On an Extremal Hypergraph Problem of Brown, Erdős and Sós

Noga Alon * Asaf Shapira †

Abstract

Let $f_r(n, v, e)$ denote the maximum number of edges in an r-uniform hypergraph on n vertices, which does not contain e edges spanned by v vertices. Extending previous results of Ruzsa and Szemerédi and of Erdős, Frankl and Rödl, we partially resolve a problem raised by Brown, Erdős and Sós in 1973, by showing that for any fixed $2 \leq k < r$, we have

$$n^{k-o(1)} < f_r(n, 3(r-k)+k+1, 3) = o(n^k).$$

1 Introduction

All the hypergraphs considered here are finite and have no parallel edges. An r-uniform hypergraph (=r-graph for short) $H = (V, E)$, is a hypergraph in which each edge contains precisely r distinct vertices of V. Denote by $f_r(n, v, e)$ the largest number of edges in an r-graph on n vertices that contains no e edges spanned by v vertices. Estimating the asymptotic growth of this function for fixed integers r, e and v is one of the most well studied problems in extremal graph theory. In particular, when $e = \binom{v}{r}$ we get the well known Turán problem of determining the maximum possible number of edges in an r-graph that contains no complete r-graph on v vertices. See the surveys [11], [8] and [20] for results and references on this and other graph and hypergraph Turán problems. In 1973, Brown, Erdős and Sós [5],[6] initiated the study of the function f for r-graphs. A general case they managed to resolve was that for every $2 \leq k < r$ and $e \geq 3$

$$f_r(n, e(r-k)+k, e) = \Theta(n^k),$$

where the upper bound follows from the observation that in any r-graph that contains no e edges on $e(r-k)+k$ vertices, any set of k vertices belongs to at most $e-1$ edges, and the lower bound is obtained by a (by now) standard application of the probabilistic deletion method. This suggested the much more difficult problem of computing the asymptotic value of

$$f_r(n, e(r-k)+k+1, e). \tag{1}$$

Even in the simplest case of (1), where $r = e = 3$ and $k = 2$, the authors of [5],[6] were only able to obtain $\Omega(n^{3/2}) = f_3(n, 6, 3) = O(n^2)$. The problem of estimating $f_3(n, 6, 3)$ became later known as

*Schools of Mathematics and Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Email: nogaa@tau.ac.il. Research supported in part by a USA-Israeli BSF grant, by the Israel Science Foundation and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University.

†School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Email: asafico@tau.ac.il. This work forms part of the author's Ph.D. Thesis.

the $(6, 3)$-problem. In one of the classical results in extremal combinatorics, Ruzsa and Szemerédi [18] resolved the $(6, 3)$-problem by proving that

$$n^{2-o(1)} < f_3(n, 6, 3) = o(n^2). \tag{2}$$

In the above, as well as throughout this paper, a $o(1)$ term will represent a quantity that approaches 0, as n tends to infinity, whereas $o(n^k)$ denotes, as usual, $o(1) \cdot n^k$. In 1986, Erdős, Frankl and Rödl [9] extended the result of [18] to arbitrary fixed r (and $e = 3$, $k = 2$ as in [18]), by showing that

$$n^{2-o(1)} < f_r(n, 3(r-3)+3, 3) = o(n^2). \tag{3}$$

Since then, the only progress on the asymptotic value of (1) was obtained by Sárközy and Selkow [19], who managed to prove some nearly tight upper bounds. Specifically, they showed that

$$f_r(n, e(r-k) + k + \lfloor \log_2(e) \rfloor, e) = o(n^k). \tag{4}$$

Note that the left hand side is obtained from (1) by replacing the 1 by $\lfloor \log_2(e) \rfloor$. As $\lfloor \log_2(3) \rfloor = 1$ this gives upper bounds for $e = 3$ and arbitrary $2 \le k < r$ in (1). No lower bounds were given since the result of [9]. Our main goal in this paper is to prove the following theorem, which extends the result of Erdős, Frankl and Rödl (3) (and therefore also the result of Ruzsa and Szemerédi (2)) by determining the asymptotic behavior of (1) for $e = 3$ and arbitrary $2 \le k < r$ as follows.

Theorem 1 *For any fixed $2 \le k < r$ we have,*

$$n^{k-o(1)} < f_r(n, 3(r-k) + k + 1, 3) = o(n^k).$$

As we have mentioned above, the upper bound given in Theorem 1 can be derived from (4). However, as observed in Section 3, this special case (that includes the upper bound of (3) as well), can be proved by a simple reduction to the upper bound of the (6,3)-problem.

The main difficulty in the proof of Theorem 1 is the proof of the lower bound. As in [18] and [9], one of our tools is a number theoretic construction, which is closely related to that of Behrend [4]. In Section 3 we use this number theoretic construction in order to construct the r-graphs needed to prove the lower bound of Theorem 1. In Section 4 we prove the main technical lemma needed in order to prove the correctness of the construction, namely that these r graphs do not contain 3 edges spanned by $3(r-k)+k+1$ vertices. Unlike the cases studied in [18] and [9], the main difficulty in the proof is that there are many possible configurations of 3 edges spanned by $3(r-k)+k+1$ that we have to rule out, while in [18] and [9] there was (essentially) only one such possible configuration. In order to rule out all the possible configurations, we give in Section 2 an algebraic construction of a certain pseudo-random matrix which we also use in our construction. This is done by using some properties of multivariate polynomials. The main ideas behind the construction given in Section 3, as well as the proof of its correctness in Section 4 are somewhat motivated by the ideas of [3], though the proof here is more involved and critically relies on the construction given in Section 2. In Section 5 we discuss some open problems as well as some additional observations about the asymptotic value of (1).

2 The Matrix

In this section we discuss the construction of a pseudo-random matrix, which will be a central ingredient in the construction of the r-graph required to obtain the lower-bound in Theorem 1. This will be done in Lemma 2.2. We first discuss *variable matrices*, namely matrices whose entries contain unknowns $x_{i,j}$ rather than real numbers. To do so, we define a certain type of matrix which we call a *proper matrix*. This type of matrix will be useful in the analysis of the construction which is given in Section 3. For an integer $k \geq 2$ we say that a $(2k-1) \times (2k-1)$ variable matrix is a proper-matrix if we can partition its columns into 3 groups T_1, T_2, T_3 of sizes t_1, t_2, t_3 respectively, such that:

1. For $1 \leq i \leq 3$ we have $1 \leq t_i \leq k-1$.

2. Any column v_i that belongs to T_1 is of the form $(x_{1,i}, x_{2,i}, \ldots, x_{k,i}, x_{2,i}, \ldots, x_{k,i})$, that is, the last $k-1$ variables must be the same as those that appear in entries $2, \ldots, k$, respectively.

3. Any column v_i that belongs to T_2 is of the form $(x_{1,i}, x_{2,i}, \ldots, x_{k,i}, 0, \ldots, 0)$, that is, the last $k-1$ variables must be identically zero.

4. Any column v_i that belongs to T_3 is of the form $(0, 0, \ldots, 0, x_{2,i}, \ldots, x_{k,i})$, that is, the variables that appear in entries $1, \ldots, k$ must be identically zero.

5. All the variables that appear in the upper k rows of the matrix are distinct.

6. All the variables that appear in the lower $k-1$ rows of the matrix are distinct.

7. There are at least k columns in $T_2 \cup T_3$ that have no common variables. Namely, if T_2 and T_3 have d columns that share some variables, then
$$t_3 - d \geq k - t_2. \tag{5}$$
Moreover, the only way a column $v \in T_3$ can share variables with a column $u \in T_2$, is that $u = (x_{i,1}, x_{i,2}, \ldots, x_{i,k}, 0, \ldots, 0)$ and $v = (0, 0, \ldots, 0, x_{i,2}, \ldots, x_{i,k})$, that is, the last $k-1$ variables of v are the variables numbered $2, \ldots, k$ of v in the same order as they appear in v.

Figure 1 depicts a proper-matrix of size 9×9, where $k = 5$ and $t_1 = t_2 = t_3 = 3$. The reader is advised to verify that it indeed satisfies all the properties of a proper-matrix. In what follows, the *degree* of a multivariate polynomial will denote the largest exponent of any variable in the expansion of the polynomial as a sum of monomials. We need the following simple yet somewhat technical claim.

Claim 2.1 *The determinant of any proper-matrix is a non-zero multivariate polynomial of degree at most 2 in each variable.*

Proof: The fact that the determinant is a multivariate polynomial of degree 2 follows from the definition of the determinant by observing that each variable appears at most twice in any proper matrix. We thus only have to show that it is not identically zero. It is clearly enough to show that for any proper-matrix P we can assign its variables 0/1 values such that the determinant of the resultant matrix is ± 1. Let P be a proper-matrix. Note that as $1 \leq t_1, t_2, t_3 \leq k-1$ (property 1) we

$$
\begin{pmatrix}
x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} & x_{1,5} & x_{1,6} & 0 & 0 & 0 \\
x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} & x_{2,5} & x_{2,6} & 0 & 0 & 0 \\
x_{3,1} & x_{3,2} & x_{3,3} & x_{3,4} & x_{3,5} & x_{3,6} & 0 & 0 & 0 \\
x_{4,1} & x_{4,2} & x_{4,3} & x_{4,4} & x_{4,5} & x_{4,6} & 0 & 0 & 0 \\
x_{5,1} & x_{5,2} & x_{5,3} & x_{5,4} & x_{5,5} & x_{5,6} & 0 & 0 & 0 \\
x_{2,1} & x_{2,2} & x_{2,3} & 0 & 0 & 0 & x_{2,4} & x_{2,7} & x_{2,8} \\
x_{3,1} & x_{3,2} & x_{3,3} & 0 & 0 & 0 & x_{3,4} & x_{3,7} & x_{3,8} \\
x_{4,1} & x_{4,2} & x_{4,3} & 0 & 0 & 0 & x_{4,4} & x_{4,7} & x_{4,8} \\
x_{5,1} & x_{5,2} & x_{5,3} & 0 & 0 & 0 & x_{5,4} & x_{5,7} & x_{5,8}
\end{pmatrix}
\begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 & 0 & 0 & 4 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \\
6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

Figure 1: A proper matrix on the left, and the matrix in the proof of Claim 2.1

have $t_i + t_j \geq k$ for any $i, j \in \{1, 2, 3\}$. Assume for simplicity that we arrange the columns of P as in Figure 1, such that the leftmost columns are from T_1 while the rightmost columns are from T_3. We describe the process for assigning the values in the following 6 stages. In Figure 1, we use $i \in [6]$ in order to denote the 1s that are assigned in the i^{th} stage to some of the variables of the matrix that appears on the left. For brevity, when we say that $x_{i,j}$ is *set*, we mean that we assign it the value 1.

1. Set all the variables on the diagonal that starts at P_{1,t_1+t_2-k+1} and ends at P_{k,t_1+t_2}. In Figure 1, this is the diagonal that starts at $x_{1,2}$ and ends at $x_{5,6}$. Note that we can set these variables as $t_1 + t_2 \geq k$ and as all the variables in the upper k rows are distinct (property 5). As some of the variables in this diagonal may appear in other entries of the matrix we must set them as well. This is done in the next two stages.

2. For every column of type 1 for which we have set a variable in the upper k rows, we must set the corresponding variable in the lower $k-1$ rows. The only exception is column t_1+t_2-k+1 as in this column we have set P_{1,t_1+t_2-k+1}, and it does not appear in the lower $k-1$ rows (recall property 2). In Figure 1 the only variable that is set in this stage is $x_{2,3}$. It follows that the number of variables that are set in this stage is
$$r_1 = k - t_2 - 1. \tag{6}$$

3. There may be columns of type 3 with the same variables as some of the columns of type 2, so we have to set them as well. In Figure 1 the only such case is $x_{3,4}$. It is crucial to observe that the variables that are set in this stage do not belong to any of the rows to which the variables that were set in the previous stage belong. This is due to properties 1 and 7, and the fact that in stage 1 of this process the variables that are set form a diagonal. Denote the number of variables set in this stage by d. We thus get that at the end of this stage, out of the lower $k-1$ rows of the columns of T_3, in
$$r_2 = k - 1 - d \tag{7}$$
rows we have not yet set any variable. Similarly, out of the t_3 columns of T_3, the number of columns in which we have not yet set any variable is
$$r_3 = t_3 - d. \tag{8}$$

4. We now arrive at the main step of the process. Assume the r_1 variables that were set in stage 2 belong to a set of rows denoted by R_1. We claim that we can find r_1 variables that belong to distinct columns of T_3, such that (i) each of these variables belongs to a distinct row of R_1 (ii) None of the other variables of T_3 that were previously set, belongs to the same row or column to which any of these variables belongs. To see this, first observe that by property 7 we have $r_3 > r_1$ which means that we have enough columns in which none of the variables was set to 1. Furthermore, as $t_3 \leq k-1$ we also have $r_2 \geq r_3 \geq r_1$, thus we also have enough rows in which none of the variables was set. We can thus find such a set of r_1 variables. In Figure 1, the only such variable is $x_{2,7}$.

5. Set more variables in the columns of T_3 as long as in the row and column to which they belong none of the variables were set. We can do this as by property 6 all the variables in the lower $k-1$ rows are distinct. In Figure 1, the only such variable is $x_{4,8}$.

6. Out of the lower $k-1$ rows, in $k-1-t_3$ none of the variables were set. Out of the leftmost t_1 columns, in $t_1-(k-t_2)$ none of the variables were set ($k-t_2$ is the number of variables set in the first t_1 columns in stage 1). As $t_1-(k-t_2) = k-1-t_3$, we can find $k-1-t_3$ variables that belong to distinct rows and columns and set them. We can do this as by property 6 all the variables in the lower $k-1$ rows are distinct. In Figure 1, the only variable set in this stage is $x_{5,1}$.

As in Figure 1, the variables that are not set in the above process are assigned the value 0. A key observation now is that the only 1s that appear in a column in which there are other 1s are those from stages 1,2. Similarly, the only 1s that appear in a row in which there are other 1s are those from stages 4,2. However as those from stage 1 are the only 1s in their rows, and those from stage 4 are the only 1s in their columns, the determinant of the matrix thus defined is ± 1. ■

For an $r \times k$ variable matrix M, in which all variables are pairwise distinct, let $\mathcal{P}(M)$ be the following set of $(2k-1)\times(2k-1)$ matrices: For any $1 \leq t_1, t_2, t_3 \leq k-1$ such that $t_1+t_2+t_3 = 2k-1$, we pick 3 sets of rows of M, denoted T_1, T_2, T_3, of sizes t_1, t_2, t_3 respectively that satisfy the following properties (i) $T_1 \cap T_2 = \emptyset$ (ii) $T_1 \cap T_3 = \emptyset$ (iii) $|T_2 \cup T_3| \geq k$. We now use the sets T_1, T_2, T_3 in order to define a matrix P as follows: For every $i \in T_1$ we put the column $(M_{i,1}, M_{i,2}, \ldots, M_{i,k}, M_{i,2}, \ldots, M_{i,k})$. For every $i \in T_2$ we put the column $(M_{i,1}, M_{i,2}, \ldots, M_{i,k}, 0, \ldots, 0)$. For every $i \in T_3$ we put the column $(0, 0, \ldots, 0, M_{i,2}, \ldots, M_{i,k})$.

Claim 2.2 *For any $r \times k$ matrix M, all the matrices in $\mathcal{P}(M)$ are proper. Also, $|\mathcal{P}(M)| \leq r^{2k-1}$*

Proof: Consider any $P \in \mathcal{P}(M)$ defined using the sets of columns T_1, T_2, T_3 of M. The matrix P satisfies the first property of a proper matrix as by definition $1 \leq t_1, t_2, t_3 \leq k-1$. Properties 2,3 and 4 follow from the definition of the matrices in $\mathcal{P}(M)$. Properties 5 and 6 follow from the fact that $T_1 \cap T_2 = \emptyset$ and $T_1 \cap T_3 = \emptyset$, and property 7 follows from the fact that $|T_2 \cup T_3| \geq k$. Finally, the upper bound on $|\mathcal{P}(M)|$ follows from the number of ways to choose $2k-1$ rows from M (possibly, with repetitions) which is clearly an upper bound for the size of $\mathcal{P}(M)$. ■

We now turn to prove the main lemma of this section. In what follows we will also refer to a set $\mathcal{P}(M)$ where M is a matrix with integer values rather than unknown variables. This should be understood as the set $\mathcal{P}(M)$ defined above, where we replace each variable $M_{i,j}$ in each of the matrices in $\mathcal{P}(M)$ with the value assigned to $M_{i,j}$. We need the following lemma of Zippel (c.f., e.g. [15]).

Lemma 2.1 *Let F be an arbitrary field, and let $f = f(x_1, \ldots, x_n)$ be a non-zero polynomial in $F[x_1, \ldots, x_n]$. Suppose the degree of f in each variable is at most d. Then, if S is a subset of F with $|S| > d$, there are at least $(|S| - d)^n$ assignments $x_1 \in S, \ldots, x_n \in S$ so that $f(x_1, \ldots, x_n) \neq 0$.*

Lemma 2.2 *For any $2 \leq k < r$ there is an $r \times k$ matrix M with the following properties:*

1. *All the entries of M are positive integers bounded by r^{2r}.*
2. *Any k rows of M are linearly independent.*
3. *All the matrices in $\mathcal{P}(M)$ are non-singular.*

Proof: Let M be an $r \times k$ variable matrix. We will show that there is an assignment to the rk entries of M that satisfies the three requirements of the lemma. Note, that requiring a certain set of k rows to be linearly independent is equivalent to requiring that a certain multivariate polynomial, namely the determinant of the corresponding matrix, will be non zero. Observe, that as M consists of rk distinct variables, any such polynomial is not identically zero. Similarly, requiring all the matrices in $\mathcal{P}(M)$ to be non-singular is equivalent to requiring that their determinants will be non-zero.

For each set S of k rows of M let f_S be the multivariate polynomial that computes its determinant. Note, that as the variables of M are distinct, f_S is a non-zero polynomial of degree 1 in each variable. Also, for any matrix $P \in \mathcal{P}(M)$ let f_P be the multivariate polynomial that computes its determinant. Recall that by Claim 2.1 the degree of each of these polynomials in each variable is at most 2. Also, for any matrix $P \in \mathcal{P}(M)$ the polynomial f_p is not identically zero. Finally, let F be the product of all these polynomials. As each of the factors of F is of degree at most 2 in each variable, it follows by Claim 2.2 that each of the variables of F has degree at most $2(|\mathcal{P}(M)| + \binom{r}{k}) \leq 2(r^{2k-1} + \binom{r}{k}) < r^{2r}$. In addition, as each of the factors of F is not identically zero, F is also not identically zero. Note, that as each of the requirements (2) and (3) is equivalent to requiring that one of the factors of F is non zero, it is enough to show that there are rk integers bounded by r^{2r}, on which F evaluates to a non-zero integer. Finally, observe that this follows immediately from Lemma 2.1 while working over $\mathbf{R}$ by taking $S_i = \{1, \ldots, r^{2r}\}$ for each of the rk variables of F. ■

We mention that a slightly better dependency on r in the above Lemma can be obtained by using the so called Combinatorial Nullstellensatz [1]. As this will only change the constants hidden in the $o(1)$ term in Theorem 1 we used Lemma 2.1 instead. As we have commented above, the matrix we construct in the above lemma has properties one would expect to find in a random matrix. In fact, one can show that for a large enough prime $p = p(k, r)$, a random $r \times k$ matrix over $GF(p)$ satisfies requirements (2) and (3) of Lemma 2.2, and hence satisfies them over the reals as well.

3 The Construction

In this section we describe the construction of the r-graph which will establish the lower bound of Theorem 1. In what follows, we say that a set $Z \subseteq [n] = \{1, \ldots, n\}$ is *h-sum-free* if for every pair of positive integers $a, b \leq h$ the only solution of the equation

$$az_1 + bz_2 = (a+b)z_3 \tag{9}$$

with $z_1, z_2, z_3 \in Z$ is one in which $z_1 = z_2 = z_3$. Note that a solution of the form $z_1 = z_2 = z_3$ is always a valid solution to equations of this type, hence an h-sum-free set is one that contains no non-trivial solution to equations of this type as long as their coefficients are bounded by h. For our construction we will need the following lemma whose proof, which is based on the construction of Behrend [4], can be found in [9] or [2].

Lemma 3.1 *For every integer h there is a constant $c = c(h)$, such that for every n there is an h-sum-free subset $Z \subset [n]$ of size at least $n/e^{c\sqrt{\log n}}$.*

We turn to define the r-graph H, which will establish the lower bound of Theorem 1. Given integers n and $2 \leq k < r$ let M be an $r \times k$ matrix which satisfies the three assertions of Lemma 2.2. Let Z be an r^{4r^2}-sum-free subset of $[n/r^{3r}]$. By Lemma 3.1, we can find such a set of size at least

$$\frac{n/r^{3r}}{e^{c\sqrt{\log(n/r^{3r})}}} \geq \frac{n}{e^{c'\sqrt{\log n}}} = n^{1-o(1)}, \tag{10}$$

where $c' = c'(r) > 0$. Consider the following definition of an r-graph $H = H(n, k, r, Z, M)$: The vertex set of H consists of r pairwise disjoint sets of vertices $V_1, \ldots, V_r$, where, with a slight abuse of notation, we think of each of these sets as being the set of integers $1, \ldots, n/r$. For every k dimensional vector $z = (z_1, \ldots, z_k) \in Z^k$, we put an edge in H that contains the vertices $v_1 \in V_1, \ldots, v_r \in V_r$ where for $1 \leq i \leq r$ we take v_i to be the integer $(Mz)_i \in V_i$. In what follows we denote by $E(z_1, \ldots, z_k)$, the edge that we put in H when we picked $z = (z_1, \ldots, z_k) \in Z^k$. Note, that we put precisely $|Z|^k$ edges in H and that each of these edges has precisely one vertex in each of the sets $V_1, \ldots, V_r$. Recall that by Lemma 2.2 item (1), the entries of M are integers bounded by r^{2r}. Furthermore, the integers in Z are bounded by n/r^{3r}, hence for every $z \in Z^k$ and $1 \leq i \leq r$, we have $(Mz)_i \leq r \cdot r^{2r} \cdot n/r^{3r} \leq n/r$. Therefore, the vertices "fit" into the sets $V_1, \ldots, V_r$.

Claim 3.1 *Any pair of edges in H share at most $k-1$ vertices. In particular H contains $|Z|^k = n^{k-o(1)}$ edges.*

Proof: It is clearly enough to show that any k vertices of an edge uniquely determine the other $r-k$ vertices of it. Suppose $v_{t_1} \in V_{i_1}, \ldots, v_{t_k} \in V_{t_k}$ are k vertices of the edge $E(z_1, \ldots, z_k)$. Denote $z = (z_1, \ldots, z_k)$, $v = (v_{t_1}, \ldots, v_{t_k})$ and observe that from the definition of H it follows that for $1 \leq i \leq k$ we have $M_{t_i} \cdot z = v_{t_i}$. Let A be the $k \times k$ matrix whose i^{th} row contains the t_i^{th} row of M. We thus get that $Az = v$. As any k rows of M are linearly independent (property 2 in Lemma 2.2), A is invertible. Hence, $z_1, \ldots, z_k$ are uniquely determined by $v_{t_1}, \ldots, v_{t_k}$. In particular, they determine the other vertices of the edge. We thus get that H contains precisely $|Z|^k$ distinct edges. As by (10) $|Z| = n^{1-o(1)}$ the claim follows. ■

In the next section we prove the following lemma, which is the key ingredient in the proof of Theorem 1.

Lemma 3.2 (The Key Lemma) *Suppose we construct $H = H(n, k, r, Z, M)$ as above. If the edges $E(a_1, \ldots, a_k)$, $E(b_1, \ldots, b_k)$ and $E(c_1, \ldots, c_k)$, are spanned by $3(r-k)+k+1$ vertices, and if for some $1 \leq i \leq k$, we have $a_i \leq c_i \leq b_i$ then there are positive integers $\beta_1, \beta_2 \leq r^{4r^2}$ such that*

$$\beta_1 a_i + \beta_2 b_i = (\beta_1 + \beta_2) c_i.$$

The lower bound of Theorem 1 will follow by combining Claim 3.1 and Lemma 3.2.

Proof of Theorem 1: We start with the lower bound. Given n, k and r, construct the r-graph $H = H(n, k, r, Z, M)$ as above. By Claim 3.1 it contains $n^{k-o(1)}$ edges. Suppose indirectly that it contains 3 edges spanned by $3(r-k)+k+1$ vertices and denote these edges by $E(a_1, \ldots, a_k)$, $E(b_1, \ldots, b_k)$ and $E(c_1, \ldots, c_k)$. Consider any $1 \leq i \leq k$ and assume without loss of generality that $a_i \leq c_i \leq b_i$. By Lemma 3.2, there are positive integers $\beta_1, \beta_2 \leq r^{4r^2}$ such that $\beta_1 a_i + \beta_2 b_i = (\beta_1 + \beta_2) c_i$. As Z is r^{4r^2}-sum-free, we have $a_i = b_i = c_i$. As this holds for all $1 \leq i \leq k$, we conclude that E_1, E_2, E_3 were defined using the same set of k integers from Z, which is impossible. This completes the proof of the lower bound.

For the upper bound we use a simple transformation to the upper bound of the (6,3)-problem given in (2). Assume indirectly that for some $2 < k < r$, there is a constant γ and infinitely many integers $n_1, n_2, \ldots$ for which there is an r-graph H_i on n_i vertices with γn_i^k edges and no 3 edges spanned by $3(r-k)+k+1$ vertices. By averaging, each of these r-graphs has $k-2$ vertices $v_i^1, \ldots, v_i^{k-2}$ that belong to at least γn_i^2 of the edges of H_i. We can thus create for each i, an $(r-(k-2)) = (r-k+2)$-graph T_i on n_i vertices that contains all the edges that contain $v_i^1, \ldots, v_i^{k-2}$ in H_i after removing $v_i^1, \ldots, v_i^{k-2}$ from them. It is clear that T_i contains γn_i^2 edges, and that it cannot contain 3 edges spanned by $3(r-k)+k+1-(k-2) = 3(r-k)+3$ vertices. As a consequence, we also conclude that each set of 3 vertices belongs to at most 2 edges, as otherwise 3 edges containing the same 3 vertices are spanned by at most $3 + 3((r-k+2)-3) < 3(r-k)+3$ vertices, which is impossible by the previous argument.

Finally, for each n_i create a 3-graph G_i, where for every edge $e \in T_i$ we put a 3-edge e' in G_i that contains an arbitrary subset of 3 vertices from e. It is easy to see that G_i contains no 3 edges spanned by 6 vertices. Indeed, if e'_1, e'_2, e'_3 are 3 such edges then let e_1, e_2, e_3 be the three edges in T_i that contain the three vertices of these vertices, respectively. We thus get that e_1, e_2, e_3 are 3 edges of T_i spanned by at most $6+3((r-k+2)-3) = 3(r-k)+3$ vertices which contradicts the properties of T_i. As we have previously established that each set of 3 vertices in T_i belongs to at most 2 edges, each G_i contains at least $\gamma n_i^2/2$ edges, hence $f_3(6,3) = \Omega(n^2)$ contradicting (2). ■

4 Proof of the Key Lemma

In this section we give the proof of Lemma 3.2. Let $H = H(n, k, r, Z, M)$ be the r-graph defined as in the previous section. In what follows we denote by a, b, c the vectors $(a_1, \ldots, a_k)$, $(b_1, \ldots, b_k)$, $(c_1, \ldots, c_k)$. We also write M_t for the t^{th} row of M. Suppose H contains 3 edges $E_1 = E(a_1, \ldots, a_k)$, $E_2 = E(b_1, \ldots, b_k)$ and $E_3 = E(c_1, \ldots, c_k)$ spanned by a set T' of $3(r-k)+k+1$ vertices. Remove from T' any vertex that is not contained in any of the edges E_1, E_2, E_3 to obtain a new set T of at most $3(r-k)+k+1$ vertices each of which is contained in at least one of these edges.

Claim 4.1 *For $1 \leq i \leq 3$, there are at least k vertices in T in which E_i intersects another E_j.*

Proof: Assume E_3 intersects either E_1 and/or E_2 in t vertices. Then there are $r-t$ vertices that belong solely to E_3. The edge E_2 contains another set of r vertices. By Claim 3.1, E_1 and E_2 have at most $k-1$ common vertices, thus there are at least $r-k+1$ additional vertices on which E_1, E_2 and E_3 are spanned. This is a total of $3r-k-t+1$ vertices, which is larger than $3(r-k)+k+1$ for any $t<k$. ■

We now arrive at the main step of the proof in which we express the intersections between E_1, E_2 and E_3 as a set of linear equations. Assume that vertex $v_t \in V_t \cap T$ is common to both E_1 and E_2. By the definition of H in Section 3 it follows that $M_t a = M_t b$, or equivalently that

$$a_1 M_{t,1} + a_2 M_{t,2} + a_3 M_{t,3} + \ldots + a_k M_{t,k} = v_t = b_1 M_{t,1} + b_2 M_{t,2} + b_3 M_{t,3} + \ldots + b_k M_{t,k} \tag{11}$$

In what follows we say that edge E_i *belongs* to a linear equation as in (11) if the equation is due to some vertex belonging to E_i and another edge. We will say that an equation *contains* an edge, if the edge belongs to that equation. We will also say that an equation is *due* to vertex v, if the edges that belong to the equation have v in common. In (11), E_1 and E_2 belong to the equation (therefore, it contains them) while E_3 does not, and this equation is due to vertex t. In what follows, it will be more convenient to write (11) as

$$a_1 M_{t,1} + a_2 M_{t,2} + a_3 M_{t,3} + \ldots + a_k M_{t,k} - b_1 M_{t,1} - b_2 M_{t,2} - b_3 M_{t,3} - \ldots - b_k M_{t,k} = 0 \tag{12}$$

Define Φ' to be the set obtained by writing an equation as (12) for each of the vertices of T that lies in $E_i, E_j \in \{E_1, E_2, E_3\}$. If a vertex lies in the three edges E_1, E_2, E_3 we write one equation that contains E_1 and E_3 and another that contains E_2 and E_3.

Comment 1 *It is important for the rest of the proof that the non-symmetry between E_3 and E_1, E_2 caused by this choice is totaly arbitrary, and that we can and will later exchange the roles of, say, E_3 and E_1, if needed.*

Claim 4.2 *The set Φ' contains at least $2k-1$ equations.*

Proof: For each vertex $v \in T$ let d_v be the number of edges out of E_1, E_2, E_3 that contain v (thus $1 \le d_v \le 3$). Note that due to v, the set Φ' contains $d_v - 1$ equations. Thus

$$|\Phi'| = \sum_{v \in T} (d_v - 1) = 3r - |T| \ge 3r - (3(r-k)+k+1) = 2k-1$$

where we used double-counting to get $\sum_v d_v = 3r$ and the fact that $|T| \le 3(r-k)+k+1$. ■

Claim 4.3 *There is $\Phi \subseteq \Phi'$ of size $2k-1$ that satisfies the following properties:*

1. *For $i \in \{1,2\}$, all the equations in Φ that contain E_i are due to distinct vertices.*
2. *Φ contains all the equations that contain E_3, which belonged to Φ'.*
3. *For any two distinct $i, j \in \{1,2,3\}$, the set Φ contains at least one and at most $k-1$ equations that contain E_i and E_j.*

Proof: We first observe that as by Claim 3.1 each pair of edges have at most $k-1$ common vertices, for any i,j, Φ' contains at most $k-1$ equations that contain E_i and E_j. In particular, Φ' contains at most $2k-2$ equations that contain E_3. Hence, we can remove some of the equations that contain both E_1 and E_2 and thus get a set Φ that contains all the equations that contained E_3. This gives (2), and the upper bound of (3). For the lower bound of (3) we again use the fact that there are at most $2k-2$ equations that contain one of the edges, to infer that there is at least one equation that contains the other two. For (1), just observe that by construction of Φ' we put at most one equation in Φ' for each vertex that contains E_1 or E_2, and as $\Phi \subseteq \Phi'$ we get (1). ■

In the rest of the proof we show how to obtain the required linear equation that contains a_1, b_1 and c_1. The other cases are identical. To prove Lemma 3.2, we will simply show that there is a linear combination of the equations of Φ from Claim 4.3, which results in the required linear equation relating a_1, b_1, c_1. We will call such a linear combination *good*. Denote the linear equations of Φ by $\ell_1, \dots, \ell_{2k-1}$. In order to get a good linear combination, we introduce unknowns $\alpha_1, \dots, \alpha_{2k-1}$, where α_i will be the coefficient of ℓ_i. For $1 \le i \le k$ let A_i be the homogenous linear equation in unknowns $\alpha_1, \dots, \alpha_{2k-1}$ that requires the coefficient of a_i to vanish in a linear combination of $\ell_1, \dots, \ell_{2k-1}$ with coefficients $\alpha_1, \dots, \alpha_{2k-1}$. For $1 \le i \le k$ define B_i and C_i to be the analogous equations with respect to b_i and c_i.

Claim 4.4 *For $1 \le i \le k$ we have $A_i + B_i + C_i = 0$.*

Proof: Just observe that the coefficient of α_j in C_i is the coefficient of c_i in the j^{th} equation of Φ. The same applies for A_i and B_i. For example, if (12) is equation j in Φ than for $1 \le i \le k$ the coefficient of α_j in C_i is 0 because E_3 does not belong to this equation. Also for $1 \le i \le k$ the coefficient of α_j in A_i is $M_{t,i}$ and the coefficient of α_j in B_i is $-M_{t,i}$. Given these observations the claim is trivial. ■

In order to get the required equation in Lemma 3.2 the coefficients of the integers $a_2, \dots, a_k$, $b_2, \dots, b_k$ and $c_2, \dots, c_k$ must vanish. This amounts to a set of $3k-3$ homogenous linear equations A_i, B_i, C_i for $2 \le i \le k$ defined above. However, by Claim 4.4 we may remove equations $C_2, \dots, C_k$ and thus get a set of $2k-2$ linear equations. Call this set Ψ. We will need the following well known result which follows from Cramer's rule and Hadamard Inequality (see, e.g., [13]).

Lemma 4.1 *Let Ψ be a set of p homogenous linear equations in q variables with integer coefficients. If $p < q$ and each of the coefficients in these equations has absolute value at most d, then Ψ has a* **non zero** *solution $\{\alpha_1, \dots, \alpha_q\}$, where all the α_is are* **integers** *with absolute value at most $(d^2p)^{p/2}$.*

As Ψ is a set of $2k-2$ equations in $2k-1$ unknowns $\alpha_1, \dots, \alpha_{2k-1}$ and each of the coefficients in Ψ is bounded by r^{2r} (recall that these coefficients are entries of M, see (12) and Lemma 2.2 item (1)), we get from the above lemma that:

Claim 4.5 *There are integers $\alpha_1, \dots, \alpha_{2k-1} \le r^{2r^2}$, not all equal to zero, such that in a linear combination of Φ with coefficients $\alpha_1, \dots, \alpha_{2k-1}$, for $2 \le i \le k$ the coefficients of a_i, b_i, c_i vanish.*

Note that the sum of the coefficients in each of the equations of Φ is zero (see (12)). Hence, the sum of the coefficients in a linear combination of these equations must also be zero. It follows that if for $2 \le i \le k$ the coefficients of a_i, b_i, c_i vanish while the coefficients of a_1, b_1, c_1 do not, then we get the required equation. Claim 4.5 almost guarantees the existence of a good linear combination. It guarantees that the coefficients of a_1, b_1, c_1 are integers bounded by $(2k-1)r^{2r^2} \le r^{4r^2}$, and that the coefficient of all the other a_i, b_i, c_i vanish. The only thing that can go wrong is that the coefficients of a_1, b_1, c_1 will also vanish. To argue that this is impossible we will first show that the coefficients of a_1 and b_1 do not vanish. To this end, we show that the equations of Ψ and any one of the equations A_1, B_1 are linearly independent. As we chose a non-zero vector of coefficient in Claim 4.5, it cannot satisfy $2k-1$ linearly independent homogenous linear equations in $2k-1$ unknowns. This will immediately imply that the coefficients of a_1 and b_1 do not vanish.

Claim 4.6 *The set Ψ with either A_1 or B_1 is a set of linearly independent linear equations.*

Proof: Consider the matrix P whose upper k columns are the coefficients of $\alpha_1, \dots, \alpha_{2k-1}$ in equations $A_1, \dots, A_k$ and whose lower $k-1$ columns are the coefficients of $\alpha_1, \dots, \alpha_{2k-1}$ in equations $B_2, \dots, B_k$. As by Lemma 2.2 item (3) all the matrices in $\mathcal{P}(M)$ are non-singular, it is enough to show that $P \in \mathcal{P}(M)$. For a column vector v of P, denote by v^a the k dimensional vector that contains the upper k entries of v and by v^b the $k-1$ dimensional vector that contains the lower $k-1$ entries of v. Observe that if v is the j^{th} column of P, then v^a and v^b contain the coefficients of $a_1, \dots, a_k$ and $b_2, \dots, b_k$ respectively in equation ℓ_j. Note further, that if E_1 does not appear in ℓ_j then $v^a = 0$ and if it does, then $v^a = (M_{t,1}, M_{t,2}, \dots, M_{t,k})$ where V_t is the cluster in which E_1 intersects another edge (recall the definition of $H(n, k, r, Z, M)$ in Section 3). Similarly, either $v^b = 0$ or $v^b = (M_{t,2}, \dots, M_{t,k})$. This means that there are three types of columns: (i) Columns v that correspond to equations that contain both E_1 and E_2. In these columns $v^a \neq 0$ and $v^b \neq 0$. Moreover, observe that in these columns the entries of v^b are precisely the last $k-1$ entries of v^a. (ii) Columns that correspond to equations that contain both E_1 and E_3. In these columns $v^a \neq 0$ while $v^b = 0$. (iii) Columns that correspond to equations that contain both E_2 and E_3. In these columns $v^a = 0$ while $v^b \neq 0$. Denote by t_1, t_2, t_3 the number of columns of type (i),(ii) and (iii) respectively. We claim that the columns of types (i),(ii),(iii) can play the role of the sets of columns T_1, T_2, T_3 in the definition of a proper matrix. Indeed, by the above discussion they satisfy properties 2,3,4. By Claim 4.3 item (3) we get that $1 \le t_1, t_2, t_3 \le k-1$, hence property 1 is also satisfied. Properties 5,6 follow from Claim 4.3 item (1). Finally, from Claim 4.1 and Claim 4.3 item (2) we get property 7. We conclude that $P \in \mathcal{P}(M)$ as needed. The proof for Ψ and B_1 is identical where we replace A_1 in the above argument by B_1. ■

Proof of Lemma 3.2: As we have commented above, all the cases $1 \le i \le k$ are identical, thus we prove the case $i = 1$. By Claim 4.5 we can find a linear combination of the equations of Φ in which for $2 \le i \le k$ the coefficients of a_i, b_i, c_i vanish. By Claim 4.6 the coefficients of a_1 and b_1 do not vanish in such a linear combination. If the coefficient of c_1 also does not vanish we are done. By the discussion preceding the proof of Claim 4.6 we conclude that if it does, then the coefficient of a_1 must be equal to the inverse of the coefficient of b_1, thus $a_1 = b_1$. In this stage we can rerun the argument of this section while exchanging the roles of E_1 and E_3 (recall Comment 1). We will thus either get the required equation, or that $b_1 = c_1$. In the former case the lemma will follow, while in the latter we will get that $a_1 = b_1 = c_1$ (thus they satisfy the equation $a_1 + b_1 = 2c_1$). In either case we get the required equation. ■

5 Concluding Remarks and Open Problems

- Given the previous results and the results of this paper, the following conjecture seems plausible:

  Conjecture 1 *For every fixed $2 \leq k < r$ and $3 \leq e$ we have*

  $$n^{k-o(1)} < f_r(n, e(r-k)+k+1, e) = o(n^k). \tag{13}$$

  It will be very interesting to extend our construction for arbitrary number of edges and thus prove the lower bound of (13). Recall that one of the main ingredients of the construction was a dense set of integers that contains no non-trivial solution to equations of the form $az_1 + bz_2 = (a+b)z_3$ where a, b are small integral constants. As the proof of Theorem 1 suggests, in order to extend the construction for arbitrary e we will have to use a dense set of integers (i.e, one of size $n^{1-o(1)}$), which contains no non-trivial solution to equations of the form

  $$a_1z_1 + \ldots + a_{e-1}z_{e-1} = (a_1 + \ldots + a_{e-1})z_e. \tag{14}$$

  However, we can only construct dense sets which contain no non-trivial solution to equations of the above type as long as $a_1, \ldots, a_{e-1}$ are positive. In fact, it is easy to see that the largest subset of the first n integers without a solution to the equation $z_1 + z_2 - z_3 = z_4$ is $O(\sqrt{n})$. Note, that for three edges we do not have to worry about the sign of the coefficients as we can always "switch sides" in order to get an equation with positive coefficients. It thus follows that the only (natural) way to extend our technique to arbitrary number of edges is to extend Lemma 3.2 by showing that given e edges spanned by $e(r-k)+k+1$ vertices we can find a linear combination as in (14) with positive coefficients. This seems to be a hard task. See [3] for a solution of a similar problem. See also [17] for some constructions that may be relevant.

- Though we are currently unable to extend our lower bounds to arbitrary number of edges, in some settings we can obtain lower bounds for more than 3 edges.

  Proposition 5.1 *Suppose that for some integers $e \geq 3$, $k \geq 2$ and $r = k+1$ we have*

  1. *$n^{k-o(1)} < f_r(n, (e-1)(r-k)+k+1, e-1)$.*
  2. *$e/\lceil (e(r-k)+k+1)/r \rceil < 2$.*

  then we also have $n^{k-o(1)} < f_r(n, e(r-k)+k+1, e)$.

  Proof: By item (1) there are infinitely many integers n_i for which there is an r-graph H_i on n_i vertices with $n_i^{k-o(1)}$ edges that contains no $e-1$ edges spanned by $(e-1)(r-k)+k+1$ vertices. We may clearly assume that these r-graphs are r-partite as it is easy and well known that every r-graph with $|E|$ edges contains an r-partite subgraph with at least $r!|E|/r^r$ edges. (See, e.g., [15], page 67).

  We claim that this family of r-graphs establishes that $n^{k-o(1)} < f_r(n, e(r-k)+k+1, e)$. Indeed, suppose one of these r-graphs H_i contains e edges spanned by $e(r-k)+k+1$ vertices.

By item (2), this set contains a vertex that belongs to at most one edge. Removing this vertex and the edge to which it belongs, we get a set of $e(r-k)+k+1-1 = (e-1)(r-k)+k+1$ vertices (recall that $r-k=1$) that span at least $e-1$ edges. This contradicts our assumption on H_i. ■

Using this proposition with $r=3$, $k=2$, $e=4$ and the fact that $n^{2-o(1)} < f_3(n,6,3)$ one immediately gets that $n^{2-o(1)} < f_3(n,7,4)$. A similar estimate was mentioned (without proof) in [18]. Reusing the above proposition with $r=3$, $k=2$, $e=5$ and the fact that $n^{2-o(1)} < f_3(n,7,4)$ we get that $n^{2-o(1)} < f_3(n,8,5)$. Several other lower bounds can be obtained using this process, but provide no new cases of equality in the left side of (13).

- It will also be very interesting to prove the upper bound of (13) for arbitrary value of e. As we observe below, to this end it is enough to resolve only the cases of $k=2$.

Proposition 5.2 *If for any e and $3 \le r$ we have $f_r(n, e(r-2)+3, e) = o(n^2)$ then for any e and $2 \le k < r$ we have $f_r(n, e(r-k)+k+1, e) = o(n^k)$.*

Proof: Assume indirectly that for some $e \ge 3$ and $2 < k < r$, there is a constant γ and infinitely many integers $n_1, n_2, \ldots$ for which there is an r-graph H_i on n_i vertices with γn_i^k edges and no e edges spanned by $e(r-k)+k+1$ vertices. By averaging, each of these r-graphs has $k-2$ vertices $v_i^1, \ldots, v_i^{k-2}$ that belong to at least γn_i^2 of the edges of H_i. We can thus create for each i, an $(r-(k-2))$-graph T_i on n_i vertices that contains all the edges that contain $v_i^1, \ldots, v_i^{k-2}$ in H_i after removing $v_i^1, \ldots, v_i^{k-2}$ from them. It is clear that T_i contains γn_i^2 edges. Moreover, it is easy to see that it cannot contain e edges spanned by $e(r-k)+k+1-(k-2) = e((r-k+2)-2)+3$ vertices. This implies that $f_{r-k+2}(n, e((r-k+2)-2)+3, e) = \Omega(n^2)$, which contradicts our initial assumption. ■

- In [18], Ruzsa and Szemerédi used ideas similar to the ones used to resolve the (6,3)-problem in order to construct graphs that contain $\Theta(n)$ induced matchings each of size $n^{1-o(1)}$. It will be very interesting to estimate the maximum possible number of induced matchings of size $n^{1-o(1)}$ in a k-graph on n vertices.

- The $(6,3)$-problem (under different disguises) has found many applications in extremal combinatorics, some examples of which are [12] and [7]. It has also found applications in theoretical computer science. Some examples are PCP analysis and Linearity Testing [14], Communication Complexity [16] and Monotonicity Testing [10]. It may be interesting to find similar applications of our extension of the $(6,3)$-problem.

References

[1] N. Alon, Combinatorial Nullstellensatz, Combinatorics, Probability and Computing 8 (1999), 7-29.

[2] N. Alon and A. Shapira, A characterization of easily testable induced subgraphs, Proc. of the 15^{th} Annual ACM-SIAM SODA, ACM Press (2004), 935-944.

[3] N. Alon and A. Shapira, Linear equations, arithmetic progressions and hypergraph property-testing, submitted.

[4] F. A. Behrend, On sets of integers which contain no three terms in arithmetic progression, *Proc. National Academy of Sciences USA* 32 (1946), 331–332.

[5] W. G. Brown, P. Erdős and V.T. Sós, Some extremal problems on r-graphs, New Directions in the Theory of Graphs, Proc. 3^{rd} Ann Arbor Conference on Graph Theorey, Academic Press, New York, 1973, 55-63.

[6] W. G. Brown, P. Erdős and V.T. Sós, On the existence of triangulated spheres in 3-graphs and related problems, Periodica Mathematica Hungaria, 3 (1973), 221-228.

[7] L. H. Clark, R. C. Entringer, J. E. McCanna and L. Székely, Extremal problems for local properties of graphs. Combinatorial Mathematics and Combinatorial Computing (Palmerston North, 1990). Australas. J. Combin. 4 (1991), 25–31.

[8] 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.

[9] P. Erdős, P. Frankl and V. Rödl, The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent, Graphs Combin. 2 (1986) 113-121.

[10] E. Fischer, E. Lehman, I. Newman, S. Raskhodnikova, R. Rubinfeld and A. Samorodnitsky, Monotonicity testing over general poset domains, Proc. of The 34^{th} STOC (2002), 474-483.

[11] Z. Füredi, Turán-type problems, Surveys in Combinatorics, London Math. Soc. Lecture Ser., A.D. Keedwell, Ed., Cambridge Univ. Press (1991), 253-300.

[12] Z. Füredi, The maximum number of edges in a minimal graph of diameter 2. J. Graph Theory 16 (1992), no. 1, 81–98.

[13] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, 6th Edition, Academic Press (2000), p. 1110.

[14] J. Håstad, and A. Wigderson, Simple analysis of graph tests for linearity and PCP. IEEE Conference on Computational Complexity, 2001, 244-254.

[15] S. Jukna, *Extremal Combinatorics With Applications in Computer Science* 2001, Springer-Verlag.

[16] P. Pudlák and J. Sgall, An upper bound for a communication game related to time-space tradeoffs. The mathematics of Paul Erdős, I, (R. L. Graham and J. Nešetřil, eds.), Algorithms Combin., 13, Springer, Berlin, (1997), 393-399.

[17] I. Z. Ruzsa, Solving a linear equation in a set of integers I, Acta Arithmetica 65 (1993), 259-282.

[18] I. Z. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles, in Combinatorics (Keszthely, 1976), Coll. Math. Soc. J. Bolyai 18, Volume II, 939-945.

[19] G. N. Sárközy and S. M. Selkow, On a Turán-type hypegraph problem of Brown, Erdős and T. Sós, manuscript, 2003.

[20] A. F. Sidorenko, What we do know and what we do not know about Turán numbers, Graphs and Combinatorics, 11 (1995), 179-199.