) < \left(s' - s + 2^{k-1}(s-1)\right)\binom{n}{k-1}. \tag{2.1}$$

Proof. It is enough to construct a graph $H \in \mathrm{SAT}(n, \mathscr{F})$ whose size does not exceed the stated bound. Our construction will be by means of an algorithm.

Our algorithm works in the following way. Let us agree that the vertex set is $X = \{1, \dots, n\}$ with the usual ordering. Given $x \in X$ and $B \subset X$, we write $B < x$ if every vertex in B is smaller than x. By $U_x = \{y \in X : y > x\}$ we denote the *upper shadow of* x and in the obvious way we define the *lower shadow* L_x. If $|B| \leqslant k$, say B consists of elements $b_1 < \dots < b_i$, $i \leqslant k$, then we define its *tail*

$$\mathscr{T}_B = \{\{b_1, \dots, b_i, x_{i+1}, \dots, x_k\} : b_i < x_{i+1} < \dots < x_k\} \subset X^{(k)}. \tag{2.2}$$

We construct an $\mathscr{F}$-saturated graph H by starting with the empty hypergraph H on X and adding to H one by one certain families of edges until we obtain $H \in \mathrm{SAT}(n, \mathscr{F})$.

The algorithm is rather simple. We take, one by one in order, the vertices of X. For every vertex x, we consider all of the i-subsets of L_x, beginning with $i = 0$ and increasing i until $i = k - 1$. For every such subset $A < x$, we consider $\mathscr{T}_B$, $B = A \cup \{x\}$, which is, by definition, the family of k-subsets having B as an initial segment. If at this moment $\mathscr{T}_B \not\subset E(H)$ and the addition of $\mathscr{T}_B$ to the edge set of H does not create any forbidden subgraph, we add $\mathscr{T}_B$ to H. This is a crucial feature of the algorithm: for every x and A we either add *all* of $\mathscr{T}_B$ or we add *nothing*.

Another important detail is the order of the steps. The outermost cycle has x *increasing* from 1 to n. The next cycle runs for i *increasing* from 0 to $k - 1$. In the innermost cycle we consider all i-subsets of L_x and here we are free to choose them in any order, but for uniformity let us agree that we use here the colex order.

In the course of the algorithm we define, on the vertex set X, auxiliary hypergraphs $H_1, \dots, H_n$ and $G_1, \dots, G_k$, which we need for an estimation of $e(H) = |E(H)|$. The k-hypergraph H_x contains precisely those edges which were added whilst considering vertices from 1 to x inclusive. The i-hypergraph G_i contains as edges those i-subsets B for which the set $\mathscr{T}_B$ was added to H.

We claim that the resulting graph $H = H_n$ is an $\mathscr{F}$-saturated graph. Indeed, H is $\mathscr{F}$-admissible, as we were adding edges only if they did not produce any forbidden subgraphs. On the other hand, take any k-subset E not in $E(H)$. We did not use the opportunity to add E to $E(H)$ when $x = \max E$, $i = k - 1$ and $A = E \setminus \{x\}$ (when $\mathscr{T}_B = \{E\}$). The only reason for our not doing so is that the addition of E would have created a forbidden subgraph F. Then, certainly, $H + E$ contains F, which shows $H \in \mathrm{SAT}(n, \mathscr{F})$.

We now show that $e(G_1) \leqslant s' - 1$ and

$$e(G_i) \leqslant (s-1)\binom{n}{i-1}, \quad i = 2, \dots, k. \tag{2.3}$$

This will establish the theorem, as then

$$\begin{aligned} e(H) &\leqslant \sum_{i=1}^{k}\binom{n-i}{k-i}e(G_i) < (s'-s)\binom{n}{k-1} + (s-1)\sum_{i=1}^{k}\binom{n-i+1}{k-i}\binom{n}{i-1} \\ &= (s'-s)\binom{n}{k-1} + (s-1)\sum_{i=1}^{k}\binom{n}{k-1}\binom{k-1}{i-1} \\ &= \left(s'-s+2^{k-1}(s-1)\right)\binom{n}{k-1}. \end{aligned}$$

Assume that, for some i, $2 \leqslant i \leqslant k$, estimate (2.3) is not true. Then there is some $(i-1)$-set $V = \{v_1,\ldots,v_{i-1}\}$, $v_1 < \cdots < v_{i-1}$, which is the initial segment of at least s edges of G_i. Let $E_1,\ldots,E_s \in E(G_i)$ be s distinct edges containing V as initial segment, say $E_j = V \cup \{z_j\}$, $j = 1,\ldots,s$, $V < z_1 < \cdots < z_s$.

Since $E_1 \in E(G_i)$, all edges whose initial segment is E_1 were added to H at the moment when $x = z_1$ and $A = V$. It follows that $V \notin E(G_{i-1})$, for otherwise these edges would already be present in H. The only reason that we did not add V to $E(G_{i-1})$ earlier when $x = v_{i-1}$ and $A = \{v_1,\ldots,v_{i-2}\}$ must have been that the hypergraph $H' = H_{v_{i-1}} + \mathscr{T}_V$ contains some forbidden subgraph F. Note that $U_{v_{i-1}}$ is an independent set in H', therefore $|V(F) \cap U_{v_{i-1}}| \leqslant s$, by the assumption on $\mathscr{F}$.

By the way the algorithm works, any permutation σ of X affecting only the upper shadow U_z of a vertex $z \in X$ (that is, $\sigma(y) = y$ for all $y \leqslant z$) is an automorphism of H_z, because any $\mathscr{T}_B \subset X^{(k)}$ with $\max B \leqslant z$ is σ-invariant.

Applying this remark to $z = v_{i-1}$, we see that we may assume, since $|V(F) \cap U_{v_{i-1}}| \leqslant s$, that

$$V(F) \cap U_{v_{i-1}} \subset Z = \{z_1,\ldots,z_s\}. \tag{2.4}$$

Now let $E \in E(F)$. Then either $E \in E(H_{v_{i-1}}) \subset E(H)$, or else letting $z_j = \min(E \cap \{z_1,\ldots,z_s\})$ we see by (2.4) that $E \in \mathscr{T}_{E_j}$. Since $E_j \in E(G_i)$ we obtain that in both cases $E \in E(H)$. But then $F \subset H$, which is a contradiction, so (2.3) is proved for $2 \leqslant i \leqslant k$.

The case $i = 1$ does not fall into general scheme of the proof. However, it is rather trivial, for if we have at least s' edges (one-element subsets) in G_1, say $\{v_1\},\ldots,\{v_{s'}\} \in E(G_1)$, then these vertices span a complete k-graph in H, because if $E \in \{v_1,\ldots,v_{s'}\}^{(k)}$ then $E \in \mathscr{T}_{\{\min E\}} \subset E(H)$. Therefore H contains every k-graph of order s', which is certainly a contradiction. □

Corollary 2.1. *For any finite family $\mathscr{F}$ of k-graphs, we have* $\mathrm{sat}(n,\mathscr{F}) = O(n^{k-1})$. □

Remark. Our construction is not generally best possible. For example, for the 2-graph consisting of 2 disjoint edges the sat-function equals 3, while our algorithm gives $n-1$.

Kászonyi and Tuza [8] showed that (1.4) is, in fact, valid for any infinite family of 2-graphs. Thus, an interesting question which still remains open is whether the estimate (1.4) is true for any infinite family $\mathscr{F}$, $k \geqslant 3$. Our construction does not settle this question, although possibly some modification of it can do the job.

The algorithm has an interesting 'convergence' property which is true for any (finite or infinite) family $\mathscr{F}$. Namely, for any fixed natural m and any $i = 1, \ldots, k$, if n is sufficiently large, $G_i([m])$ (the graph spanned by the first m vertices of G_i) does not depend on n. Consequently, any finite part of the resulting graph $H \in \mathrm{SAT}(n, \mathscr{F})$ is eventually constant, which allows us to define the 'limiting' $\mathscr{F}$-saturated graph H_∞ of infinite order. Unfortunately, we have not found any interesting application of this object so far.

2.2. The asymptotic behaviour of the sat-function

Let us turn to another conjecture from [12], namely that, for every 2-graph F, the limit

$$c(\mathscr{F}) = \lim_{n\to\infty} \frac{1}{n}\mathrm{sat}(n, \mathscr{F}), \tag{2.5}$$

exists, for $\mathscr{F} = \{F\}$.

We can show that (2.5) is not generally true for an infinite family $\mathscr{F}$ of 2-graphs.

Let us take as forbidden graphs $K_{1,m}$ (a star), some fixed $m \geqslant 4$, and cycles

$$C_{b_i}, C_{b_i+1}, \ldots, C_{2b_i}, \qquad i \in \mathbb{N},$$

for some 'fast growing' sequence (b_i). For $n = 2b_i + 1$ we have $\mathrm{sat}(n, \mathscr{F}) \leqslant n$ because a cycle on n vertices does not contain any forbidden subgraph, but the addition of any chord produces a cycle of length at least $n/2$, so $C_n \in \mathrm{SAT}(n, \mathscr{F})$. Thus $\underline{\lim}\ \mathrm{sat}(n, \mathscr{F})/n \leqslant 1$.

Now, for $n = b_i - 1$ any $G \in \mathrm{SAT}(n, \mathscr{F})$ can have at most

$$v = \sum_{l=0}^{2b_{i-1}-1} (m-2)^l < (m-2)^{2b_{i-1}}$$

vertices of degree strictly less than $m-1$. Indeed, let x be a vertex of degree at most $m-2$. As $\Delta(G) \leqslant m-1$ (which is equivalent to $K_{1,m} \not\subset G$), there are at most $v-1$ vertices at distance at most $2b_{i-1} - 1$ from x. Now, for any $y \in G$ with $d(x, y) \geqslant 2b_{i-1}$, the addition of edge $\{x, y\}$ to G can create neither a cycle of length at most $2b_{i-1}$ nor a cycle of length greater than $|G| = b_i - 1$. So, $d(y) = m-1$ and all vertices of degree at most $m-2$ are confined to the ball around x containing in total at most v vertices. Therefore

$$e(G) \geqslant \frac{(n-v)(m-1)}{2} = \frac{m-1+o(1)}{2} n$$

if our sequence satisfies $(m-2)^{2b_{i-1}} = o(b_i)$.

From these considerations we conclude that the limit (2.5) does not necessarily exist for an infinite family $\mathscr{F}$.

3. Butterflies

For $k \geqslant 3$ many totally new difficulties arise that are not present when $k = 2$. There are some links with design theory, and, as designs are quite hard to handle, the same applies to saturated hypergraphs. We will provide an example to demonstrate the point.

A *Steiner triple system* $S_\lambda(2, 3, v)$ is a 3-hypergraph of order v in which every two distinct vertices are covered by exactly λ edges. Later we will be interested in the existence of an $S_\lambda(2, 3, v)$. Dehon [4] characterized all possible pairs (λ, v), as follows.

Theorem 3.1. *Let λ and v be positive integers. There exists an $S_\lambda(2,3,v)$ if and only if $\lambda \leqslant v-2$, $\lambda v(v-1) \equiv 0 \pmod 6$ and $\lambda(v-1) \equiv 0 \pmod 2$.* □

We consider $\mathrm{sat}(n, B_m^3)$, where the 3-graph B_m^3 (called a *butterfly*) has vertex set $[m+2]$ and m edges $\{1,2,i\}$, $i = 3,\ldots,m+2$. Given $H \in \mathrm{SAT}(n, B_m^3)$, we build a 2-graph G, so that $\{x,y\} \in E(G)$ if there are $m-1$ edges in H containing both x and y. Let $e = e(G)$. Note that if some vertices $x,y,z \in G$ do not span a single edge then $\{x,y,z\} \in E(H)$, and so

$$e(H) \geqslant (m-1)e/3 + k_3(\overline{G}), \tag{3.1}$$

where $k_3(\overline{G})$ means the number of triangles in the complement of G. Corollary VI.1.6 from [2] (which can be deduced from the results of Moon and Moser [10]) states that $k_3(\overline{G}) \geqslant \bar{e}(4\bar{e}-n^2)/(3n)$, where $\bar{e} = \binom{n}{2} - e$ denotes the number of edges in the complement of G. We aim to show that

$$\mathrm{sat}(n, B_m^3) \geqslant \frac{(m-1)e(\overline{T_2(n)})}{3}, \tag{3.2}$$

where $T_2(n)$ is the Turán graph of order n, so $e(\overline{T_2(n)}) = \binom{n}{2} - \lfloor n^2/4 \rfloor$.

Assume that $n = 2v$ is even, then $e(\overline{T_2(n)}) = v(v-1)$. By (3.1) we are home if $e \geqslant v(v-1)$, so assume the contrary. Note that $\bar{e} = v(2v-1) - e$ and that $e < v(v-1)$ implies $\bar{e} > v^2$. Then we have

$$\begin{aligned} e(H) - (m-1)v(v-1)/3 &\geqslant k_3(\overline{G}) + \frac{m-1}{3}(e - v(v-1)) \\ &\geqslant (\bar{e} - v^2)\left(\frac{2\bar{e}}{3v} - \frac{m-1}{3}\right) \\ &> (\bar{e} - v^2)(2v - m + 1)/3 \geqslant 0 \end{aligned}$$

for $v \geqslant (m-1)/2$, that is, $n \geqslant m-1$.

By similar reasoning we can show that (3.2) is true for any odd $n \geqslant 4(m-1)/3$.

On the other hand, suppose that there exist $H_i = S_{m-1}(2,3,v_i)$, $i = 1,2$, $v_1 = \lfloor n/2 \rfloor$ and $v_2 = \lceil n/2 \rceil$; define H to be their disjoint union, $H = H_1 \sqcup H_2$.

We claim that $H \in \mathrm{SAT}(n, B_m^3)$. Indeed, any 3-edge $E \notin E(H)$ has at least 2 vertices in one of the components, say $x, y \in H_1$. By definition, H_1 has $m-1$ edges containing $\{x,y\}$ which, together with the edge E, form a butterfly. This shows that for such values of n we have $\mathrm{sat}(n, B_m^3) \leqslant (m-1)e(\overline{T_2(n)})/3$.

The above discussion combined with Theorem 3.1 gives a proof of the following.

Theorem 3.2. *Let $m \geqslant 2$ and n be natural numbers with $n \geqslant m+2$ if n is even or $n \geqslant 4(m-1)/3$ if n is odd. Then the inequality* (3.2) *is true.*

We have equality in (3.2) *if and only if $n \geqslant 2m+2$ and there is an integer l such that $m-1 = 2l$ and $n \equiv 0,1,7$ or 8 $\pmod{12}$, $m-1 = 3l$ and $n \equiv 10 \pmod{12}$, $m-1 = l$ and $n \equiv 2$ or 6 $\pmod{12}$, or $m-1 = 6l$ and $n \equiv 3,4,5,9$ or 11 $\pmod{12}$.* □

Characterization of non-isomorphic Steiner triple systems is in the embryo state, so we cannot say much more about extremal graphs in the cases covered by Theorem 3.2 than that $G = G(H)$ is isomorphic to the complement of the Turán graph and any 2 vertices of H are contained in either $m-1$ edges of H or none.

It is hard to find exact values of $\mathrm{sat}(n, B_m^3)$ for all n. We will only provide some upper bound which may be considered as a satisfactory answer to the problem.

If n is even, take a representation $n = n_1 + n_2$ with n_1 and n_2 being congruent to 1 or 3 modulo 6 and $|n_1 - n_2| \leqslant 6$. Then $S_{m-1}(2,3,n_1) \sqcup S_{m-1}(2,3,n_2) \in \mathrm{SAT}(n, B_m^3)$ and

$$\mathrm{sat}(n, B_m^3) \leqslant (m-1)\left(\binom{n_1}{2} + \binom{n_2}{2}\right)/3 \leqslant (m-1)\left(e(\overline{T_2(n)}) + 9\right)/3.$$

If n is odd, we represent similarly $n + 1 = n_1 + n_2$, take the union of the corresponding designs that share one vertex and deduce that

$$\mathrm{sat}(n, B_m^3) \leqslant (m-1)\left(e(\overline{T_2(n)}) + \tfrac{n-1}{2} + 9\right)/3.$$

Of course, if we know that 2 or 3 divides $m-1$ then we can do better.

4. Directed hypergraphs

Here we shall consider, roughly speaking, k-hypergraphs with the additional structure of directed edges.

Actually, many different but natural definitions suggest themselves, but we will consider in more detail the following classes (defined below): acyclic directed hypergraphs and ordered hypergraphs. We shall ask whether the estimate

$$\mathrm{sat}(n, \mathscr{F}) = O(n^{k-1}) \tag{4.1}$$

is true for a general k-hypergraph family $\mathscr{F}$ and for the appropriately defined sat-function.

Note that we can give a uniform definition of saturatedness: *cf.* Tuza [11]. Suppose we have a class of objects $\mathscr{C}$, with a binary relation '$\subset$' and a rank function $r : \mathscr{C} \to \mathbb{N}$. Given a family $\mathscr{F} \subset \mathscr{C}$, we say that $H \in \mathscr{C}$ is *$\mathscr{F}$-admissible* if H does not contain an $F \in \mathscr{F}$ as a subobject. Now, let $\mathrm{SAT}(n, \mathscr{F})$ be the family of maximal $\mathscr{F}$-admissible objects of rank n; H is called *$\mathscr{F}$-saturated* if $H \in \mathrm{SAT}(r(H), \mathscr{F})$.

In our cases, $\mathscr{C}$ is the class of hypergraphs with some additional structure, $G \subset H$ holds if H contains a copy of G as a substructure, and $r(H) = |H|$, $G, H \in \mathscr{C}$.

4.1. Cycle-free directed hypergraphs

To obtain a *directed hypergraph* we take a usual hypergraph and on every one of its edges introduce some *orientation*, that is, a linear order. In fact, estimate (4.1) is not generally true for directed hypergraphs, which is exhibited by the *directed* 3-*cycle* C_3 consisting of edges $(1,2)$, $(2,3)$ and $(3,1)$: improving previous bounds of Katona and Szemerédi [9], Füredi, Horak, Pareek and Zhu [7] showed that $\mathrm{sat}(n, C_3) \approx n \log_2 n$.

But let us consider cycle-free hypergraphs. We say that a directed hypergraph H is *cycle-free* or *acyclic* if there is no *cycle* which is by definition an alternating sequence of vertices and edges $(x_1, E_1, \ldots, x_l, E_l, x_{l+1} = x_1)$ such that x_i precedes x_{i+1} in E_i. Equivalently, H is cycle-free if we can order its vertices in a way compatible with the ordering of the edges.

By definition, H is $\mathscr{F}$-saturated if no $F \in \mathscr{F}$ is a subgraph of H but the addition of any new (ordered) edge to G creates either a forbidden subgraph or a directed cycle. For a directed hypergraph H let $\alpha(H)$ be equal to the maximum number of vertices in H not spanning an edge.

Theorem 4.1. *In the class of the cycle-free k-graphs for any family $\mathscr{F}$ with bounded $\{\alpha(F) : F \in \mathscr{F}\}$ we have* $\mathrm{sat}(n, \mathscr{F}) = O(n^{k-1})$.

Proof. We proceed essentially in the same way as in the proof of Theorem 2.1, so we briefly describe the differences.

Consider one by one $x \in X = [n]$, $i = 0, \dots, k-1$, $A \in L_x^{(i)}$. Let $B = A \cup \{x\}$ and let $\mathscr{T}_B$ be defined by (2.2). An orientation of the edges in $\mathscr{T}_B$ is called *symmetric* if any order preserving injections $f, g : [k] \to [n]$ with $f([k]), g([k]) \in \mathscr{T}_B$ induce identical orientations of $[k]$.

If $\mathscr{T}_B \not\subset E(H)$ (as unoriented k-tuples) and there exists a symmetric orientation of $\mathscr{T}_B$ such that $H + \mathscr{T}_B$ does not contain a forbidden subgraph or a cycle, then we add $\mathscr{T}_B$ (with this orientation) to the edge set of H.

That is the algorithm. The obtained hypergraph H does not contain a forbidden configuration. As every k-subset $E \subset X$ was tested (for $B = E$ we had $\mathscr{T}_B = \{E\}$ and every orientation was symmetric), we conclude that $H \in \mathrm{SAT}(n, \mathscr{F})$.

As in Theorem 2.1 we define the auxiliary hypergraphs H_x (directed) and G_i (undirected). We have to show that $e(G_i) = O(n^{i-1})$.

First, suppose that $E(G_1) = \{\{x_1\}, \dots, \{x_l\}\}$. One can easily check that, as H is cycle-free, there is no choice for the orientation of the edges of $\mathscr{T}_{\{x_i\}}$, $2 \leqslant i \leqslant l$ and H contains the complete cycle-free k-graph on l vertices, which implies $l = O(1)$.

Suppose that $e(G_i) \neq O(n^{i-1})$, for some $1 < i \leqslant k$. Then, for some $(i-1)$-tuple $V \subset X$, we can find an arbitrarily large set $Z = \{z_1, \dots, z_s\} \subset U_x$, $x = \max V$, such that $V \cup \{z_i\} \in E(G_i)$, $i \in [s]$, and the orientation of $\cup_{i \in [s]} \mathscr{F}_{V \cup \{z_i\}} \subset E(H)$ extends to a symmetric orientation '$\prec$' of $\mathscr{T}_V$. As $V \notin E(G_{i-1})$ we conclude that $H' = H_x + (\mathscr{T}_V, \prec)$ contains a forbidden subgraph F or a cycle. If a copy of F is present we follow the proof of Theorem 2.1. Otherwise let $C = (y_1, E_1, \dots, y_l, E_l, y_{l+1} = y_1)$ be a shortest cycle in H'.

We claim that C can be chosen so that $|W| \leqslant 3k - 5$, where $W = (\cup_{i \in [l]} E_i) \cap U_x$. Then for $s \geqslant 3k - 5$ we may assume that $W \subset Z$, and the argument of Theorem 2.1 shows that $C \subset H$, which is a contradiction, thus proving the theorem.

If $Y = \{y_1, \dots, y_l\} \subset U_x$ then $l \leqslant 2$ and the claim is true. Indeed, there is an $i \in [l]$ such that y_{i+1} is larger than y_i and y_{i+2} in $[n]$ but it follows y_i in E_i and precedes y_{i+2} in E_{i+1} which, by the symmetry of $U_x \subset H'$, implies that any two $y, y' \in U_x$ form a 2-cycle.

Next, $|Y \cap U_x| \leqslant 1$; otherwise pick $y_h, y_i \in U_x \cap Y$, $h < i$, with $y_{i+1} \in Y \setminus U_x$ and obtain a strictly shorter cycle through $(y_1, \dots, y_h, y_{i+1}, \dots, y_{l+1} = y_1)$ as $U_x \subset H'$ is 'symmetric'. The two edges containing the point (if one exists) in $Y \cap U_x$ contribute at most $2k - 3$ to $|W|$. By the symmetry of U_x, we can assume that for the remaining edges $E_i \cap U_x$ lies within some fixed $(k-2)$-subset of U_x, which shows that $|W| \leqslant 3k - 5$. □

For $k = 2$ we can prove a stronger result which includes all infinite families.

Theorem 4.2. *In the class of the cycle-free digraphs we have* $\mathrm{sat}(n, \mathscr{F}) = O(n)$ *for any family* $\mathscr{F}$.

Proof. It is enough to provide a construction. Order the vertex set $X = \{x_1, \ldots, x_n\}$. Repeat the following as long as no forbidden subgraph appears: take the next vertex x of X and add all of $\mathscr{T}_x$. Here, $\mathscr{T}_x$ is the set of the (oriented) edges of the form xy, $y \in U_x$.

Suppose we have repeated the iteration $m = m(n)$ times. Let $G' = G'(n)$ be the graph received after these m steps. As $[m] \subset V(G')$ spans the complete digraph, the number of iterations is bounded by a constant not depending on n: namely, $m < u$, where $u = \min\{|F| : F \in \mathscr{F}\}$.

Obviously, $m(n)$ is non-increasing as a function of n for $n \geqslant u$, so it is constant for n sufficiently large. Then, the reason for terminating the procedure is that the addition of $\mathscr{T}_{x_{m+1}}$ would create a forbidden subgraph F and it will be the case for any subsequent n, that is, $G'(n) + \mathscr{T}_{x_{m+1}}$ contains the same subgraph F.

Now we add edges to G' in any order as long as we create neither a cycle nor a forbidden subgraph. In the resulting graph G, no $d = |V(F) \cap U_{x_{m+1}}|$ edges can start at the same vertex $y \in U_{x_m}$, as otherwise we have a subgraph isomorphic to F.

So, the number of edges in G is at most $m(n-1) - \binom{m}{2} + (n-m)(d-1) = O(n)$. □

Actually, one can argue that, for sufficiently large n,

$$m = \min\{|F| - \alpha'(F) : F \in \mathscr{F}\} - 1,$$

where $\alpha'(F)$ is the maximum size of $A \subset V(F)$ such that no edge starts in A. Equivalently, m is the minimum number of vertices one needs to remove from some $F \in \mathscr{F}$ to obtain a *directed star* (a digraph whose edges start at a common vertex). For d we can take the size of any such star. This observation allows us to write more explicitly the bound of Theorem 4.2.

4.2. Ordered hypergraphs

We can introduce yet another class: *ordered k-graphs*. An ordered k-graph is a usual (unoriented) k-graph with an extra structure: we have a fixed ordering on the vertex set and the vertices of a subgraph inherit their order from the original graph. To avoid confusion note that an ordered graph comes equipped with a *fixed* vertex ordering while a cycle-free graph is one that admits at least one compatible vertex ordering.

Without any difficulties we can restate word by word the proof of Theorem 2.1 (except that now we have already been given an order on the vertex set and in the construction we take the vertices in this order). We change the definition of the independence number for ordered hypergraphs to obtain a stronger result. Namely, we define

$$\alpha(F) = \max\{|U_x| : U_x \text{ does not span an edge, } x \in V(F)\}.$$

Theorem 4.3. *Let* $\mathscr{F}$ *be a family of ordered k-graphs such that the set* $\{\alpha(F) : F \in \mathscr{F}\}$ *is bounded. Then* $\mathrm{sat}(n, \mathscr{F}) = O(n^{k-1})$. □

Using the ideas of Theorem 4.2 one can see that for $k = 2$ our result can be extended to all infinite families.

Theorem 4.4. *For any family $\mathscr{F}$ of ordered 2-graphs we have* $\mathrm{sat}(n, \mathscr{F}) = O(n)$. □

It is easy to give an example of an ordered k-hypergraph with sat-function of order n^{k-1}: for instance, complete hypergraphs for which the answer is given by formula (1.3), the same as for usual hypergraphs.

Trivial examples show that the estimate (1.4) is not generally true if we enlarge any of the above classes by allowing edges of different sizes (up to k) and/or multiple edges (and defining the sat-function respectively).

Acknowledgements

The author would like to express his gratitude to Andrew Thomason for very stimulating discussions and the remarks that greatly improved the final version of the paper.

References

[1] Bollobás, B. (1965) On generalized graphs. *Acta Math. Acad. Sci. Hungar.* **16** 447–452.
[2] Bollobás, B. (1978) *Extremal Graph Theory*, Academic Press, London.
[3] Bollobás, B. (1995) Extremal graph theory. In *Handbook of Combinatorics* (R. L. Graham, M. Grötschel and L. Lovász, eds), North-Holland, pp. 1231–1292.
[4] Dehon, M. (1983) On the existence of 2-designs $S_\lambda(2, 3, v)$ without repeated blocks. *Discrete Mathematics* **43** 155–171.
[5] Erdős, P., Füredi, Z. and Tuza, Z. (1991) Saturated r-uniform hypergraphs. *Discrete Mathematics* **98** 95–104.
[6] Erdős, P., Hajnal, A. and Moon, J. W. (1964) A problem in graph theory. *Amer. Math. Monthly* **71** 1107–1110.
[7] Füredi, Z., Horak, P., Pareek, C. M. and Zhu, X. (1998) Minimal oriented graphs of diameter 2. *Graphs Combin.* **14** 345–350.
[8] Kászonyi, L. and Tuza, Z. (1986) Saturated graphs with minimal number of edges. *J. Graph Theory* **10** 203–210.
[9] Katona, G. and Szemerédi, E. (1967) On a problem of graph theory. *Studia Sci. Mat. Hungar.* **2** 23–28.
[10] Moon, J. W. and Moser, L. (1962) On a problem of Turán. *Publ. Math. Inst. Hungar. Acad. Sci.* **7** 283–287.
[11] Tuza, Z. (1986) A generalization of saturated graphs for finite languages. In *Proc. IMYCS'86, Smolenice Castle* (*Czechoslovakia*), Vol. 185 of MTA SZTAKI Studies, pp. 287–293.
[12] Tuza, Z. (1988) Extremal problems on saturated graphs and hypergraphs. *Ars Combinatoria* **25B** 105–113.
[13] Zykov, A. A. (1949) On some properties of linear complexes. *Mat. Sbornik N.S.* **24** 163–188. (In Russian.)