RADO PARTITION THEOREM FOR RANDOM SUBSETS OF INTEGERS

Vojtech Rödl and Andrzej Ruciński

Department of Mathematics and Computer Science
Emory University, Atlanta GA 30322
email: rodl@mathcs.emory.edu, andrzej@mathcs.emory.edu

Abstract. For an $l \times k$ matrix $A = (a_{ij})$ of integers, denote by $\mathcal{L}(A)$ the system of homogenous linear equations $a_{i1}x_1 + a_{ik}x_k = 0$, $1 \leq i \leq l$. We say that A is *density regular* if every subset of $\mathbf{N}$ with positive density, contains a solution to $\mathcal{L}(A)$. For a density regular $l \times k$ matrix A, an integer r and a set of integers F, we write

$$F \to (A)_r$$

if for any partition $F = F_1 \cup .... \cup F_r$ there exists $i \in \{1, 2, ..., r\}$ and a column vector $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$ and all entries of $\mathbf{x}$ belong to F_i. Let $[n]_N$ be a random N-element subset of $\{1, 2, ..., n\}$ chosen uniformly from among all such subsets. In this paper we determine for every density regular matrix A a parameter $\alpha = \alpha(A)$ such that $\lim_{n\to\infty} \mathbf{P}([n]_N \to (A)_r) = 0$ if $N = O(n^\alpha)$ and 1 if $N = \Omega(n^\alpha)$.

1. Introduction

Partition theorems for sets of integers constitute an important part of Ramsey theory. They typically embody the following pattern: for every finite coloring of the set of integers $\mathbf{N}$, there exist $x_1, ..., x_n \in \mathbf{N}$ of the same color, which satisfy some prescribed conditions. The first theorem of this type was proved by I. Schur in 1916.

Theorem [8]. *If* $\mathbf{N}$ *is finitely colored then there exist* x, y, z *having the same color such that* $x + y = z$.

In 1927 B.L. van der Waerden published a proof of the following result.

Research supported by NSF grant INT-940671. In addition, the first author supported by NSF grant DMS-9401559 and the second author supported by KBN grant 2 P03A 023 09. The second author on leave from Adam Mickiewicz University, Poznań, Poland.

Theorem [10]. *If* $\mathbf{N}$ *is finitely colored then one of the color classes contains arbitrarily long arithmetic progressions.*

As arithmetic progressions of length k correspond to distinct-valued solutions of the system

$$x_1 - x_2 = x_2 - x_3 = ... = x_{k-1} - x_k \ , \tag{1.1}$$

we see that configurations considered in both, Schur's and van der Waerden's theorems correspond to integer solutions of systems of homogenous linear equations.

In 1930 R.Rado published a far reaching generalization of both these results.

For an $l \times k$ matrix $A = (a_{ij})$ of integers, denote by $\mathcal{L} = \mathcal{L}(A)$ the system of homogenous linear equations

$$\sum_{j=1}^{k} a_{ij}x_j = 0, \qquad 1 \le i \le l \ .$$

We say that $\mathcal{L}$ (and also A) is *partition regular* if for any finite coloring of $\mathbf{N}$, there is always a solution to $\mathcal{L}$ with all x_i having the same color.

The matrix A is said to satisfy *the column condition* if it is possible to re-order the column vectors $\mathbf{a_1}$,...,$\mathbf{a_k}$, so that for some choice of indices $0 = k_0 < k_1 < ...k_t = k$, setting $\mathbf{b_i} = \sum_{j=k_{i-1}+1}^{k_i} \mathbf{a_j}$, $i = 1, ..., t$,

(i) $\mathbf{b_1} = \mathbf{0}$

and

(ii) for each $i = 1, ..., t$, $\mathbf{b_i}$ can be expressed as a rational linear combination of $\mathbf{a_j}$, $1 \le j \le k_{i-1}$.

The classical result of Rado asserts the following:

Theorem [6]. *The system* $\mathcal{L}(A)$ *is partition regular if, and only if* A *satisfies the column condition.*

Note that a matrix A consisting of a single row satisfies the column condition if, and only if the equation $\mathcal{L}(A)$ has a nontrivial 0-1 solution. Indeed, then condition (ii) is trivially fulfilled, while (i) is also true if $\mathbf{b_1}$ is the sum of the coefficients corresponding to the 1's in this solution and $\mathbf{b_2}$ is the sum of all the other coefficients. Thus, the simplest example of a non-partition equation is $x + y = 3z$.

Without loss of generality we may clearly assume that matrix A is of full rank, i.e. its rank, denoted here by $h(A)$, satisfies the equation

$$h(A) = l \; . \tag{1.2}$$

This assumption will be imposed throughout the paper.

Moreover, if for some $i \neq j$ all solutions of $A\mathbf{x} = \mathbf{0}$ satisfy $x_i = x_j$ then this system is equivalent to another system $A'\mathbf{x} = \mathbf{0}$ which contains the equation $x_i = x_j$. (Indeed, assuming without loss of generality, that $i = k-1$ and $j = k$, and applying Gaussian elimination to A, we obtain another matrix $A' = (a'_{ij})$. If for some $j < k-1$, $a'_{kj} \neq 0$, then we would be free to choose a solution with $x_{k-1} \neq x_k$, which is a contradiction.) Eliminating this equation together with one of the variables x_i or x_j, leads to a system of $l-1$ equations and $k-1$ variables, the solutions of which, viewed as set of integers rather than vectors, coincide with the solutions of the original system. Therefore, in what follows we will be, without loss of generality, assuming that the system $A\mathbf{x} = \mathbf{0}$ has, for each $i \neq j$, a solution $\mathbf{x} = (x_1, ..., x_k)$ such that $x_i \neq x_j$. Systems (and matrices) satisfying the above condition will be called *irredundant.*

The value of k is at least $l+1$, since otherwise there would not be any positive solution. If, however, $k = l+1$, then, after Gaussian elimination, the last equation involves x_{k-1} and x_k only, but, by the irredundancy, it cannot be of the form $x_{k-1} = x_k$. Then, however it does not satisfy the column condition and, so, it is not partition regular. Hence, we always have $k \geq l+2$.

A natural question in connection with the van der Waerden theorem was raised by Erdős and Turán [2] more than 60 years ago. Assume that $F \subseteq \mathbf{N}$ is a set having positive upper density, i.e.

$$\bar{d}(F) := \lim\sup_{n\to\infty} \frac{|F \cap [n]|}{n} > 0 \; .$$

Does F have to contain an arithmetic progression of arbitrary length? In 1975 Szemerédi settled this conjecture in the affirmative proving one of his great results.

Theorem [9]. *If* $F \subseteq \mathbf{N}$ *satisfies* $\bar{d}(F) > 0$ *then* F *contains arbitrarily long arithmetic progressions.*

This is a powerful strengthening of van der Waerden's theorem implying that for any finite partition of $\mathbf{N}$ the "biggest" color class always contains arbitrarily long arithmetic progressions. Analogous strengthening of Schur's theorem is clearly not valid, and therefore it is natural to ask which partition regular systems admit the "density" versions. Such systems were characterized in [3].

The system $\mathcal{L}(A)$ (and matrix A) is said to be *density regular* if every set of integers F of positive density $\bar{d}(F) > 0$ contains a distinct-valued solution of $\mathcal{L}(A)$.

Theorem [3]. *An irredundant system* $\mathcal{L}(A)$ *is density regular if, and only if it has a solution* $x_1 = ... = x_k = 1$.

The simplest example of a density regular system is the system (1.1) corresponding to arithmetic progressions of length k.

We will find it convenient to utilize the following "arrow notation" introduced by Erdős and Rado. For a partition regular $l \times k$ matrix A, an integer r and a set of integers F, we write

$$F \to (A)_r$$

if for any partition $F = F_1 \cup .... \cup F_r$ there exists $i \in \{1, 2, ..., r\}$ and a column vector $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$ and all entries of $\mathbf{x}$ belong to F_i.

Fixing A and r, observe that the property $F \to (A)_r$ is monotone, i.e. it implies $F' \to (A)_r$ for any $F' \supset F$. Also, while trivially $\emptyset \to (A)_r$ is not true for any A, an easy compactness argument, together with Rado,s theorem, shows that $[n] \to (A)_r$ holds for $n \geq n(A, r)$, where $n(A, r)$ is a constant depending on A and r.

In this paper we study the question of for which $N = N(n)$, $1 \leq N \leq n$, almost all N-element subsets $F \subset [n]$ have the property $F \to (A)_r$. More formally, let $[n]_N$ be a random N-element subset of $[n]$ chosen uniformly from among all $\binom{n}{N}$ such subsets of $[n]$. The aim of this paper is to determine a threshold function $N_0 = N_0(A, r, n)$ such that

$$\lim_{n\to\infty} \mathbf{P}([n]_N \to (A)_r) = \begin{cases} 0 \text{ if } N = O(N_0) \\ 1 \text{ if } N = \Omega(N_0) \end{cases},$$

for any density regular matrix A.

In order to formulate the result of this paper we need to define a matrix parameter, measuring, in some sense, a relative degree of freedom of the system $\mathcal{L}(A)$.

Let Q be a subset of the set of columns of A. (We will be identifying these columns with their indices $1, 2, ..., k$.) We denote by A^Q the matrix obtained by deleting all columns belonging to Q.

Let $h(A)$ stand for the rank of matrix A. We will use the abbreviation $h_Q = h(A^Q)$. We set $h_Q = 0$ if $|Q| = k$.

Definition 1.1 For an irredundant, partition regular matrix A with $h(A) = l$ let

$$m_A = \max_{2\le q\le k} \max_{Q\subseteq[k],|Q|=q} \frac{q-1}{q-1+h_Q-l}.$$

It will be proved in the next section that $q + h_Q - l \geq 2$ for every $q \geq 2$.

To better accommodate this new concept, suppose for a moment that, in addition to all previous assumptions, every $l \times l$ submatrix of A has rank l.

Then, if $q \leq k - l$, $h_Q = l$, while for $q > k - l$, $h_Q = k - q$. Thus, under this additional assumption, the parameter m_A becomes

$$m_A = \max\left\{1, \frac{k-1}{k-l-1}\right\} = \frac{k-1}{k-l-1}. \tag{1.3}$$

This corresponds to the case when the maximum in Definition 1.1 is achieved by taking $Q = [k]$.

The additional condition is, for example, satisfied by the matrix A corresponding to the system (1.1), the solutions of which are arithmetic progressions of length k.

We are now ready to formulate our result.

Theorem 1.1. *For every $l\times k$ irredundant, density regular matrix A of rank $h(A) = l$ and for every integer $r \geq 2$, there are constants c and C such that*

$$\lim_{n\to\infty} \mathbf{P}([n]_N \to (A)_r) = \begin{cases} 0 & \text{if } N < cn^{1-1/m_A} \\ 1 & \text{if } N > Cn^{1-1/m_A} \end{cases}.$$

We will prove the “0-statement” for all partition regular matrices. However, as far as the “1-statement” is concerned, it is only proved for density regular matrices.

For matrices corresponding to the van der Waerden theorem, i.e. matrices A such that the solutions to $A\mathbf{x} = \mathbf{0}$ are arithmetic progressions of length k, this theorem was stated in [7] with only an outline of proof. Let $F \to (k)_r$ stand for the property that for every partition of a set F into r classes, at least one of the classes contains an arithmetic progression of length k. As a special case of Theorem 1.1 we obtain

Theorem [7]. *For all $k \geq 3$ and $r \geq 2$ there exist constants c and C such that*

$$\lim_{n\to\infty} \mathbf{P}([n]_N \to (k)_r) = \begin{cases} 0 \text{ if } N \leq cn^{\frac{k-2}{k-1}} \\ 1 \text{ if } N \geq Cn^{\frac{k-2}{k-1}} \end{cases} .$$

We strongly believe that the following is true:

Conjecture. *The “1-statement” of Theorem 1.1 remains true for any partition regular system $A\mathbf{x} = \mathbf{0}$.*

Unfortunately, our method does not allow for such a generalization. A crucial ingredient of our argument is Theorem B (cf. Section 2), valid only for density regular systems of equations. The main difficulty we face is the lack of its analog for an arbitrary partition regular system. So far we have failed to find a way around this obstacle.

Note that the proof of the “0-statement” of Theorem 1.1 given in Section 7 works for all partition regular matrices.

The simplest case of a partition regular system which is not covered by our Theorem 1.1 is the system consisting of a single equation $x + y - z = 0$. This case, which corresponds to Schur’s Theorem, is investigated in our forthcoming paper with R.L. Graham [4]. The result proved there supports our conjecture.

The paper is organized as follows. In the next section some tools from linear algebra, partition theory and probability theory are collected. Section 3 contains an outline of the proof of the “1-statement” of our theorem, where it is reduced to a statement $S(r, s)$. This statement is proved by double induction in Sections 4-6. Finally, in Section 7 we prove the "0-statement".

2. PRELIMINARIES

In this section we review and introduce some tools that will be utilized in the proof. We begin with some elementary algebraic properties of systems of linear equations. For a matrix A let $h(A)$ denote the rank of A. The following elementary result can be derived in many ways.

Proposition 2.1. *For an arbitrary integer $l \times k$ matrix A and an l-dimensional integer vector $\mathbf{b}$, the system of linear equations $A\mathbf{x} = \mathbf{b}$ has no more than $n^{k-h(A)}$ integer solutions with all entries belonging to the set $[n] = \{1, 2, ..., n\}$.* $\square$

As for the lower bound, in general, there may be no solution at all. Since we are actually interested in partition regular systems, we restrict our attention to them, and derive a matching lower bound from a partition result. The following strengthening of Rado's Theorem was proved in [3].

Theorem A [3]. *Let A be an $l \times k$ matrix of rank l which satisfies the Rado columns condition. Then for any r there exists $c_r(A)$ such that in any r-coloring of $[n]$ there are at least $c_r(A)n^{k-l}$ monochromatic solutions to the system $A\mathbf{x} = \mathbf{0}$.* $\square$

In particular, the above result implies that there are at least $c_1(A)n^{k-l}$ solutions of $A\mathbf{x} = \mathbf{0}$ in $[n]$, which, together with Proposition 2.1, yields that the number of such solutions is, indeed, of the order of n^{k-l}.

For an $l \times k$ matrix A with columns $\mathbf{c_1}$, $\mathbf{c_2}$,...,$\mathbf{c_k}$ and $Q \subseteq [k]$, let A^Q be the matrix with columns $\mathbf{c_i}$, $i \in [k] \setminus Q$.

Proposition 2.2. *Let A be an $l \times k$ irredundant, partition regular matrix of rank l. Then*

(i) For every Q, $|Q| = 1$, we have $h(A^Q) = l$;

(ii) For every Q, $|Q| = q \geq 2$, the inequality $l - h(A^Q) + 2 \leq q$ holds.

Note: Part (ii) of the above proposition verifies the correctness of Definition 1.1.

Proof.

(i) Suppose that $h(A^Q) = l - 1$ for some Q, $|Q| = 1$. Without loss of generality assume that $Q = \{k\}$ and hence A^Q consists of columns $\mathbf{c_1}$,...,$\mathbf{c_{k-1}}$. Apply Gaussian elimination to A to obtain a matrix $A' = (a'_{ij})$. Since $h(A^Q) < l$, $a'_{lj} = 0$ for $j = 1, 2, ..., k - 1$. On the other hand, due to the fact that A has full rank, we have $a'_{lk} \neq 0$. This, however, means that all solutions $\mathbf{x} = (x_1, ..., x_k)$ of the system $A'\mathbf{x} = \mathbf{0}$ (which is equivalent to $A\mathbf{x} = \mathbf{0}$) satisfy $x_k = 0$, and thus there is no positive solution. This is a sharp contradiction with Theorem A.

(ii) We will proceed by induction on q. Assume first that there is Q, $|Q| = 2$, with $h(A^Q) < l$. Without loss of generality assume that $Q = \{k - 1, k\}$. Using an argument similar to that used in part (i), we infer that every solution to $A\mathbf{x} = \mathbf{0}$ satisfies either $x_{k-1} = 0$ or $x_k = 0$ or $x_{k-1} = \alpha x_k$, where α is a positive rational number. The first two cases contradict, as before, the existence of positive solutions ensured by Theorem A. In the third case, due to the fact that A is irredundant, α must be different from 1. Then, however, the single equation $x_{k-1} - \alpha x_k = 0$ is not partition regular (has no nontrivial 0-1 solution), contradicting our assumption.

Assume now that $q \geq 3$ and that the statement (ii) is true for $q - 1$. As the rank of a matrix can drop by at most one when a column is deleted, the required inequality follows by induction. ☐

Finally, we state a result which is an important tool in our proof. This is a strengthening of the Schur Theorem stated in Introduction.

We call a solution of $A\mathbf{x} = \mathbf{0}$, $\mathbf{x} = (x_1, ..., x_k)$, *all-distinct* if all the entries x_i are distinct, $i = 1, ..., k$. It was shown in [3] that an irredundant, density regular system always contains an all-distinct solution. In fact, much more was proved there, and the theorem below will play a crucial role in our proof.

Theorem B [3]. *Let A be an $l \times k$ matrix of rank l which is irredundant and density regular. Then for every $\epsilon > 0$, there are constants $\sigma = \sigma(A, \epsilon)$ and $n_0 = n_0(A, \epsilon)$*

such that if $n > n_0$ *and* $X \subseteq [n]$, $|X| > \epsilon n$, *then the set* X *contains at least* σn^{k-l} *all-distinct solutions* $\mathbf{x}$ *to* $A\mathbf{x} = \mathbf{0}$. □

Let X be a random variable with binomial distribution with expectation np. The following bound is provided by Chernoff's inequality:

$$\mathbf{P}(|X - np| > \epsilon np) < \exp\{-C(\epsilon)np\} ,$$

where $C(\epsilon)$ is a positive constant.

In a few places in our proof we need exponentially small bounds on tails of sums of not necessarily independent random variables. If the dependence is relatively weak, the bounds for the lower tail are provided by Janson's inequality.

Throughout the paper $\mathbf{E}(X)$ stands for the expectation of a random variable X. Let F be a finite set, from which a subset is drawn randomly in such a way that the inclusions of individual elements are independent. Further, let $\mathcal{S}$ be a family of subsets of F and for each $B \in \mathcal{S}$ let I_B be equal to 1 if B is entirely included into the random subset and 0 otherwise. Finally, let $X = \sum_{B \in \mathcal{S}} I_B$. Then

Lemma 2.1 [5]**.** *For every* $0 < \epsilon \leq 1$

$$\mathbf{P}(X \leq (1-\epsilon)\mathbf{E}(X)) \leq \exp\left\{ -\frac{(\epsilon \mathbf{E}(X))^2}{2\sum\sum_{B_1 \cap B_2 \neq \emptyset} \mathbf{E}(I_{B_1} I_{B_2})} \right\} \qquad \square$$

Unfortunately, the upper tail counterpart of Lemma 2.1 is not true in general. As an exponential bound is often needed also for the upper tail, to cope with this situation we developed in [7] an approach based on the following elementary lemma which deals with a somewhat simplified case when all elements are included to the random set with the same probability p and all members of $\mathcal{S}$ are of the same size s.

Lemma 2.2 [7]**.** *Let* F *be a finite set and* $\mathcal{S}$ – *a family of* s*-element subsets of* F. *For* $0 < p < 1$, *let* F_p *be a random subset of* F *obtained by independent inclusion of each element with probability* p. *Then, for any integer* t, *with probability at least* $1 - 2^{-t/s}$, *there exists a set* $E \subset F_p$ *of size* t *such that* $F_p \setminus E$ *contains at most* $2|\mathcal{S}|p^s$ *sets from* $\mathcal{S}$. □

Hence, exceeding two times the expectation is exponentially unlikely, provided we are allowed to destroy some of the sets in $\mathcal{S}$ by deleting a certain number of elements from the random set. Then, of course, there is a danger of losing other properties held by the random set. It turns out, however, that monotone properties held with exponential probabilities survive the deletion. The next lemma, also from [7], makes this precise.

For a family $\mathcal{Q}$ of subsets of a set F and an integer t, let

$$\mathcal{Q}_t = \{B : \forall D \subseteq B,\ \text{if } |D| \le t,\ \text{then } B \setminus D \in \mathcal{Q}\}\ .$$

Lemma 2.3 [7]. *Let F be a set of m elements, $0 < p < 1$, and b' and δ satisfy*

$$\delta(1 + \log_2 e - \log_2 \delta) < b'(1 - \delta)\ . \tag{2.1}$$

Then, for every increasing family $\mathcal{Q} = \mathcal{Q}(m)$ of subsets of F and for $0 < t \le \delta mp/2$, if $\mathbf{P}(F_{(1-\delta)p} \in \neg\mathcal{Q}) < 2^{-(1-\delta)b'mp}$ then $\mathbf{P}(F_p \in \neg\mathcal{Q}_t) < 2^{-b''mp}$, provided mp is large enough, where $b'' = b''(b', \delta) = \frac{1}{2}\min\{\frac{1}{2}(1-\delta)b', (log_2 e)C(\delta/2)\}$. □

Warning: The property $\mathcal{Q}$ must not depend on p, but on m only!

These two lemmas complement each other and for future reference we derive a corollary from them.

Corollary 2.1. *Let F be a set, $|F| = m$, $0 < p < 1$, and let δ and b' satisfy (2.1). Furthermore, let $p = p(m)$ and $pm \to \infty$, and let $t = \frac{1}{2}\delta mp$ be an integer. Let $\mathcal{S}$ be a family of s-element subsets of F and $\mathcal{Q} = \mathcal{Q}(m)$ be an increasing family of subsets of F. Then, for m large enough and with b'' as in Lemma 2.3, if*

$$\mathbf{P}(F_{(1-\delta)p} \in \neg\mathcal{Q}) < 2^{-b'(1-\delta)mp}$$

then, with probability at least

$$1 - 2^{-t/s} - 2^{-b''mp}\ ,$$

there exists a set $E_0 \subset F_p$, $|E_0| = t$ such that

(i) $F_p \setminus E_0 \in \mathcal{Q}$

and

(ii) $F_p \setminus E_0$ contains at most $2|\mathcal{S}|p^s$ sets from $\mathcal{S}$. □

3. SKETCH OF THE PROOF

As usual when working with random structures, we will find it convenient to switch to the binomial model F_p defined in Section 2 with $F = \{1, 2, ..., n\} = [n]$.

It is well-known (see [1]) that for monotone properties the binomial and uniform models are asymptotically equivalent provided $N \sim np$. As the sets $F \subset [n]$ satisfying $F \to (A)_r$ form an increasing family, we can formulate our Theorem 1.1 in its equivalent form.

Theorem 3.1. *For every $l \times k$ irredundant, density regular matrix A of full rank $h(A) = l$, and for every integer $r \geq 1$, there are constants c and C such that*

$$\lim_{n\to\infty} \mathbf{P}([n]_p \to (A)_r) = \begin{cases} 0 & \textit{if } p < cn^{-1/m_A} \\ 1 & \textit{if } p > Cn^{-1/m_A} \end{cases} .$$

The 0-statement, which asserts the existence of a constant c such that $\lim_{n\to\infty} \mathbf{P}([n]_p \to (A)_r) = 0$ for $p < cn^{-1/m_A}$, will be proved in Section 7. Here we concentrate on the 1-statement. It will follow from the stronger technical lemma below. The lemma will be proved by induction. To describe the induction hypothesis, we will find it convenient to apply to A Gaussian elimination with both rows and columns in reverse order, to obtain a matrix A' with the property that $a'_{ij} = 0$ whenever $j > k-l+i$, $i = 1, ..., l$. Notice that $h(A'_Q) = h(A_Q)$ and, consequently, $m_{A'} = m_A$. Furthermore, by permuting the columns of A' we obtain a matrix A'', which, in addition, satisfies the condition $a''_{l-i,k-i} \neq 0$, $i = 0, ..., l-1$ (this follows from the assumption (1.2) that A is of full rank). Notice that again $m_{A''} = m_{A'} = m_A$. Because of that and also because neither of the above operations changes the set of solutions of $\mathcal{L}(A)$, we will be assuming in our proof that the matrix A satisfies both these conditions, i.e.

$$(3.1) \quad a_{ij} = 0 \text{ whenever } j > k-l+i,\ i = 1, ..., l, \qquad \text{and } a_{l-i,k-i} \neq 0,\ i = 0, ..., l-1 .$$

For $s = 1, 2, ..., l$ let A_s be the matrix consisting of the s first rows of A truncated to the first $s+k-l$ columns of A, i.e. A_s consists of the first s rows of the matrix A^Q, where $Q = \{s + k - l + 1, ..., k\}$. For convenience, we denote by A_0 the $1 \times (k - l)$

matrix consisting of zeros only. Thus any $k-l$ tuple $(x_1,\dots,x_{k-l})$ is a solution of $A_0\mathbf{x}=\mathbf{0}$. Matrix A_0 has rank 0 and the Definition 1.1 does not apply to it. However, for the sake of uniformity, we set $m_{A_0}=1$.

An all-distinct solution $(x_1,\dots,x_{s+k-l})$ of $A_s\mathbf{x}=\mathbf{0}$ is called *nested* in a given set $F\subseteq[n]$, if there exists a vector $\mathbf{y}=(y_1,\dots,y_k)$ such that

(i) $A\mathbf{y}=\mathbf{0}$

(ii) $y_i=x_i$ for $i=1,\dots,s+k-l$

(iii) $y_i\in F$, for $i=1,\dots,k$.

For ease of presentation, we refer to an all-distinct solution of $A_s\mathbf{x}=\mathbf{0}$, which is nested in F, as to an *(s,F)-solution.* When $s=l$ $(A_l=A)$, we call these solutions simply *F-solutions.*

Lemma 3.1. *For every irredundant, density regular $l\times k$ matrix A of full rank l, for any integer $r\geq 1$ and for each $s=0,1,\dots,l$, the following statement $S(r,s)$ is true:*

Statement $\mathbf{S(r,s)}$. *For every $d>0$ there exist constants n_0, $a=a(r,s,d)$, $b=b(r,s,d)$ and $C=C(r,s,d))$ such that for every $n>n_0$ and for every set $F\subseteq[n]$ with $|F|>dn$, if $p>Cn^{-1/m_{A_s}}$ then, with probability at least $1-e^{-bnp}$, for every r-coloring of F_p, there are at least $an^{k-l}p^{s+k-l}$ monochromatic (s,F)-solutions.*

Note. Because for every subset Q of the first $k-l+s$ columns, we have $h(A^Q)-h(A_s^Q)\leq l-s$, we also have $m_{A_s}\leq m_A$.

Lemma 3.1, with $s=l$, implies the 1-statement of Theorem 3.1. We will prove the lemma by double induction on r and s.

4. Proof of Statement $S(1,s)$

By Theorem B, there are at least σn^{k-l} F-solutions of $A\mathbf{x}=\mathbf{0}$. Due to the fact that the last l columns of A are linearly independent (cf. (3.1)), these solutions are in 1-1 correspondence with a subset of (s,F)-solutions. Therefore, there are at least σn^{k-l} (s,F)-solutions.

Let X_s be the number of (s,F)-solutions which are contained in the random subset F_p. Thus, $\mathbf{E}(X_s) \geq \sigma n^{k-l} p^{s+k-l}$.

In order to prove Statement $S(1,s)$ we just need to show that, with high probability, X_s is close to its expectation. We will use Janson's inequality (Lemma 2.1). For an (s,F)-solution $\mathbf{x}$, let $I_{\mathbf{x}} = 1$ if $\mathbf{x} \in F_p$ and 0 otherwise.

Let t_n be the number of intersecting pairs of (s,F)-solutions. Then, by Proposition 2.1, $t_n \leq n^{k-l} u_n$, where u_n is the maximum number of (s,F)-solutions $\mathbf{x} = (x_1, ..., x_{k-l+s})$ intersecting a given (s,F)-solution $\mathbf{z} = (z_1, ..., z_{k-l+s})$.

To bound u_n, one needs to fix a nonempty subset of coordinates of $\mathbf{z}$ (say, $\{z_{i_1}, ... z_{i_q}\}$, $i_1 < ... < i_l$) on which the two solutions coincide together with an ordered q-tuple $Q = (j_1, ..., j_q)$ of columns of A_s such that $x_{j_1} = z_{i_1}, ..., x_{j_q} = z_{i_q}$. Having fixed Q and taking into account that matrix A_s^Q has $k-l+s-q$ columns,we see that there are, again by Proposition 2.1, at most $n^{k-l+s-q-h(A_s^Q)}$ (s,F)-solutions $\mathbf{x}$ intersecting $\mathbf{z}$ in the prescribed way. Hence,

$$u_n = O(\max_{1 \leq q \leq k} \max_{|Q|=q} n^{k-l+s-q-h(A_s^Q)})$$

and

$$\sum \sum_{\mathbf{x} \cap \mathbf{z} \neq \emptyset} \mathbf{E}(I_{\mathbf{x}} I_{\mathbf{z}}) = O((2n)^{k-l} \max_{1 \leq q \leq k} \max_{|Q|=q} n^{k-l+s-q-h(A_s^Q)} p^{2k-p}) .$$

Plugging it all into Janson's inequality, and using the fact that $h(A_s^Q) = s$ if $q = 1$ (cf. Proposition 2.2(i)), we obtain, for any $\epsilon > 0$,

$$\mathbf{P}(X_s \leq (1-\epsilon)\mathbf{E}(X_s)) < \exp\{-\Omega(\max_{1 \leq q \leq k} \max_{|Q|=q} n^{q+h(A_s^Q)-s} p^q)\} \leq \exp\{-\Omega(np)\} ,$$

by the definition of m_{A_s}. $\square$

5. Proof of Statement $S(r, 0)$

As case $r = 1$ was already treated in the previous section, we assume here that $r \geq 2$ and proceed by induction on r.

Set

$$\mathcal{T} = \{(x_1, ..., x_{k-l}) : \exists (x_{k-l+1}, ..., x_k) \in F, A\mathbf{x} = \mathbf{0}\} .$$

Thus, $\mathcal{T}$ is the set of $(0, F)$-solutions of $A\mathbf{x} = \mathbf{0}$. As was explained in the previous section,

$$|\mathcal{T}| \geq \sigma n^{k-l} . \tag{5.1}$$

Statement $S(r, 0)$ is equivalent to the fact that there exist constants a, b, and C such that for any $F \subset [n]$, $|F| > dn$, the following holds with probability at least $1 - e^{-bnp}$:

$$\text{for every } r\text{-coloring of } F_p \text{ there are at least}$$
$$a n^{k-l} p^{k-l} \text{ monochromatic } (k-l)\text{-tuples of } \mathcal{T}. \tag{5.2}$$

This will follow from a technical Lemma 5.1 below with constants a, b and C specified later. Fix a set $F \subset [n]$, $|F| > dn$.

Let us set $\mathcal{T}_{k-l} = \mathcal{T}$ and for each $1 \leq i \leq k - l - 1$ define sets $N(x_1, ...x_{k-l-i})$ and families of ordered $(k - l - i)$-tuples $\mathcal{T}_{k-l-i}$ recursively by

$$N(x_1, ...x_{k-l-i}) = \{x_{k-l-i+1} : (x_1, ..., x_{k-l-i+1}) \in \mathcal{T}_{k-l-i+1}\}$$

and

$$\mathcal{T}_{k-l-i} = \{(x_1, ..., x_{k-l-i}) : |N(x_1, ..., x_{k-l-i})| > \frac{\sigma}{2^i} n\} .$$

We will need the following estimate.

Proposition 5.1. *For every* $1 \leq i \leq k - l - 1$,

$$|\mathcal{T}_{k-l-i}| \geq \frac{\sigma}{2^i} n^{k-l-i} .$$

Proof.

Let us start with $i = 1$ and suppose that

$$|\mathcal{T}_{k-l-1}| < \frac{\sigma}{2} n^{k-l-1} .$$

Then, by Proposition 2.1

$$|\mathcal{T}| \leq |\mathcal{T}_{k-l-1}|n + n^{k-l-i}\frac{\sigma}{2}n < \sigma n^{k-l} ,$$

which contradicts (5.1).

We continue by induction on i in a similar way. Assume that

$$|\mathcal{T}_{k-l-i+1}| \geq \frac{\sigma}{2^{i-1}}n^{k-l-i+1}$$

and suppose that

$$|\mathcal{T}_{k-l-i}| < \frac{\sigma}{2^i}n^{k-l-i} .$$

Then, again by Proposition 2.1,

$$|\mathcal{T}_{k-l-i+1}| \leq |\mathcal{T}_{k-l-i}|n + n^{k-l-i}\frac{\sigma}{2^{i-1}}n < \frac{\sigma}{2^{i-1}}n^{k-l-i+1} ,$$

another contradiction. □

Our plan for the proof of (5.2) is to expose the set F in $k-l$ rounds with probabilities p_1,...,p_{k-l}, where $p_i = c_i p$, $c_1 + c_2 + ... + c_{k-l} = 1$ and $c_1 \ll c_2 \ll ... \ll c_{k-l}$ are to be specified later (cf. inequality (5.10)).

Lemma 5.1. *For every $j = 1, ..., k-l$ the following Statement $W(j)$ is true.*

Statement W(j) *: There exist constants a_j and b_j such that, with probability at least $1 - e^{-b_j np}$, for every r-coloring of the random subset $F_{p_1+...+p_j}$ either*

(i) $F_{p_1+...+p_j}$ contains at least $an^{k-l}p^{k-l}$ monochromatic elements of $\mathcal{T} = \mathcal{T}_{k-l}$

or

(ii) $F_{p_1+...+p_j}$ contains at least $a_j n^j p_1...p_j$ monochromatic elements of $\mathcal{T}_j$.

Note that for the statement $W(k-l)$, in view of our choice of a (cf. (5.11)), the alternative (ii) implies (i), which, in turn, is equivalent to (5.2).

Proof.

We will prove $W(j)$ by induction on j. Observe that the alternative (ii), say with $a_1 = \frac{1}{r}\frac{1}{2}\frac{\delta}{2^{k-l-1}}$, of Statement $W(1)$ follows immediately by Chernoff's inequality and the fact that $|\mathcal{T}_1| \geq \frac{\delta}{2^{k-l-1}}n$ (cf. the Proposition 5.1 above).

Now assume that $W(j-1)$ is true. Let $\mathcal{A}$ be the event that there is an r-coloring of $F_{p_1+...+p_j}$ for which both, (i) and (ii) of Statement $W(j)$ are false. Let $\mathcal{B}_{(i)}$ and $\mathcal{B}_{(ii)}$ be the events that, for every r-coloring of $F_{p_1+...+p_{j-1}}$, part (i) and part (ii), respectively, of Statement $W(j-1)$ holds and, moreover, that $|F_{p_1+...+p_{j-1}}| < 2dn(p_1 + ... + p_{j-1})$. If the alternative (i) holds for $W(j-1)$ then it does for $W(j)$. Hence the events $\mathcal{A}$ and $\mathcal{B}_{(i)}$ are disjoint and

$$\mathbf{P}(\mathcal{A}) \leq \mathbf{P}(\neg\mathcal{B}_{(i)} \cap \neg\mathcal{B}_{(ii)}) + \sum_{K\in\mathcal{B}_{(ii)}} \mathbf{P}(\mathcal{A}|K)\mathbf{P}(K) \ . \tag{5.3}$$

Conditional on the outcome $K = F_{p_1+...+p_{j-1}}$ of the first $j-1$ rounds, for every r-coloring $h : K \to [r]$, let $\mathcal{A}_h$ be the event that there is an extension $\bar{h} : F_{p_1+...+p_j} \to [r]$ of h such that $\bar{h} \equiv h$ on K and both, (i) and (ii) of Statement $W(j)$ are false. Then

$$\mathbf{P}(\mathcal{A}|K) = \mathbf{P}(\bigcup_h \mathcal{A}_h|K) \leq r^{2dn(p_1+...+p_{j-1})}\mathbf{P}(\mathcal{A}^{h_0}|K) \ , \tag{5.4}$$

where h_0 maximizes the conditional probability.

By the inductive assumption and Chernoff's inequality,

$$\mathbf{P}(\neg\mathcal{B}_{(i)} \cap \neg\mathcal{B}_{(ii)}) < e^{-b_{j-1}np} + e^{-C(1)dn(p_1+...p_{j-1})} \ . \tag{5.5}$$

It remains to estimate $\mathbf{P}(A^h|K)$ for a fixed instance K of $F_{p_1+...+p_{j-1}}$ such that $K \in \mathcal{B}_{(ii)}$ and for a fixed r-coloring h of K.

So, we know that K contains at least $a_{j-1}n^{j-1}p_1...p_{j-1}$ red elements of $\mathcal{T}_{j-1}$. Set $d_j = \frac{\sigma}{2^{k-l-j+2}}$ and consider the set D_j of all elements of F which belong to $N(x_1, ..., x_{j-1})$ for at least $a_{j-1}d_jn^{j-1}p_1...p_{j-1}$ red $(j-1)$-tuples $(x_1, ..., x_{j-1})$ of $\mathcal{T}_{j-1}$.

Proposition 5.2. *$|D_j| \geq d_jn$.*

Proof.

Suppose $|D_j| < d_jn$ and consider the bipartite graph with $a_{j-1}n^{j-1}p_1...p_{j-1}$ red $(j-1)$-tuples of $\mathcal{T}_{j-1}$ on the left and the elements of F on the right. An edge joins a tuple $(x_1, ..., x_{j-1})$ with x_j if, and only if $x_j \in N(x_1, ..., x_{j-1})$. On one hand, by the definition of $\mathcal{T}_{j-1}$, this bipartite graph contains more than

$$a_{j-1}n^{j-1}p_1...p_{j-1}2d_jn$$

edges. On the other hand, by the definition of D_j, it contains less than

$$d_jna_{j-1}n^{j-1}p_1...p_{j-1} + na_{j-1}d_jn^{j-1}p_1...p_{j-1}$$

edges, which is a contradiction. $\square$

Let $D'_j = D_j \setminus K$. As the set K is only of order of magnitude of np, we have $|D'_j| > d'_jn$, for some $.999d_j < d'_j < d_j$. Now we shall expose D'_j with probability $(1 - \delta_j)p_j$, where δ_j satisfies inequality (2.1) with $b' = b(r-1, 0, d'_j)/d'_j$.

By our inductive assumption, Statement $S(r-1, 0)$ is true, which implies that, with probability at least $1 - e^{-(1-\delta_j)b(r-1,0,d'_j)np_j}$, for every $(r-1)$-coloring of $(D_j)_{(1-\delta_j)p_j}$, there are at least $a(r-1, 0, d'_j)n^{k-l}((1-\delta_j)p_j)^{k-l}$ monochromatic members of $\mathcal{T}$, provided $(1-\delta_j)p_j > C(r-1, 0, d'_j)$.

Consequently, by Lemma 2.3, applied with $F = D'_j$, $p = p_j$ and $\delta = \delta_j$, the above property remains true for $(D'_j)_{p_j}$ with probability $1 - e^{-b''np_j}$, where $b'' = b''(b', \delta_j)$ as in Lemma 2.3, even after deleting up to $\frac{1}{2}\delta_jd'_jnp_j$ elements from $(D'_j)_{p_j}$, i.e. for every set $E \subset (D'_j)_{p_j}$, $|E| < \frac{1}{2}\delta_jd'_jnp_j$, and for every $(r-1)$-coloring of $(D'_j)_{p'_j} \setminus E$, there are at least

$$a(r-1, 0, d'_j)n^{k-l}((1-\delta_j)p_j)^{k-l}$$

monochromatic members of $\mathcal{T}$.

Now we ask an adversary to choose an extension $\bar{h}$ of the given r-coloring h of K. If she colors less than $\frac{1}{2}\delta_jd'_jnp_j$ elements of $(D'_j)_{p_j}$ red, then, denoting by E the set of these elements, we are facing an $(r-1)$-coloring of $(D'_j)_{p_j} \setminus E$ and the alternative (i) of Statement $W(j)$ follows for any

$$a \leq a(r-1, 0, d'_j)((1-\delta_j)c_j)^{k-l} \ . \tag{5.6}$$

If, however, red is used on at least $\frac{1}{2}\delta_jd'_jnp_j$ elements of $(D'_j)_{p_j}$ then, as each element of D_j extends to at least $a_{j-1}d_jn^{j-1}p_1...p_{j-1}$ red j-tuples of $\mathcal{T}_j$, we have at least

$$a_{j-1}\frac{1}{2}\delta_j d_j d'_j n^j p_1 ... p_j \tag{5.7}$$

red tuples of $\mathcal{T}_j$ in $F_{p_1+...+p_j}$. Hence the alternative (ii) of $W(j)$ is true with

$$a_j = a_{j-1}\frac{1}{2}\delta_j d_j d'_j \ . \tag{5.8}$$

Putting these two alternatives together, we obtain, for a fixed instance K of the random subset $F_{p_1+...+p_{j-1}}$ and for a given r-coloring h of K, that

$$\mathbf{P}(\mathcal{A}_h|K) \le e^{-b'' d'_j n p_j} \ . \tag{5.9}$$

In order to estimate $\mathbf{P}(\mathcal{A}|K)$, we need, by (5.4), that

$$r^{2dn(p_1+...+p_{j-1})}e^{-b'' d'_j n p_j}$$

is exponentially small in $\Omega(np)$. This will be guaranteed by setting

$$c_j \ge 2\frac{2d(c_1 + ... + c_{j-1})}{b'' d'_j} \ . \tag{5.10}$$

Now, (5.3-5.5) and (5.9,5.10) give

$$\mathbf{P}(\mathcal{A}) \le e^{-b_{j-1}np} + e^{-C(1)dn(p_1+...p_{j-1})} + e^{-b'' d'_j n p_j} \le e^{-b_j n}$$

for suitably chosen b_j. This completes the proof of the implication $W(j-1) \Longrightarrow W(j)$ with a_j defined by (5.8), and hence also of Lemma 5.1.

To conclude the proof of Statement $S(r,0)$, set, by (5.6),

$$a = \min\{c_1 ... c_{k-l} a_{k-l},\ a(r-1, d'_j)((1-\delta_j)c_j)^{k-l},\ j = 1, ..., k-l\} \ , \tag{5.11}$$

$b = b_{k-l}$ and $C = \frac{2}{c_1}C(r-1, 0, d'_1)$. (Note that $C(r-1, 0, d'_1)$ is a decreasing function of d'_1 and that d_j, and therefore also d'_j, increases with j.

6. PROOF OF STATEMENT $S(r,s)$, $r \geq 2$, $s \geq 1$

In the proof of statement $S(r,s)$ we shall use the statements $S(r,s-1)$ and $S(r-1,s)$. Let $F \subseteq [n]$, $|F| > dn$.

We will employ the well known technique called *the two-round exposure.* Representing $p = p_1 + p_2 - p_1p_2$, one first generates the random subset F_{p_1}, conditions on the outcome, colors it, and only then generates F_{p_2}. We shall be assuming that both p_1 and p_2 are of the same order of magnitude as p, but that p_2 is sufficiently bigger than p_1.

We say that an $(s-1,F)$-solution $(x_1,...,x_{k-l+s-1})$ *focuses* on x_{k-l+s} if $(x_1,...,x_{k-l+s})$ is an (s,F)-solution. (Note, that x_{k-l+s}, if exists, is uniquely determined by an $(s-1,F)$-solution.) For an instance K of the random subset F_{p_1} and for an r-coloring h of K, let $H_i = H_i(K,h)$, $i = 1,...,r$, be the set of all elements of F which are in the focus of at least $\frac{1}{2}a'n^{k-l-1}p_1^{k-l+s-1}$ monochromatic $(s-1,F)$-solutions in color i, where $a' = a(r,s-1,d)$ is the constant appearing in Statement $S(r,s-1)$.

Loosely speaking, we shall show, using $S(r,s-1)$, that as a result of round 1, at least one of the sets H_i will be large and therefore many of its elements will survive through the second round. Then, if color i is used a lot on $(H_i)_{p_2}$, we obtain many (s,F)-solutions in that color. If this color is used a little, we make a twist which enables us to neglect this color altogether, and Statement $S(r-1,s)$ completes the proof.

Formally, this outline can be described as follows. Let $\mathcal{A}$ be the event that there exists an r-coloring of F_p, with less than $an^{k-l}p^{k-l+s}$ monochromatic (s,F)-solutions in F_p in each color.

Let $\mathcal{B}$ be the event that $|F_{p_1}| < 2np_1$ and that for every $h : F_{p_1} \to [r]$ there is an $i \in [r]$ such that $|H_i(F_{p_1},h)| > d'n$. Clearly,

$$\mathbf{P}(\mathcal{A}) \leq \mathbf{P}(\neg\mathcal{B}) + \sum_{K\in\mathcal{B}} \mathbf{P}(\mathcal{A}|K)\mathbf{P}(K) \ .$$

Conditioning on the outcome $K = F_{p_1}$ of the first round, for every $h : K \to [r]$, let

$\mathcal{A}_h$ be the event that there is an extension $\bar{h} : F_p \to [r]$ of h such that $\bar{h} \equiv h$ on K and there are less than $an^{k-l}p^{k-l+s}$ monochromatic (s, F)-solutions in F_p in each color. Then

$$\mathbf{P}(\mathcal{A}|K) = \mathbf{P}(\bigcup_h \mathcal{A}_h|K) \le r^{2np_1}\mathbf{P}(\mathcal{A}^{h_0}|K) \ ,$$

where h_0 maximizes the conditional probability. Thus, all we have to show is that

(A) $\mathbf{P}(\neg\mathcal{B}) = e^{-b_1np_1}$

and that

(B) for every $K \in \mathcal{B}$ and for every r-coloring h of the elements of K,

$$\mathbf{P}(\mathcal{A}_h|F_{p_1} = K) \le e^{-b_2np_2} \ ,$$

where b_1 and b_2 are constants and $4(\log r)p_1 < b_2p_2$.

The first component of the event $\mathcal{B}$, the inequality $|F_{p_1}| < 2np_1$, is an immediate consequence of Chernoff's inequality. Now we shall prove the essential part of the statement (A).

Given an r-coloring h of F_{p_1}, for each element $x \in F$, let $\alpha_x^{(i)}$ be the number of $(s-1, F)$-solutions colored by color i, focused on x, and contained in F_{p_1}. By our inductive assumption $S(r, s-1)$, we know that, with probability $1 - e^{-b'np_1}$, $b' = b(r, s-1, d)$, for every h, there is an i such that

$$\sum_{x\in F} \alpha_x^{(i)} \ge a(r, s-1, d)n^{k-l}p_1^{k-l+s-1} \ . \tag{6.1}$$

We want to show that the number t of those x's for which the parameter $\alpha_x^{(i)} = \alpha_x$ exceeds $1/(2n)$ of the right-hand side of (6.1) is at least $d'n$. (cf. the definition of of H_i and of property $\mathcal{B}$.)

Let us order the $\alpha_x's$ from high to low

$$\alpha_1 \ge \dots \ge \alpha_t \ge \frac{1}{2}a'n^{k-l-1}p_1^{k-l+s-1} > \alpha_{t+1} \ge \dots \ge \alpha_{|F|} \ .$$

Then

$$\sum_{i=1}^{t} \alpha_i \ge \frac{1}{2}a'n^{k-l}p_1^{k-l+s-1} \ .$$

If $t > \frac{1}{3}\sum_{i=1}^{t} \alpha_i$, then

$$t > \frac{1}{6} a' n^{k-l} p_1^{k-l+s-1} \ .$$

As, by (1.1) and the fact that A_s has $k-l+s$ columns and s rows, $m_{A_s} \geq \frac{k-l+s-1}{k-l-1}$, we have, with $p_1 = c_1 n^{-1/m_{A_s}}$ and $c_1 > \frac{6d'}{a'}$, that $t > d'n$.

Otherwise, i.e. when $t \leq \frac{1}{3}\sum_{i=1}^{t} \alpha_i$, we denote by X the number of pairs of $(s-1, F)$-solutions focusing on the same element and contained in F_{p_1}. Observe that $X = \sum_{x \in F} \binom{\alpha_x}{2}$. We aim toward bounding X from above. As a first step we now attempt to bound the expectation of X, $\mathbf{E}(X)$. To this end, consider a fixed $(s-1, F)$-solution $\mathbf{z} = (z_1, ..., z_{k-l+s-1})$ and let z_{k-l+s} be its focus. Any other $(s-1, F)$-solution $\mathbf{x}$ with the same focus intersects $\mathbf{z}$ in some $q-1 \geq 0$ elements. Let Q, $|Q| = q$, be the subset of columns of matrix A_s corresponding to the variables of $\mathbf{x}$ which coincide with some values of $\mathbf{z}$, including z_{k-l+s}. The probability that a pair $\mathbf{x}, \mathbf{z}$ is present in F_{p_1} is precisely $p_1^{2(k-l+s-1)-(q-1)}$. Thus, for some suitably chosen constant a_1, by Proposition 2.1,

$$\mathbf{E}(X) \leq a_1 n^{k-l} n^{k-l+s-q-h_Q} p_1^{2(k-l+s-1)-q+1}$$

where Q maximizes the above quantity and $h_Q = h(A_s^Q)$.

Assume for a moment that with probability close enough to 1, $X < 2\mathbf{E}(X)$. Then,

$$\sum_{x \in F} \binom{\alpha_x}{2} < 2\mathbf{E}(X) \leq 2a_1 n^{k-l} n^{k-l+s-q-h_Q} p^{2(k-l+s-1)-q} \ . \tag{6.2}$$

Then, by Jensen's inequality,

$$\sum_{i=1}^{t} \binom{\alpha_i}{2} \geq \frac{1}{3}\left(\frac{1}{2} a' n^{k-l} p_1^{k-l+s-1}\right)^2 \ .$$

Comparing the last inequality with (6.2), we obtain

$$t > \frac{(a')^2}{24 a_1} n^{q+h_Q-s} p_1^{q-1} > d'n$$

for sufficiently large p_1 (cf. Definition 1.1 of m_A and apply it to A_s).

Unfortunately, we cannot claim the inequality $X < 2\mathbf{E}(X)$ with sufficiently high probability. Therefore, we need to refine our approach. For $E \subset [n]$, let X_E be the number of pairs of $(s-1, F)$-solutions with the same focus contained in $F_{p_1} \setminus E$. We will show that, with probability at least $1 - e^{-b_1 np_1}$, there exists a set $E_0 \subset [n]$ such that $X_{E_0} < 2\mathbf{E}(X)$, while at the same time, for every 2-coloring of $F_{p_1} \setminus E_0$, an inequality only slightly weaker than (6.1) is valid. This will enable us to literally repeat the above argument with only minor adjustments.

To achieve that, we will apply Corollary 2.1 with $\delta > 0$ and so small that the inequality (2.1) holds with the constant b' from (6.1). Furthermore, let $\mathcal{S}$ be the family of all pairs of $(s-1, F)$-solutions focusing on the same element and contained in F_{p_1} and let property $\mathcal{Q}$ state that inequality (6.1) holds for every r-coloring, with p_1 replaced by $(1-\delta)p_1$. In other words, the property $\mathcal{Q}$ considered here is the family of all subsets $R \subseteq F$ such that for every partition $R = R_1 \cup R_2 \cup \ldots \cup R_r$ we have

$$\sum_{x \in F} \alpha_x^{(i)} \geq a' n^{k-l}((1-\delta)p_1)^{k-l+s-1}$$

in at least one of the sets R_i. Note, that by fixing $(1-\delta)p_1$, we made property $\mathcal{Q}$ independent of p and that, therefore, it is increasing.

Replacingp_1 by $(1-\delta)p_1$ in(6.1), we obtain

$$\mathbf{P}(F_{(1-\delta)p_1} \in \neg\mathcal{Q}) < 2^{-b' n(1-\delta)p_1} .$$

Now, by Corollary 2.1, with probability at least $1 - 2^{-b'' np_1} - 2^{-t/6}$, there is a set of $E_0 \subseteq F_{p_1}$, $|E_0| = t = \frac{1}{2}\delta\binom{n}{3}p_1^3$, such that both, (6.1) holds for $F_{p_1} \setminus E_0$ with the extra factor of $(1-\delta)^{k-l+s-1}$ on the right, and $X_{E_0} < 2\mathbf{E}(X)$.

Thus, all our previous estimates hold, with the additional factor $(1-\delta)$ raised to the appropriate power, and so, $t > d'n$, for small enough δ. This completes the proof of fact (A).

In order to prove (B), consider an outcome $K \in \mathcal{B}$ of the first round and a coloring h of K. By the definition of $\mathcal{B}$, we have $|K| \leq 2np_1$ and, for some $i \in [r]$, $|H_i(K, h)| > d'n$.

Set $D = H_i \setminus K$ and observe that $|D| > .999d'n = d''n$.

Let $\delta > 0$ satisfy the inequality (2.1) with $b' = b(r-1, s, d'')/d''$. We now apply Lemma 2.3 with $D = F$, $p = p_2$, and $\mathcal{Q}$ being the property from the statement $S(r-1, s)$, i.e., setting $a'' = a(r-1, s, d'')$, $\mathcal{Q}$ is the family of all subsets $R \subseteq D$ such that for every $(r-1)$-coloring $R = R_1 \cup \ldots \cup R_{r-1}$, in at least one of the color classes there are at least $a''n^{k-l}((1-\delta)p_2)^{k-l+s}$ (s, F)-solutions. Finally, we set $t = \frac{1}{2}\delta|D|p_2$.

As, by statement $S(r-1, s)$,

$$\mathbf{P}(D_{(1-\delta)p_2} \notin \mathcal{Q}) < 2^{-(1-\delta)b'|D|p_2}$$

we infer, by Lemma 2.3, that

$$\mathbf{P}(D_{p_2} \notin \mathcal{Q}_t) < 2^{-b''|D|p_2} \ , \tag{6.3}$$

i.e., with high probability, for every subset $E \subset D_{p_2}$, $|E| \leq t$, the set $D_{p_2} \setminus E$ still has the property $\mathcal{Q}$.

Now, we let the adversary finish the coloring of F_p. More formally, we consider an extension $\bar{h} : F_p \to [r]$ of $h : K \to [r]$.

Recall that $D \subset H_i(K, h)$ for some fixed color $i \in [r]$ and consider two cases.

If the coloring $\bar{h}$ uses color i less than t times on D_{p_2}, then, by (6.3), after deleting the elements of color i, the remaining set still has property $\mathcal{Q}$, i.e. there is a color j ($j \neq i$) such that there are at least $a''n^{k-l}((1-\delta)p_2)^{k-l+s}$ (s, F)-solutions in color j.

If, on the other hand, coloring $\bar{h}$ uses color i at least t times on $D_{p_2} \subset H_i$, then, recalling the definition of H_i, we infer that there are in F_p at least

$$t\frac{1}{2}a'n^{k-l-1}p_1^{k-l+s-1} \geq an^{k-l}p^{k-l-1}$$

(s, F)-solutions in color i, for a suitable choice of constant $a > 0$. Hence, fact (B) and therefore the entire statement $S(r, s)$ is proved. □

7. Proof of the 0-statement

For the proof of the negative part of Theorem 3.1 we assume that $p = cn^{-1/m_A}$, where c is a sufficiently small constant.

Let A be an $l \times k$ irredundant, partition regular matrix of full rank l. For $Q \subseteq [k]$ and $\bar{Q} = [k] \setminus Q$, let $\mathbf{r}_1^Q$, $\mathbf{r}_2^Q$,..., $\mathbf{r}_l^Q$, and $\mathbf{r}_1^{\bar{Q}}$, $\mathbf{r}_2^{\bar{Q}}$,..., $\mathbf{r}_l^{\bar{Q}}$ be the rows of $A^{\bar{Q}}$ and A^Q, respectively. (Recall that due to our peculiar notation $A^{\bar{Q}}$ consists of columns $\mathbf{c}_i$, $i \in Q$.)

Set $h = h(A^Q)$. Without loss of generality assume that the first h rows $\mathbf{r}_1^{\bar{Q}}$,..., $\mathbf{r}_h^{\bar{Q}}$ of matrix A^Q are linearly independent. Hence, for each i, $h < i \leq l$,

$$\mathbf{r}_i^{\bar{Q}} = \sum_{j=1}^{h} \delta_j^i \mathbf{r}_j^{\bar{Q}} \tag{7.1}$$

for some δ_1^i,...,δ_h^i.

The following definition is crucial in our proof.

Definition 7.1 The matrix $B = B(A, Q)$ is the matrix consisting of $|Q|$ columns and $l - h$ rows

$$\mathbf{r}_i^Q - \sum_{j=1}^{h} \delta_j^i \mathbf{r}_j^Q$$

$i = h + 1, ..., l$.

For a vector $\mathbf{x} = (x_1, ..., x_k)$ and $\emptyset \neq Q \subseteq [k]$, let $\mathbf{x}_Q = (x_i)_{i \in Q}$. The following claim describes an important property of matrix $B = B(A, Q)$.

Proposition 7.1. *For every $\emptyset \neq Q \subseteq [k]$, if $A\mathbf{x} = \mathbf{0}$, then $B\mathbf{x}_Q = \mathbf{0}$.*

Proof.

Let $A\mathbf{x} = \mathbf{0}$. Thus, denoting by $\mathbf{r}_i$ the i-th row of A,

$$\mathbf{r}_i \mathbf{x} = \mathbf{r}_i^Q \mathbf{x}_Q + \mathbf{r}_i^{\bar{Q}} \mathbf{x}_{\bar{Q}} \tag{7.2}$$

$i = 1, ...l$.

Combining (7.1) and (7.2), we get, for $i = 1, 2, ...l - h$, that

$$\mathbf{r}_{h+i}^Q \mathbf{x}_Q = \sum_{j=1}^{h} \delta_j^i \mathbf{r}_j^Q \mathbf{x}_Q$$

or subsequently $B\mathbf{x}_Q = \mathbf{0}$. □

Definition 7.2 We will call an $l \times k$, irredundant, partition regular matrix A, with $h(A) = l$, *strictly balanced* if, for every $Q \subset [k]$, $2 \leq q = |Q| < k$, the inequality

$$\frac{q-1}{q-1+h(A^Q)-l} < \frac{k-1}{k-1-l}$$

holds.

For instance, the matrix A corresponding to the system of equations (1.1), the solution of which are arithmetic progresions of length k, is, by the argument leading to formula (1.3), strictly balanced. Our next lemma says that a strictly balanced matrix can be associated (in the sense of Definition 7.1) with any irredundant, partition regular, full rank matrix.

Lemma 7.1. *Let A be an $l \times k$, irredundant, partition regular matrix, with $h(A) = l$, and let q_0 be the smallest integer with the property that there exists $Q_0 \subseteq [k]$, $|Q_0| = q_0$, such that*

$$\max_Q \frac{q-1}{q-1+h(A^Q)-l} = \frac{q_0-1}{q_0-1+h(A^{Q_0})-l} .$$

Then, the matrix $B = B(A, Q_0)$ is strictly balanced.

The proof of the Lemma 7.1 will follow from the next Proposition 7.2.

Proposition 7.2. *Let $Q \subset Q_0 \subseteq [k]$ and let B^Q be the matrix obtained from $B = B(A, Q_0)$ by omitting the columns of Q. Then*

$$h(A^Q) = h(B^Q) + h(A^{Q_0}) .$$

Note. For $Q = \emptyset$ the Lemma implies that $h(A) = h(B) + h(A^{Q_0})$. If $h(A) = l$, this means that the matrix B is of full rank.

Proof of Proposition 7.2.

Set $U = \{\mathbf{u} : A^Q\mathbf{u} = \mathbf{0}\}$ and $V = \{\mathbf{v} : B^Q\mathbf{v} = \mathbf{0}\}$. Let us consider a linear mapping defined by $\phi(\mathbf{u}) = (\mathbf{u})_{Q_0 \setminus Q}$.

The proof of Proposition 7.2 will be an easy consequence of the following Proposition 7.3 7.1, the proof of which is postponed until later.

Proposition 7.3. *The mapping $\phi : U \to V$ is a homomorphism onto V.*

Set $W = \phi^{-1}(0)$. By the dimension theorem,

$$dim(U) = dim(W) + dim(V)\ . \tag{7.3}$$

Clearly, we have

$$dim(U) = k - q - h(A^Q) \tag{7.4}$$

and

$$dim(V) = q_0 - q - h(B^Q)\ . \tag{7.5}$$

As

$$\phi^{-1}(0) = \{\mathbf{u} = (\mathbf{0}, \mathbf{w}) \in U;\ \text{where } \mathbf{w} = (\mathbf{u})_{\bar{Q}_0}\} = \{(\mathbf{0}, \mathbf{w}) : A^{Q_0}\mathbf{w} = \mathbf{0}\}$$

and the last subspace of U is isomorphic to the space $\{\mathbf{w} : A^{Q_0}\mathbf{w} = \mathbf{0}\}$, we also have

$$dim(W) = k - q_0 - h(A^{Q_0})\ . \tag{7.6}$$

Combining (7.3-7.6) yields Proposition 7.2. □

Proof of Proposition 7.3.

Let us first prove that $\phi(U) \subseteq V$. For $\mathbf{u} = (\mathbf{v}, \mathbf{w}) \in U$, where $\mathbf{v} = \mathbf{u}_{Q_0 \setminus Q}$ and $\mathbf{w} = \mathbf{u}_{\bar{Q}_0}$, consider a k-dimensional vector $\mathbf{x} = (\mathbf{0}, \mathbf{u})$, where $(\mathbf{x})_Q = \mathbf{0}$. Then $A\mathbf{x} = \mathbf{0}$ and by Proposition 7.1, $B\mathbf{y} = B(A, Q_0)\mathbf{y} = \mathbf{0}$, where $\mathbf{y} = (\mathbf{x})_{Q_0}$. As $\mathbf{y} = (\mathbf{0}, \mathbf{v})$, where $(\mathbf{y})_Q = \mathbf{0}$ and $(\mathbf{y})_{Q_0 \setminus Q} = \mathbf{v}$, we conclude that $B^Q\mathbf{v} = \mathbf{0}$ and hence $\mathbf{v} \in V$.

Finally, we verify that the mapping ϕ is onto V. This means that we need to show that every $\mathbf{v} \in V$ has "an extension" $\mathbf{u} \in U$ satisfying $\mathbf{u}_{Q_0 \setminus Q} = \mathbf{0}$.

Let $\mathbf{v} \in V$, i.e. let $B^Q\mathbf{v} = \mathbf{0}$. By the definitions of matrices $B = B(A, Q_0)$ and B^Q, we have, for every $i = 1, 2, ..., l - h$,

$$\mathbf{r}_{h+i}^{Q_0\setminus Q}\mathbf{v} = \sum_{j=1}^{h} \delta_j^i \mathbf{r}_i^{Q_0\setminus Q}\mathbf{v} \tag{7.7}$$

where $h = h(A^{Q_0})$ and the coefficients δ_j^i also satisfy

$$\mathbf{r}_{h+i}^{\bar{Q}_0} = \sum_{j=1}^{h} \delta_j^i \mathbf{r}_i^{\bar{Q}_0} \ . \tag{7.8}$$

Consider now the system of linear equations

$$\mathbf{r}_i^{\bar{Q}_0}\mathbf{w} = -\mathbf{r}_i^{Q_0\setminus Q}\mathbf{v} \overset{df}{=} \mathbf{b_i} \tag{7.9}$$

$i = 1, ..., l$, with variable $\mathbf{w}$ and fixed $\mathbf{v}$.

As, due to our notation, the row vectors $\mathbf{r}_j^{\bar{Q}_0}$, $j = 1, ..., h$, are linearly independent, the system of the first h equations in (7.9), i.e. the system

$$\mathbf{r}_j^{\bar{Q}_0}\mathbf{w} = -\mathbf{r}_j^{Q_0\setminus Q}\mathbf{v} = \mathbf{b_j} \ ,$$

$j = 1, ..., h$, has a solution $\mathbf{w}$. Due to (7.7) and (7.8), this solution satisfies the remaining $l - h$ equations of (7.9) as well. Setting $\mathbf{u} = (\mathbf{v}, \mathbf{w})$, we infer by (7.9) that $A^Q\mathbf{u} = \mathbf{0}$, i.e. $\mathbf{u} \in U$. □

Proof of Lemma 7.1.

Recall first that the matrix $B = B(A, Q_0)$ has $q_0 = |Q_0|$ columns, $l - h(A^{Q_0})$ rows and the full rank $h(B) = l - h(A^{Q_0})$ (cf. Note after Proposition 7.2).

Suppose now that the matrix B is not strictly balanced, i.e. there exists $Q \subset Q_0$, $q = |Q| \geq 2$, such that

$$\frac{q-1}{q-1+h(B^Q)-(l-h(A^{Q_0}))} \geq \frac{q_0-1}{q_0-1-(l-h(A^{Q_0}))} \ .$$

By Proposition 7.2, however,

$$\frac{q-1}{q-1+h(B^Q)-(l-h(A^{Q_0}))} = \frac{q-1}{q-1-(l-h(A^{Q}))} \ ,$$

which contradicts our choice of Q_0. □

By Proposition 7.1, if $F \nrightarrow (A)_r$ then $F \nrightarrow (B)_r$. Hence, when proving the 0-statement of Theorem 3.1, we may assume without loss of generality that A is strictly balanced, since if this were not the case, one could replace A with $B = B(A, Q_0)$, where Q_0 is as in Lemma 7.1. Thus, we now have

$$m_A = \frac{k-1}{k-l-1} \quad \text{and} \quad p = cn^{-\frac{k-l-1}{k-1}} .$$

Also we may restrict our attention to the case when $r = 2$, since, trivially, for $r > 2$, $F \nrightarrow (A)_r$ implies $F \nrightarrow (A)_2$.

Our proof will consist of two statements, one deterministic, saying that the property $F \rightarrow (A)_2$ implies the existence of a certain structure in F, while the probabilistic statement will almost surely exclude that structure from the random set F_p. We shall need a few definitions first.

A simple path is a hypergraph consisting of edges E_1,...,E_l, $l \geq 1$, such that

$$|E_i \cap E_j| = \begin{cases} 1 \text{ if } j = i+1, \quad i = 1, ..., l-1 \\ 0 \text{ otherwise} . \end{cases}$$

A fairly simple cycle is a hypergraph which consists of a simple path $(E_1, ..., E_l)$, $l \geq 2$, and an edge E_0 such that

$$|E_0 \cap E_i| = \begin{cases} 1 \text{ if } i = 1 \\ 0 \text{ for } i = 2, ..., l-1 \\ s \text{ if } i = l \ , \end{cases}$$

where $s \geq 1$. A fairly simple cycle is said to be *simple* if $s = 1$. A fairly simple but not simple cycle will be called *spoiled.*

A simple path P of a hypergraph H is called *spoiled* if it is not an induced subhypergraph of H, i.e. there is an edge E in H such that $E \notin E(P)$ but $E \subset V(P)$.

A subhypergraph H_0 of H is said to have *a handle* if there is a edge E in H such that $|E| > |E \cap V(H_0)| \geq 2$.

For a set of integers F and a matrix A, let $H(F, A)$ be the hypergraph with the vertex set $V(F)$ whose edges are the sets of all-distinct solutions of $A\mathbf{x} = \mathbf{0}$ which are contained in F.

Deterministic Lemma. *If* $F \to (A)_2$ *then the hypergraph* $H(F, A)$ *contains either a fairly simple cycle with a handle or a spoiled simple path.*

Probabilistic Lemma. *If* p *and* A *are as above then, almost surely, the random hypergraph* $H(F_p, A)$ *contains neither a fairly simple cycle with a handle nor a spoiled simple path.*

The proof of Deterministic Lemma.

Assume that $F \to (A)_2$. This is equivalent to saying that the chromatic number of $H(F, A)$ is at least 3. We may assume that $H(F, A)$ is edge-critical with respect to that property or otherwise we could replace $H(F, A)$ with its 3-edge-critical subgraph, ignoring some F-solutions of $A\mathbf{x} = \mathbf{0}$. As such, it satisfies the following property.

Proposition 7.4. *If H is a 3-edge-critical hypergraph then for every edge $E \in H$ and for every vertex $v \in E$ there is $E' \in H$ such that $E \cap E' = \{v\}$.*

Proof.

Let H be a 3-edge-critical hypergraph, and suppose that there is an edge $E \in H$ and a vertex $v \in E$, so that every edge E' containing v contains also another vertex of E. By the 3-edge-criticality, the edges of G can be blue-red colored in such a way that only E is monochromatic, say blue. Now, by changing the color of v to red, E ceases to be monochromatic and, at the same time, no other edge of H becomes monochromatic, since every edge containing v contains also a vertex colored blue. This is, however, a contradiction with the fact that the chromatic number of H is 3. □

Let P be the longest simple path in $H = H(F, A)$. By Proposition 7.4, P contains at least two edges of H. Let x and y be two vertices which belong to only the first edge of P, and let E_x and E_y be two edges of H (read: solutions to $A\mathbf{x} = \mathbf{0}$) whose existence is guarantied by Proposition 7.4, i.e. $E_z \cap E_1 = \{z\}$, $z = x, y$.

By the maximality of P, $h_z = |V(P) \cap E_z| \geq 2$, $z = x, y$. Let $i_z = \min\{i \geq 2 : E_z \cap E_i \neq \emptyset\}$, $z = x, y$, and assume that, say, $i_y \leq i_x$. If $h_z = k$ for some z, then P is a spoiled simple path. Otherwise, the edges $E_1, ...E_{i_x}, E_x$ form a fairly simple cycle for which E_y is a handle. □

The proof of Probabilistic Lemma.

Let X, Y, Z, and W be random variables counting, respectively, simple paths of length at least $B \log n$, spoiled cycles, simple cycles of length less than $B \log n + 1$ with handles, and spoiled simple paths of length less than $B \log n$ in the random superhypergraph $H(F_p, A)$, where $B = B(c, A)$ is a big enough constant. Straightforward estimates show that their expectations all converge to 0 as $n \to \infty$. Below we frequently use the equation

$$n^{k-l-1} p^{k-l} = c \ .$$

Indeed, by Proposition 2.2(i),

$$\mathbf{E}(X) < O\left(\sum_{t > B \log n} n^{k-l} n^{(t-1)(k-l-1)} p^{k+(t-1)(k-1)} \right) = O\left(np \sum_{t > B \log n} c^t \right) = o(1) \ .$$

For estimating both, Y and Z we will utilize the following consequence of our assumption that the matrix A is strictly balanced. It is easy to verify that the inequality of Definition 7.2 is equivalent to the inequality

$$k - q - h_Q - \frac{1}{m_A}(k - q) < 0 \ ,$$

where $h_Q = h(A^Q)$. (This is because both these inequalities are equivalent to $l(k-q) < k(h_Q - 1)$.) This implies that, for each Q, with $2 \le q = |Q| < k$,

$$n^{k-q-h_Q} p^{k-q} = o(n^{-\varepsilon}) \tag{7.10}$$

for some $\varepsilon > 0$.

To estimate $\mathbf{E}(Y)$, we begin with a pair of edges which spoil the cycle, i.e. which intersect each other on at least $q \ge 2$ elements, and continue along the cycle until the last edge closes it by sharing one vertex with both the previous and the first edge. Thus, by Proposition 2.1, Proposition 2.2(i) and (ii), and by inequality (7.10)

$$\begin{aligned} \mathbf{E}(Y) &< O\left(\sum_{t>2} \sum_{Q} n^{k-l} p^k n^{k-q-h_Q} p^{k-q} (n^{k-l-1} p^{k-1})^{t-3} n^{k-2-l} p^{k-2} \right) \\ &= O\left(\sum_{t>2} \sum_{Q} c^{t-1} n^{k-q-h_Q} p^{k-q} \right) = o(1) \ . \end{aligned}$$

Similarly,

$$\mathbf{E}(Z) = O\left(\sum_{t=3}^{B\log n}\sum_{Q} n^{t(k-l-1)}p^{t(k-1)}(\log n)^{k-1}n^{k-q-h_Q}p^{k-q}\right) = o(1)\ ,$$

where the logarithmic factor represents the number of choices of the elements at which a handle is attached to the cycle. The spoiled simple paths can be classified into two types: those with at least one spoiling edge intersecting an edge of the path in at least two vertices, and the others. Let us denote their numbers by W_2 and W_1, respectively. We have $W = W_1 + W_2$ and, clearly, $W_2 > 0$ implies $Y > 0$. Thus we need to worry only about W_1. However, if a spoiling edge E intersects each edge of a simple path P in at most 1 vertex, then there is a subhypergraph consisting of a simple cycle C_1 (made by E and a segment of P between two consecutive intersections with E) and a simple path P_1 with its end-edges intersecting two consequtive edges of C_1, each in one vertex, but otherwise being disjoint from C_1. Let U count such configurations in $H(F_p, A)$. Thus, $W_1 > 0$ implies that $U > 0$ and we need to estimate $\mathbf{E}(U)$. Designating t_1 to represent the number of edges in C_1 and t_2 for the number of edges in P_1, we have, again by Proposition 2.2(i) and (ii),,

$$\mathbf{E}(U) = O\left(\sum_{t_1\geq 3}\sum_{t_2\geq 1} c^{t_1}(n^{k-l-1}p^{k-1})^{t_2-1}n^{k-l-2}p^{k-2}\right) = O(1/np) = o(1)\ .$$

Thus, by Markov's inequality, $\mathbf{P}(X = Y = Z = W = 0) \to 1$ as $n \to \infty$, which was to be proved. □

Acknowledgments. We would like to thank an anonymous referee and the Editor, Dr. Brenner, for many suggestions leading to an improvement of the exposition of the paper.

REFERENCES

1. B. Bollobás, *Random Graphs*, Academic Press, 1985.
2. P. Erdős and P. Turan, *On some sequences of integers*, J. London Math. Soc. **11** (1936), 261-264.
3. P. Frankl, R. Graham and V. Rödl, *Quantitative theorems for regular systems of equations*, J. Combin. Th., Series A **47(2)** (1988), 246-261.

4. R. Graham, V. Rödl and A. Ruciński, *On Schur properties of random subsets of integers*, to appear, Journal of Number Theory.
5. S. Janson, *Poisson approximation for large deviations*, Random Structures & Algorithms **1(2)** (1990), 221-229.
6. R. Rado, *Note on combinatorial analysis*, Proc. London Math. Soc. **48**, 122-160.
7. V. Rödl and A. Ruciński, *Threshold functions for Ramsey properties*, J. Amer. Math. Soc. **8(4)** (1995), 917-942.
8. I. Schur, *Über die Kongruenz* $x^m + y^m \equiv z^m (mod p)$, Jber. Deutsch. Mat. Verein. **25** (1916), 114-117.
9. E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arithmetica **27** (1975), 199-245.
10. B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Arch. Wisk. **15** (1927), 212-216.