A Note on Sparse Supersaturation and Extremal Results for Linear Homogeneous Systems

Christoph Spiegel*

Department of Mathematics
Universitat Politècnica de Catalunya
Barcelona, Spain

christoph.spiegel@upc.edu

Submitted: Jan 10, 2017; Accepted: Aug 9, 2017; Published: Aug 25, 2017
Mathematics Subject Classifications: 05D10, 05D40

Abstract

We study the thresholds for the property of containing a solution to a linear homogeneous system in random sets. We expand a previous sparse Szémeredi-type result of Schacht to the broadest class of matrices possible. We also provide a shorter proof of a sparse Rado result of Friedgut, Rödl, Ruciński and Schacht based on a hypergraph container approach due to Nenadov and Steger. Lastly we further extend these results to include some solutions with repeated entries using a notion of non-trivial solutions due to Rúzsa as well as Rué et al.

1 Introduction

A *k-term arithmetic progression* is a set of integers that can be written as $\{a, a+d, \dots, a+(k-1)d\}$ for some $a, d, k \in \mathbb{Z}$, $k \geqslant 3$ and $d \neq 0$. The Theorem of van der Waerden [26] states that every finite colouring of $[n] = \{1, \dots, n\}$ contains a monochromatic k-term arithmetic progression for n large enough. Szemerédi's Theorem [25] strengthened this result by stating that every set of integers with positive natural density contains a k-term arithmetic progression. Rado [17] generalized van der Waerden's result to certain systems of linear equations and Frankl, Graham and Rödl [6] did the same for Szemerédi's extremal result.

A common area of interest is to study 'sparse' or 'random' versions of such results. Consider the *binomial random set* $[n]_p$ where each element in $[n]$ is chosen independently with probability $p = p(n)$. That is $[n]_p$ is a random variable sampling from the finite probability space on all subsets of $[n]$ that assigns each $T \subseteq [n]$ the probability

*Supported by the Spanish Ministerio de Economía y Competitividad FPI grant under the project MTM2014-54745-P.

$\mathbb{P}([n]_p = T) = p^{|T|}(1-p)^{n-|T|}$. We say that some property holds *asymptotically almost surely* if the probability of it holding in $[n]_p$ converges to 1. Given an integer valued matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ with r rows and m columns, we largely follow the notation of Schacht [23] and let $\mathcal{S}(A) = \{\mathbf{x} \in \mathbb{Z}^m : A \cdot \mathbf{x}^T = \mathbf{0}^T\}$ be the set of all solutions and $\mathcal{S}_0(A) = \{\mathbf{x} = (x_1, \dots, x_m) \in \mathcal{S}(A) : x_i \neq x_j \text{ for } i \neq j\}$ the set of all *proper* solutions. We call A *irredundant* if $\mathcal{S}_0(A) \neq \varnothing$. Given a set of integers T and $s \in \mathbb{N}$ we write

$$T \to_s A \tag{1}$$

if for every finite partition $T_1 \dot\cup \dots \dot\cup T_s = T$ there exists $1 \leqslant i \leqslant s$ such that $T_i \cap \mathcal{S}_0(A) \neq \varnothing$. An irredundant matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ is *partition regular* if for every $s \in \mathbb{N}$ we have $[n] \to_s A$ for n large enough. Rado [17] gave the *column condition* as a characterization of partition regular matrices.

Rödl and Ruciński [18] formulated a sparse version of Rado's Theorem that was later completed by Friedgut, Rödl and Schacht [7]. To state it let $\varnothing \neq Q \subseteq [m]$ be any set of column indices and define $r_Q = \mathrm{rk}(A) - \mathrm{rk}(A^{\overline{Q}})$ where $A^{\overline{Q}}$ is the matrix obtained by keeping only the columns which are indexed by $\overline{Q}$ and the rank of a matrix A is denoted as $\mathrm{rk}(A)$. Here $A^{\varnothing}$ is the empty matrix with $\mathrm{rk}(A^{\varnothing}) = 0$. The *maximum* 1*-density* of a given matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ is defined as

$$m_1(A) = \max_{\substack{Q\subseteq[m]\\ 2\leqslant|Q|}} \frac{|Q|-1}{|Q|-r_Q-1}. \tag{2}$$

We will later see that for the specific kinds of matrices under consideration, this is indeed well-defined, that is $|Q| - r_Q - 1 > 0$ for all $Q \subseteq [m]$ satisfying $|Q| \geqslant 2$. This allows us to state the following sparse version of Rado's Theorem.

Theorem 1 (Rödl and Ruciński [18], Friedgut, Rödl and Schacht [7])**.** *For every $r, m, s \in \mathbb{N}$ and partition regular matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ there exist constants $c = c(A,s)$ and $C = C(A,s)$ such that*

$$\lim_{n\to\infty} \mathbb{P}([n]_p \to_s A) = \begin{cases} 0 & \text{if } p(n) \leqslant c\, n^{-1/m_1(A)},\\ 1 & \text{if } p(n) \geqslant C\, n^{-1/m_1(A)}. \end{cases}$$

The current proof of this theorem is quite involved. A first goal of this note is to provide a short proof of the 1-statement in Theorem 1 following the ideas of Nenadov and Steger's short proof of a sparse Ramsey Theorem [16]. This approach combines the recently developed hypergraph container framework by Balogh, Morris and Samotij [1] as well as Saxton and Thomason [22] with a supersaturation result of Frankl, Graham and Rödl [6].

Schacht [23] as well as independently Conlon and Gowers [2] also stated a sparse version of Szémeredi's Theorem. Schacht also extended it to density regular systems as well as sum-free sets, that is $A = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$. Given a set of integers T and $\epsilon > 0$ we write

$$T \to_\epsilon A \tag{3}$$

if every subset S for which $|S|/|T| \geqslant \epsilon$ satisfies $S \cap \mathcal{S}_0(A) \neq \varnothing$. An irredundant matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ is *density regular* if for all $\epsilon > 0$ we have $[n] \to_\epsilon A$ for n large enough. Frankl, Graham and Rödl [6] characterized density regular systems as those that are *invariant*, that is they satisfy $A \cdot \mathbf{1} = \mathbf{0}$. Lastly, we say that a matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ is *positive* if $\mathcal{S}(A) \cap \mathbb{N}^m \neq \varnothing$ and *abundant* if any $r \times (m-2)$ submatrix obtained from A by deleting two columns has the same rank as A. Every density regular system is clearly partition regular and partition regular systems are irredundant and positive by definition. We will later see in Lemma 6 that they are also abundant.

The second goal of this note is to extend Schacht's statement to the broadest group of matrices possible. To prove this generalization we derive a supersaturation result from a removal lemma due to Král', Serra and Vena [14] and combine it with a corollary of the hypergraph containers due to Balogh, Morris and Samotij [1].

Given some matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ let $\mathrm{ex}(n, A)$ be the size of the largest subset of $[n]$ not containing a proper solution and define $\pi(A) = \lim_{n\to\infty} \mathrm{ex}(n, A)/n$. Observe the clear parallels to the Turán number of a graph. Clearly density regular systems satisfy $\pi(A) = 0$ and for other systems systems we have $\pi(A) > 0$. One can easily bound this value away from 1, as we will later see in Lemma 5. These definitions and observations allow us to state the following sparse extremal result.

Theorem 2. *For every $r, m \in \mathbb{N}$ such that $m \geqslant 3$ and every irredundant, positive and abundant matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ and $\epsilon > \pi(A)$ there exist constants $c = c(A, \epsilon)$ and $C = C(A, \epsilon)$ such that*

$$\lim_{n\to\infty} \mathbb{P}\left([n]_p \to_\epsilon A\right) = \begin{cases} 0 & \text{if } p(n) \leqslant c\, n^{-1/m_1(A)}, \\ 1 & \text{if } p(n) \geqslant C\, n^{-1/m_1(A)}. \end{cases}$$

For singe-line equations it is common in combinatorial number theory to not just limit oneself to proper solutions, that is solutions with no repeated entries, but to also consider certain non-trivial solutions which may have some repeated entries. Ruzsa [21] gave a definition for non-trivial solutions in this scenario and more recently Rué, S. and Zumalacárregui [20] extended it to include arbitrary homogeneous linear systems of equations. A third and final goal will therefore be to extend the previously stated sparse results to include non-trivial solutions. We need to introduce some notation in order to give a formal definition of this notion.

Given a solution $\mathbf{x} = (x_1, \ldots, x_m) \in \mathcal{S}(A)$ let

$$p(\mathbf{x}) = \big\{\{1 \leqslant j \leqslant m : x_i = x_j\} : 1 \leqslant i \leqslant m\big\} \tag{4}$$

denote the set partition of the column indices $[m]$ indicating the repeated entries in $\mathbf{x}$. Note that for $\mathbf{x} \in \mathcal{S}_0(A)$ we have $p(\mathbf{x}) = \{\{1\}, \ldots, \{m\}\}$. Given some set partition p of $\{1, \ldots, m\}$, let A_p denote the matrix obtained by summing up the columns of A according to p, that is for $p = \{T_1, \ldots, T_s\}$ such that $\min(T_1) < \cdots < \min(T_s)$ for some $1 \leqslant s \leqslant m$ and $\mathbf{c}_i$ the i-th column vector of A for every $i \in [m]$, we have

$$A_p = \left(\sum_{i\in T_1} \mathbf{c}_i \;\middle|\; \sum_{i\in T_2} \mathbf{c}_i \;\middle|\; \cdots \;\middle|\; \sum_{i\in T_s} \mathbf{c}_i \right). \tag{5}$$

A solution $\mathbf{x} \in \mathcal{S}(A)$ is now defined to be *non-trivial* if $\mathrm{rk}(A_{p(\mathbf{x})}) = \mathrm{rk}(A)$. We denote the set of all non-trivial solutions by $\mathcal{S}_1(A) = \{\mathbf{x} \in \mathcal{S}(A) : \mathrm{rk}(A_{p(\mathbf{x})}) = \mathrm{rk}(A)\}$, that is we have $\mathcal{S}(A) \supseteq \mathcal{S}_1(A) \supseteq \mathcal{S}_0(A)$. Lastly, given a set of integers T, an integer $s \in \mathbb{N}$ and some $\epsilon > 0$, we write

$$T \to_s^\star A \tag{6}$$

if for every finite partition $T_1 \dot\cup \ldots \dot\cup T_s = T$ there exists $1 \leqslant i \leqslant s$ such that $T_i \cap \mathcal{S}_1(A) \neq \varnothing$ and

$$T \to_\epsilon^\star A \tag{7}$$

if every subset S for which $|S|/|T| \geqslant \epsilon$ also satisfies $S \cap \mathcal{S}_1(A) \neq \varnothing$. This is just a direct extension of the previous notation to include non-trivial solutions. We now have the following two statements.

Theorem 3. *For every $r, m, s \in \mathbb{N}$ and partition regular matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ there exists a constant $c = c(A, s)$ such that*

$$\lim_{n\to\infty} \mathbb{P}\left([n]_p \to_s^\star A\right) = 0 \quad \text{if } p(n) \leqslant c\, n^{-1/m_1(A)}.$$

Theorem 4. *For every $\epsilon > 0$, $r, m \in \mathbb{N}$ such that $m \geqslant 3$ and irredundant, positive and abundant matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ there exists a constant $c = c(A, \epsilon)$ such that*

$$\lim_{n\to\infty} \mathbb{P}\left([n]_p \to_\epsilon^\star A\right) = 0 \quad \text{if } p(n) \leqslant c\, n^{-1/m_1(A)}.$$

Observe that these results are stronger than the 0-statements of Theorem 1 and Theorem 2 and that their respective 1-statements supply matching counter-statements to Theorem 3 and Theorem 4. We will therefore only require proofs of the 1-statements of Theorem 1 and Theorem 2 as well as the 0-statements in Theorem 3 and Theorem 4.

Outline. In the remainder of this note we will first state some preliminaries in *Section 2* about linear systems of equations, their induced submatrices and supersaturation results as well as introduce hypergraph containers. We will then proceed by providing short proofs based on hypergraph container results of the 1-statements in Theorem 2 and Theorem 1 in *Sections 3* and *4* respectively. The 0-statements of Theorem 3 and Theorem 4 will be proven in *Sections 5* and *6* respectively.

Remark. Independent and simultaneous work of Hancock, Staden and Treglown [9] covers many of the same results presented here. In particular, they provide a strengthening of the 1-statement of Theorem 1 that also implies Theorem 2 as a sub-case. Their approach likewise consists of combining hypergraph containers with supersaturation results.

2 Preliminaries

Given some matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$, we have previously defined $\mathrm{ex}(n, A)$ to be the size of the largest subset of $[n]$ not containing a proper solution and $\pi(A) = \lim_{n\to\infty} \mathrm{ex}(n, A)/n$.

Erdős and Turan [5] determined that the size of the largest Sidon set, that is the case when $A = \begin{pmatrix}1 & 1 & -1 & -1\end{pmatrix}$, satisfies $\mathrm{ex}(n, A) = \Theta(\sqrt{n})$. For sum-free subsets, that is $A = \begin{pmatrix}1 & 1 & -1\end{pmatrix}$, it is also easy to see that $\pi(A) = 1/2$. Hancock and Treglown [10, 11] very recently extended this to matrices of the form $A = \begin{pmatrix}p & q & -r\end{pmatrix}$ where $p, q, r \in \mathbb{N}$ such that $p \geqslant q \geqslant r$. Unfortunately, unlike the Erdős-Stone-Simonovits Theorem [4, 3] in the graph case, no exact characterization of $\pi(A)$ is known for arbitrary matrices A. However, the following lemma shows that one can still easily bound this value away from 1 for every irredundant and positive matrix.

Lemma 5 (Folklore)**.** *Every irredundant and positive matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ satisfies $\pi(A) < 1$.*

Proof. Let $\mathbf{x} = (x_1, \dots, x_m) \in \mathcal{S}_0(A) \cap \mathbb{N}^m$. Clearly we also have $j \cdot \mathbf{x} = (jx_1, \dots, jx_m) \in \mathcal{S}_0(A) \cap \mathbb{N}^m$ for any $j \geqslant 1$. Now for $n \geqslant m \max_i(x_i)$ we observe that every $i \in [n]$ can appear in at most m of the $J = \lceil n/\max_i(x_i)\rceil$ solutions $\mathbf{x}, 2 \cdot \mathbf{x}, \dots, J \cdot \mathbf{x} \in [n]^m$, so every subset of $[n]$ that avoids $\mathcal{S}_0(A)$ is missing at least J/m elements. It follows that $\pi(A) \leqslant (n - J/m)/n \leqslant 1 - 1/(m \max_i(x_i)) < 1$. □

Partition and density regular matrices are irredundant and positive by definition. The next statement shows that Theorem 2 does indeed extend already existing results by proving that they are also abundant.

Lemma 6. *If a given $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ is partition or density regular, then it is abundant.*

Proof. Rado characterized partition regular matrices as those that satisfy the column condition, that is it is possible to re-order the column vectors $\mathbf{c}_1, \dots, \mathbf{c}_m$ of A, so that for some choice of indices $0 = m_0 < m_1 < \cdots < m_t = m$ setting $\mathbf{b}_i = \sum_{j=m_{i-1}+1}^{m_i} \mathbf{c}_j$ for $i = 1, \dots, t$ gives $\mathbf{b}_1 = \mathbf{0}$ and $\mathbf{b}_i$ can be expressed as a as a rational linear combination of $\mathbf{c}_1, \dots, \mathbf{c}_{m_{i-1}}$ for all $i \in \{2, \dots, t\}$.

Assume now that A is non-abundant, that is there exists a submatrix obtained by omitting two columns that has rank strictly smaller than $\mathrm{rk}(A)$. It follows that through basic row operations A can be transformed into a matrix of full rank whose last row contains only two non-zero entries $a, b \in \mathbb{Z} \setminus \{0\}$. As the matrix is partition regular, it is also irredundant and hence $a \neq -b$, that is $a + b \neq 0$. It follows that in order to satisfy the first requirement of the column condition, the columns need to be arranged such that there is a 0 in the last entry of the first column. However, there now must exist some $2 \leqslant i \leqslant t$ such that the last entry in $\mathbf{b}_i$ is non-zero while the last entries in $\mathbf{b}_1, \dots, \mathbf{b}_{i-1}$ are zero, violating the second requirement of the column condition. It follows that A must have been abundant. □

Let us consider some examples to illustrate these categories. $A = \begin{pmatrix}1 & -2 & 1\end{pmatrix}$, that is the matrix associated with 3-term arithmetic progression, is density regular by Roth's Theorem [19]. It therefore is trivially also partition regular, which was previously established

by van der Waerden [26]. The matrix associated with k-term arithmetic progressions

$$A = \begin{pmatrix} 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & & \ddots & \\ & & & 1 & -2 & 1 \end{pmatrix} \in \mathcal{M}_{(k-2)\times k} \tag{8}$$

is density regular and therefore abundant by Szémeredi's Theorem [25]. $A = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$ is not density regular, but by Schur's Theorem it is still partition regular. Lastly, $A = \begin{pmatrix} 1 & 1 & -r \end{pmatrix}$ for $r \in \mathbb{N} \setminus \{1, 2\}$ is neither partition nor density regular but it is abundant. For some more examples see [20].

2.1 Counting Solutions

We extend the notation from the introduction to inhomogeneous systems of linear equations. Given some matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ and column vector $\mathbf{b} \in \mathbb{Z}^r$ we write $\mathcal{S}(A, \mathbf{b}) = \{\mathbf{x} \in \mathbb{Z}^m : A \cdot \mathbf{x}^T = \mathbf{b}^T\}$, $\mathcal{S}_0(A, \mathbf{b}) = \{\mathbf{x} = (x_1, \dots, x_m) \in \mathcal{S}(A, \mathbf{b}) : x_i \neq x_j \text{ for } i \neq j\}$ and $\mathcal{S}_1(A, \mathbf{b}) = \{\mathbf{x} \in \mathcal{S}(A, \mathbf{b}) : \mathrm{rk}(A_{p(\mathbf{x})}) = \mathrm{rk}(A)\}$, so that $\mathcal{S}(A) = \mathcal{S}(A, \mathbf{0})$, $\mathcal{S}_0(A) = \mathcal{S}_0(A, \mathbf{0})$ and $\mathcal{S}_1(A) = \mathcal{S}_1(A, \mathbf{0})$. We remark that by elementary properties of systems of linear equations, we have the trivial upper bound

$$\big|\,\mathcal{S}_0(A, \mathbf{b}) \cap [n]^m\big| \leqslant \big|\,\mathcal{S}_1(A, \mathbf{b}) \cap [n]^m\big| \leqslant \big|\,\mathcal{S}(A, \mathbf{b}) \cap [n]^m\big| \leqslant n^{m-\mathrm{rk}(A)}. \tag{9}$$

The next lemma is due to Janson and Ruciński [13] and establishes a lower bound that matches this up to a constant for homogeneous systems. It can be easily extended to include non-homogeneous systems, see for example [15].

Lemma 7 (Janson and Ruciński [13])**.** *Let $r, m \in \mathbb{N}$ and a matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ be given. If $\mathcal{S}_0(A) \cap \mathbb{N}^m$ is non-empty then there exists a constant $c_0 = c_0(A) > 0$ such that*

$$|\,\mathcal{S}(A) \cap [n]^m| \geqslant |\,\mathcal{S}_1(A) \cap [n]^m| \geqslant |\,\mathcal{S}_0(A) \cap [n]^m| \geqslant c_0\, n^{m-rk(A)}. \tag{10}$$

In order to determine the exact asymptotic value of $|\,\mathcal{S}(A)\cap[n]^m|/n^{m-\mathrm{rk}(A)}$ or $|\,\mathcal{S}_0(A)\cap[n]^m|/n^{m-\mathrm{rk}(A)}$, one needs to employ Ehrhart's Theory, see for example Rué et al. [20].

Lastly, let $\mathcal{P}(A) = \{p(\mathbf{x}) : \mathbf{x} \in \mathcal{S}_1(A)\}$ denote the family of all set partitions of the column indices $[m]$ stemming from non-trivial solutions. The following lemma gives us the necessary tool to handle non-trivial solutions with repeated entries.

Lemma 8. *For every $r, m \in \mathbb{N}$, $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$, partition $p \in \mathcal{P}(A)$ and set $T \subset \mathbb{N}$ we have*

$$\big|\{\boldsymbol{x} \in \mathcal{S}_1(A) \cap T^m : p(\boldsymbol{x}) = p\}\big| \leqslant \big|\,\mathcal{S}_0(A_p) \cap T^{|p|}\big|. \tag{11}$$

Proof. Write $p = \{T_1, \dots, T_s\}$ for some $1 \leqslant s \leqslant m$ such that $\min(T_1) < \cdots < \min(T_s)$. Let $Q = \{\min(T_1), \dots, \min(T_s)\}$. Now for every $\mathbf{x} = (x_1, \dots, x_m) \in \mathcal{S}_1(A) \cap T^m$ such that $p(\mathbf{x}) = p$, we would have $\mathbf{x}^Q = (x_{\min(T_1)}, \dots, x_{\min(T_S)}) \in T^{|p|}$ as well as $\mathbf{x}^Q \in \mathcal{S}(A_p)$ as can be readily seen by the definition of A_p. Since $p = p(\mathbf{x})$, the vector $\mathbf{x}^Q$ would furthermore be proper, so that $\mathbf{x}^Q \in \mathcal{S}_0(A_p) \cap T^{|p|}$. The map $\{\mathbf{x} \in \mathcal{S}_1(A) : p(\mathbf{x}) = p\} \cap T^m \to \mathcal{S}_0(A_p) \cap T^{|p|}$, $\mathbf{x} \mapsto \mathbf{x}^Q$ is clearly injective, proving the desired statement. □

An easy corollary of this result is clearly that if there are no proper solutions to the system A_p in a set, then there can also not be non-trivial solutions to A whose repetitions are indicated by p.

2.2 Induced Submatrices

The notion of induced submatrices was originally introduced by Rödl and Ruciński [18] – though not explicitly referred to as such – while developing a sparse version of Rado's Partition Theorem. Recall the definitions from the introduction, especially $r_Q = \mathrm{rk}(A) - \mathrm{rk}(A^{\overline{Q}})$. Observe that we can without loss of generality assume that A is of full rank for this part, since the solution space is unaffected by this assumption. This will simplify notation significantly.

For a given matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ and column indices $\varnothing \subseteq Q \subseteq [m]$, we will now construct through basic row operations a matrix that tries to encapsulate the information contained in A through the columns indexed by Q. Denote the rows of A by $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ so that the rows of A^Q and $A^{\overline{Q}}$ are respectively $\mathbf{a}_1^Q, \mathbf{a}_2^Q, \ldots, \mathbf{a}_r^Q$ and $\mathbf{a}_1^{\overline{Q}}, \mathbf{a}_2^{\overline{Q}}, \ldots, \mathbf{a}_r^{\overline{Q}}$. Here we allow for empty vectors and matrices. If $\mathrm{rk}(A^{\overline{Q}}) < \mathrm{rk}(A)$, then we can express exactly $r_Q > 0$ of the r rows of $A^{\overline{Q}}$ as linear combinations of the rest, that is there are indices $i_1 < \cdots < i_{r_Q} \in [m]$ and integers $d_i, d_i^j \in \mathbb{Z}$ for $i \in \{i_1, \ldots, i_{r_Q}\}$ and $j \in [m] \setminus \{i_1, \ldots, i_{r_Q}\}$ so that

$$d_i\, \mathbf{a}_i^{\overline{Q}} = \sum_{j\in[m]\setminus\{i_1,\ldots,i_{r_Q}\}} d_i^j\, \mathbf{a}_j^{\overline{Q}} \qquad \text{for } i \in \{i_1, \ldots, i_{r_Q}\}. \tag{12}$$

Consider now the following integer-valued matrix with r_Q rows and $|Q|$ columns

$$A[Q] = \begin{pmatrix} d_{i_1}\, \mathbf{a}_{i_1}^Q - \sum\limits_{j\in[m]\setminus\{i_1,\ldots,i_{r_Q}\}} d_{i_1}^j\, \mathbf{a}_j^Q \\ \vdots \\ d_{i_{r_Q}} \mathbf{a}_{i_{r_Q}}^Q - \sum\limits_{j\in[m]\setminus\{i_1,\ldots,i_{r_Q}\}} d_{i_{r_Q}}^j\, \mathbf{a}_j^Q \end{pmatrix} \in \mathcal{M}_{r_Q\times|Q|}(\mathbb{Z}). \tag{13}$$

To illustrate this construction further, note that if we assume that the column indices are appropriately ordered, that is $Q = \{1, \ldots, |Q|\}$, then the matrix A (without the assumption of being of full rank) can be rewritten as

$$B = \left(\begin{array}{cc} \multicolumn{2}{c}{\multirow{2}{*}{$A[Q]$}} & \left.\vphantom{\begin{matrix}a\\a\end{matrix}}\right] \mathrm{rk}(A) - r_Q \\ & & \left.\vphantom{\begin{matrix}a\\a\end{matrix}}\right] r_Q \\ \mathbf{0} & \mathbf{0} & \left.\vphantom{\begin{matrix}a\\a\end{matrix}}\right] r - \mathrm{rk}(A) \end{array}\right) \tag{14}$$

through elementary row operations, that is $A = P_1^{-1} \cdot B$ where $P_1 \in \mathcal{M}_{r\times r}$ is an invertible rectangular matrix. We have $\mathcal{S}(B) = \mathcal{S}(A)$, that is the homogeneous solution space remains unchanged, at least up to the column permutation necessary to ensure that $Q = \{1, \ldots, |Q|\}$.

Observe that the matrix $A[Q]$ is only well defined up to our choices of indices i_k and coefficients d_i, d_i^j. However, the homogeneous solution space $\mathcal{S}(A[Q])$ is independent

of these, so we pick one representative for each $\varnothing \subseteq Q \subseteq [m]$ and refer to it as the *submatrix of A induced by Q*. The notation $A[Q]$ will refer to this particular representative. We state the following simple observations, that are immediately clear by considering Equation (14).

Remark 9. $A[Q]$ is of full rank, that is $\mathrm{rk}(A[Q]) = r_Q$ for any $Q \subseteq [m]$ satisfying $r_Q > 0$. If A was irredundant, positive or abundant, then $A[Q]$ also fulfils these properties for any $Q \subseteq [m]$ such that $r_Q > 0$.

The following lemma now establishes some results regarding the rank of the induced submatrices of an abundant matrices. It also verifies that the maximum 1-density parameter given in the introduction is indeed well-defined for abundant matrices.

Lemma 10 (Kusch et al. [15])**.** *For any $r, m \in \mathbb{N}$, abundant matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ and selection of column indices $Q \subseteq [m]$ the following holds. If $|Q| \geqslant 2$ then we have $|Q| - r_Q - 1 > 0$, that is the parameter $m_1(A)$ is well-defined. If $|Q| \leqslant 2$ then we have $r_Q = 0$.*

The next lemma is crucial and establishes that a lack of non-trivial solutions to an induced submatrix of A also implies a lack of non-trivial solutions to the full matrix. Note that this was previously proven by Rödl and Ruciński for proper solutions [18].

Lemma 11. *For any $r, m \in \mathbb{N}$, matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ and set $T \subset \mathbb{N}$ the following holds. If there exists a selection of column indices $Q \subseteq [m]$ such that $r_Q > 0$ and $\mathcal{S}_1(A[Q]) \cap T^{|Q|} = \varnothing$ then $\mathcal{S}_1(A) \cap T^m = \varnothing$.*

Proof. We prove the statement by showing that $\mathbf{x} \in \mathcal{S}_1(A) \cap T^m$ would imply $\mathbf{x}^Q \in \mathcal{S}_1(A[Q]) \cap T^{|Q|}$, giving a contradiction. It is easy to see through Equation (14) and the definition of $A[Q]$ that $\mathbf{x} \in \mathcal{S}_1(A) \cap T^m$ implies $\mathbf{x}^Q \in \mathcal{S}(A[Q]) \cap T^{|Q|}$, so let us show that $\mathbf{x}^Q$ is non-trivial for $A[Q]$. Let $\phi_Q : Q = \{i_1 < \cdots < i_{|Q|}\} \to \{1, \ldots, |Q|\},\ i_j \mapsto j$ be the bijective map canonically mapping the indices in $Q \subseteq [m]$ to the first $|Q|$ integers. Writing $p(\mathbf{x}) = \{T_1, \ldots, T_t\}$ for some $1 \leqslant t \leqslant m$ we define $p' = \{\phi_Q(T_1 \setminus \overline{Q}), \ldots, \phi_Q(T_t \setminus \overline{Q})\} \setminus \{\varnothing\}$. Observe that in fact $p' = p(\mathbf{x}^Q)$. Equation (14) illustrates that $\mathrm{rk}(A) - \mathrm{rk}(A_{p(\mathbf{x})}) \geqslant \mathrm{rk}(A[Q]) - \mathrm{rk}(A[Q]_{p'})$. As $\mathbf{x}$ was assumed to be non-trivial, that is $\mathrm{rk}(A) = \mathrm{rk}(A_{p(\mathbf{x})})$, and since we have previously observed that $p' = p(\mathbf{x}^Q)$, we therefore have $\mathrm{rk}(A[Q]) \leqslant \mathrm{rk}(A[Q]_{p'}) = \mathrm{rk}(A[Q]_{p(\mathbf{x}^Q)})$ so that the desired statement follows. □

We end our observations about induced submatrices by stating the following easy proposition. It covers some trivial cases not considered by Theorem 4 and will in fact be needed later in the proof of it. For a reference for the inequalities of Markov, Chernoff or Chebyshev, see for example [12].

Proposition 12. *For every $r, m \in \mathbb{N}$ such that $m \geqslant 2$, irredundant an positive bot not abundant matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ and $\epsilon > 0$ we have $\lim_{n\to\infty} \mathbb{P}\left([n]_p \to^\star_\epsilon A\right) = 0$ for any $p(n) = o(1)$.*

Proof. Since A is not abundant but positive and irredundant, there exists some $Q \subseteq [m]$ satisfying $|Q| = 2$ such that $A[Q] = (\ a \quad -b\ )$ for some $a, b \in \mathbb{N}$, $a \neq b$. By Lemma 11 we can replace A with $A[Q]$. It follows by Equation (9), Lemma 7 as well as the linearity of expectation that $\mathbb{E}\big(|\mathcal{S}(A) \cap [n]_p^2|\big) = \Theta(n\,p^2)$ while $\mathbb{E}(|[n]_p|) = np$. If $np = O(1)$ then $\mathbb{E}\big(|\mathcal{S}(A) \cap [n]_p^2|\big) = o(1)$ and the result trivally holds by Markov's Inequality. If $np \to \infty$ then by Chernoff $|[n]_p| \geqslant np/2$ asymptotically almost surely. Since $p = o(1)$ we have $\mathbb{E}\big(|\mathcal{S}(A) \cap [n]_p^2|\big) = o(np/2)$ and therefore for any given set of positive density, we can remove one element per solution and still asymptotically almost surely have a solution-free set of that same density. This proves the desired result. □

2.3 Removal Lemma and Supersaturation Results

A common ingredient to proving sparse results are robust versions of the deterministic statement, referred to as supersaturation results. In the graph setting such a result is folklore and easy to prove. A number theoretical counterpart is Varnavides [27] robust version of Szemerédi's Theorem which states that a set of positive density contains not just one, but a positive proportion of all k-term arithmetic progressions. Frankl, Graham and Rödl [6] formulated such results both for partition and density regular systems.

Lemma 13 (Theorem 1 in Frankl, Graham and Rödl [6])**.** *For a given partition regular matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ and $s \in \mathbb{N}$ there exists $\zeta = \zeta(A, s) > 0$ such that for any partition $[n] = T_1 \dot\cup \ldots \dot\cup T_s$ and n large enough we have $|\mathcal{S}_0(A) \cap T_1^m| + \cdots + |\mathcal{S}_0(A) \cap T_s^m| \geqslant \zeta\,|\,\mathcal{S}_0(A) \cap [n]^m|$.*

Lemma 14 (Theorem 2 in Frankl, Graham and Rödl [6])**.** *For a given density regular matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ and $\delta > 0$ there exists $\zeta = \zeta(A, \delta) > 0$ such that any subset $T \subseteq [n]$ satisfying $|T| \geqslant \delta n$ contains at least $\zeta\,|\,\mathcal{S}_0(A) \cap [n]^m|$ proper solutions for n large enough.*

We will extend Lemma 14 to cover the scope of this note by using an *arithmetic removal lemma*. Green [8] first formulated such a statement for linear equations in an abelian group. Later Shapira [24] as well as independently Král', Serra and Vena [14] proved a removal lemma for linear maps in finite fields. We will state it here in a simplified version.

Theorem 15 (Removal Lemma [14])**.** *Let $\mathbb{F}_q$ be the finite field of order q. Let X be a subset of $\mathbb{F}_q$ and $A \in \mathcal{M}_{r\times m}(\mathbb{F}_q)$ a matrix of full rank. For $\mathcal{S} = \{\boldsymbol{x} \in \mathbb{F}_q^m : A \cdot \boldsymbol{x}^T = \mathbf{0}^T\}$ and every $\epsilon > 0$ there exists an $\eta = \eta(\epsilon, r, m)$ such that if $|\mathcal{S} \cap X^m| < \eta\,|\mathcal{S}|$ then there exists a set $X' \subset X$ with $|X'| < \epsilon q$ and $\mathcal{S} \cap (X \setminus X')^m = \varnothing$.*

Applying this result, we formulate the following extension of Lemma 14. Note that one could also obtain this result through a direct application of a removal lemma for colored hypergraphs as for example Theorem 2 in [14].

Lemma 16 (Supersaturation)**.** *For a given $r, m \in \mathbb{N}$, positive and irredundant matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ and $\delta > \pi(A)$ there exists $\zeta = \zeta(\delta, A) > 0$ such that any subset $T \subseteq [n]$ satisfying $|T| \geqslant \delta n$ contains at least $\zeta\,|\,\mathcal{S}_0(A) \cap [n]^m|$ proper solutions for n large enough.*

Proof. Let $q = q(A, n)$ be a prime number between $2mn \max(|A|)$ and $4mn \max(|A|)$ and $\mathbb{F}_q$ the finite field with q elements. Here $\max(|A|)$ refers to the maximal absolute entry in A. Note that such a prime number exists for example because of the Bertrand-Chebyshev Theorem. We have $\mathbb{F}_q \cong \mathbb{Z}_q$ and we can identify the integers with their corresponding residue classes in $\mathbb{F}_q$. The matrix A now defines a map from $\mathbb{F}_q^m$ to $\mathbb{F}_q^r$. A solution in $\mathcal{S}(A)$ clearly lies in the $\mathcal{S}$ and, as we have chosen q large enough, all canonical representatives from $\mathcal{S} \cap [n]^m$ also lie in $\mathcal{S}(A) \cap [n]^m$ for $n \geqslant \max |A|$.

Next, set $\delta' = (\delta + \pi(A))/2$ and let n be large enough such that any subset of density at least δ' in $[n]$ contains a proper solution. Note that $\delta > \delta' > \pi(A)$. Given a subset $T \subseteq [n]$ satisfying $|T| \geqslant \delta n$ consider the corresponding set X of residue classes in $\mathbb{F}_q$. One needs to remove at least $(\delta - \delta')n$ elements from T in order for T^m to avoid $\mathcal{S}_0(A)$ in $[n]$, so one needs to remove at least an

$$\epsilon = \frac{(\delta - \delta')n}{q} \geqslant \frac{(\delta - \delta')}{4m \max(|A|)} > 0$$

proportion of elements in $\mathbb{F}_q$ from X so that X^m avoids $\mathcal{S}$ in $\mathbb{F}_q$. It follows from Theorem 15 that $|\mathcal{S} \cap X^m| \geqslant \eta |\mathcal{S}|$ for some $\eta = \eta(\epsilon, r, m)$. Since we have chosen q large enough, it follows that T contains at least an η proportion of $\mathcal{S}(A) \cap [n]^m$. An easy consequence of Equation (9) and Lemma 7 is that $\lim_{n\to\infty} |\mathcal{S}_0(A) \cap [n]^m| / |\mathcal{S}(A) \cap [n]^m| \geqslant c_0$ for $c_0 = c_0(A) > 0$ as given by Lemma 7. It follows result holds for n large enough and $\zeta = \zeta(\delta, A) = (c_0\, \eta)/2$. □

2.4 Hypergraph Containers

The development of *hypergraph containers* by Balogh, Morris and Samotij [1] as well as independently Thomason and Saxton [22] has opened a new, easy and unified framework to proving sparse results. Let us start by stating the Hypergraph Container Theorem as given by Balogh, Morris and Samotij.

Given a hypergraph $\mathcal{H}$ we denote its vertex set by $V(\mathcal{H})$ and its set of hyperedges by $E(\mathcal{H})$. The cardinality of these sets will be respectively denoted by $v(\mathcal{H})$ and $e(\mathcal{H})$. Given some subset of vertices $A \subseteq V(\mathcal{H})$ we denote the subgraph it induces in $\mathcal{H}$ by $\mathcal{H}[A]$ and its degree by $\deg_{\mathcal{H}}(A) = |\{e \in E(\mathcal{H}) : A \subseteq e\}|$. For $\ell \in \mathbb{N}$ we denote the *maximum ℓ-degree* by $\Delta_\ell(\mathcal{H}) = \max\{\deg_{\mathcal{H}}(A) : A \subseteq V(\mathcal{H}) \text{ and } |A| = \ell\}$. Let the set of independent vertex sets in $\mathcal{H}$ be denoted by $\mathcal{I}(\mathcal{H})$. Lastly, let $\mathcal{H}$ be a uniform hypergraph, $\mathcal{F}$ an increasing family of subsets of $V(\mathcal{H})$ and $\epsilon > 0$. We say that $\mathcal{H}$ is $(\mathcal{F}, \epsilon)$-*dense* if $e(\mathcal{H}[A]) \geqslant \epsilon\, e(\mathcal{H})$ for every $A \in \mathcal{F}$.

Theorem 17 (Hypergraph Containers, Theorem 2.2 in [1])**.** *For every $m \in \mathbb{N}$, $c > 0$ and $\epsilon > 0$, there exists a constant $C = C(m, c, \epsilon) > 0$ such that the following holds. Let $\mathcal{H}$ be an m-uniform hypergraph and let $\mathcal{F} \subseteq 2^{V(\mathcal{H})}$ be an increasing family of sets such that $|A| \geqslant \epsilon v(\mathcal{H})$ for all $A \in \mathcal{F}$. Suppose that $\mathcal{H}$ is $(\mathcal{F}, \epsilon)$-dense and $p \in (0, 1)$ is such that, for every $\ell \in \{1, \ldots, k\}$,*

$$\Delta_\ell(\mathcal{H}) \leqslant c\, p^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})}. \tag{15}$$

Then there exists a family $\mathcal{T} \subseteq \binom{V(\mathcal{H})}{\leqslant Cp\,v(\mathcal{H})}$ *and functions* $f : \mathcal{T} \to \overline{\mathcal{F}}$ *and* $g : \mathcal{I}(\mathcal{H}) \to \mathcal{T}$ *such that for every* $I \in \mathcal{I}(\mathcal{H})$*,*

$$g(I) \subseteq I \quad \text{and} \quad I \setminus g(I) \subseteq f(g(I)). \tag{16}$$

The statement gives the existence of a small number of *containers* $\overline{\mathcal{F}}$ and some *fingerprints* $\mathcal{T}$ so that every independent set I in $\mathcal{H}$ is identified with a fingerprint $g(I)$ that determines a container $f(g(I))$ which contains $I \setminus g(I)$.

Next, let $\mathcal{H} = (\mathcal{H}_n)_{n\in\mathbb{N}}$ be a sequence of m-uniform hypergraphs and let $\alpha \in [0, 1)$. We say that $\mathcal{H}$ is *α-dense* if for every $\delta > 0$, there exist some $\epsilon > 0$ such that for $U \subseteq V(\mathcal{H}_n)$ which satisfies $|U| > (\alpha + \delta)\, v(\mathcal{H}_n)$ we have $e(\mathcal{H}_n[U]) > \epsilon\, e(\mathcal{H}_n)$ for n large enough. Balogh, Morris and Samotij proved the following consequence of their container statement.

Theorem 18 (Sparse Sets through Hypergraph Containers, Theorem 5.2 in [1])**.** *Let* $\mathcal{H} = (\mathcal{H}_n)_{n\in\mathbb{N}}$ *be a sequence of* m*-uniform hypergraphs,* $\alpha \in [0, 1)$ *and let* $C > 0$*. Suppose that* $q = q(n)$ *is a sequence of probabilities such that for all sufficiently large* n *and for every* $\ell \in \{1, \dots, m\}$ *we have*

$$\Delta_\ell(\mathcal{H}_n) \leqslant C\, q(n)^{\ell-1}\, \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}. \tag{17}$$

If $\mathcal{H}$ *is* α*-dense, then for every* $\delta > 0$*, there exists a constant* $c = c(C, \alpha, m) > 0$ *such that if* $p(n) > c\, q(n)$ *and* $p(n)\, v(\mathcal{H}_n) \to \infty$ *as* $n \to \infty$*, then asymptotically almost surely*

$$\alpha\big(\mathcal{H}_n[V(\mathcal{H}_n)_{p(n)}]\big) \leqslant (\alpha + \delta)\, p(n)\, v(\mathcal{H}_n). \tag{18}$$

We will make use of this statement in order to obtain a proof for the 1-statement of Theorem 2. For a proof of the 1-statement of Theorem 1 such a ready-made statement does not exist and we will follow Nenadov and Steger's [16] short proof of a sparse Ramsey statement by applying Theorem 17.

3 Proof of the 1-statement in Theorem 2

Let $\mathcal{H}_n$ be the hypergraph with vertex set $V(\mathcal{H}_n) = [n]$ and edge multiset

$$E(\mathcal{H}_n) = \big\{\big\{\{x_1, \dots, x_m\} : (x_1, \dots, x_m) \in \mathcal{S}_0(A) \cap [n]^m\big\}\big\}.$$

Observe that H_n can be a multigraph, that is multiple edges are allowed, but the multiplicity of each edge is clearly bounded by $m!$. We do this to simplify counting, since this way we have $|E(\mathcal{H}_n)| = |\,\mathcal{S}_0(A) \cap [n]^m|$. We observe that we can limit ourselves to proper solutions when proving the 1-statement.

Corollary 16 now states that $\mathcal{H} = (\mathcal{H}_n)_{n\in\mathbb{N}}$ is $\pi(A)$-dense. The statement of Theorem 2 follows if Theorem 18 can be applied. In order to apply Theorem 18, it remains to determine a sequence $q = q(n)$ satisfying the required condition. The following lemma gives us upper bounds for the maximum ℓ-degrees in $\mathcal{H}_n$.

Lemma 19. *For $1 \leqslant \ell \leqslant m$ we have $\Delta_\ell(\mathcal{H}_n) \leqslant \ell!\, m^\ell \max_{Q\subseteq[m],\,|Q|=\ell} n^{(m-rk(A))-(|Q|-r_Q)}$.*

Proof. For $\mathcal{H} = (\mathcal{H}_n)$ as defined above and $\ell \in \{1,\dots,m\}$ we have

$$\begin{aligned}
\Delta_\ell(\mathcal{H}_n) &\leqslant \max_{x_1,\dots,x_\ell\in[n]} \left|\{\mathbf{x}\in\mathcal{S}_0(A)\cap[n]^m : \exists Q\subseteq[m], \pi\in S(\ell) \text{ s.t. } \mathbf{x}^Q = (x_{\pi(1)},\dots,x_{\pi(\ell)})\}\right| \\
&\leqslant \ell! \binom{m}{\ell} \max_{\substack{(x_1,\dots,x_\ell)\in[n]^\ell \\ Q\subseteq[m],|Q|=\ell}} \left|\{\mathbf{x}\in[n]^{m-l} \mid A^{\overline{Q}}\cdot\mathbf{x}^T = -A^Q\cdot(x_1,\dots,x_\ell)^T\}\right| \\
&\leqslant \ell!\, m^\ell \max_{\substack{Q\subseteq[m]\\|Q|=\ell}} \max_{\mathbf{b}\in\mathbb{Z}^r} \left|\,\mathcal{S}(A^{\overline{Q}},\mathbf{b})\cap[n]^m\right| \leqslant \ell!\, m^\ell \max_{\substack{Q\subseteq[m]\\|Q|=l}} n^{|\overline{Q}|-\mathrm{rk}(A^{\overline{Q}})} \\
&= \ell!\, m^\ell \max_{\substack{Q\subseteq[m]\\|Q|=l}} n^{(m-\mathrm{rk}(A))-(|Q|-r_Q)}
\end{aligned}$$

where $S(\ell)$ denotes the set of permutations of ℓ elements. We have also made extensive use of the notation defined in the introduction as well as as the trivial upper bound for the number of solutions stated in Equation (9). □

Note that $r_Q = 0$ for any $Q\subseteq[m]$ satisfying $|Q| = 1$ due to Lemma 10 and that there exists $c_0 = c_0(A) > 0$ such that $e(\mathcal{H}_n) \geqslant c_0\, n^{m-\mathrm{rk}(A)}$ due to Lemma 7. Using Lemma 19 we now observe that

$$\Delta_1(\mathcal{H}_n) \leqslant m\, n^{m-\mathrm{rk}(A)-1} \leqslant m/c_0\, \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}.$$

For $\ell\in\{2,\dots,m\}$ we again apply Lemma 19 to see that

$$\begin{aligned}
\Delta_\ell(\mathcal{H}_n) &\leqslant \ell!\, m^\ell \max_{Q\subseteq[m],\,|Q|=\ell} n^{(m-\mathrm{rk}(A))-(|Q|-r_Q)} \\
&= \ell!\, m^\ell \left(\max_{Q\subseteq[m],\,|Q|=\ell} n^{-\frac{|Q|-r_Q-1}{|Q|-1}}\right)^{\ell-1} n^{m-\mathrm{rk}(A)-1} \\
&\leqslant \ell!\, m^\ell \left(n^{-1/m_1(A)}\right)^{\ell-1} n^{m-\mathrm{rk}(A)-1} \\
&\leqslant (\ell!\, m^\ell)/c_0 \left(n^{-1/m_1(A)}\right)^{\ell-1} \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}.
\end{aligned}$$

Lastly we observe that $n^{-1/m_1(A)}\, v(\mathcal{H}_n) = n^{1-1/m_1(A)} \to \infty$ as $m_1(A) > 1$. It follows that the prerequisites of Theorem 18 hold for $C = (m!\, m^m)/c_0$, $q = q(n) = n^{-1/m_1(A)}$ and we can choose the $c = c(A,\epsilon)$ in Theorem 2 to be equal to the $c = c(C,\pi(A),m)$ as given by Theorem 18.

4 A Short Proof of the 1-statement in Theorem 1

As stated in the introduction, this result was previously proven by Friedgut, Rödl and Schacht [7] as well as independently Conlon and Gowers [2]. This proof merely serves as

a short version that follows the short proof of a sparse Ramsey result due to Nenadov and Steger [16].

We will need two ingredients in order to prove the 1-statement of Theorem 1. The first will be the following easy corollary to Lemma 13.

Corollary 20. *For a given partition regular matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ and $s \in \mathbb{N}$ there exist $\epsilon = \epsilon(A,s)$ and $\delta = \delta(A,s) > 0$ such that for any $T_1,\dots,T_s \subseteq [n]$ satisfying $|\mathcal{S}_0(A)\cap T_i^m| \leqslant \epsilon\,|\mathcal{S}_0(A)\cap[n]^m|$ for $1 \leqslant i \leqslant s$ we have $\big|[n]\setminus(T_1\cup\dots\cup T_s)\big| \geqslant \delta n$ for n large enough.*

Proof. Let $\zeta = \zeta(A,s+1)$ be as in Lemma 13 and $\epsilon = \epsilon(A,s) = \zeta/2s$. Set $\tilde{T}_i = T_i\setminus\bigcup_{j=1}^{i-1} T_j$ for $1 \leqslant i \leqslant s$ and $\tilde{T}_{s+1} = [n]\setminus\bigcup_{j=1}^{r} T_j$ and consider the partition $[n] = \tilde{T}_1\,\dot\cup\,\dots\,\dot\cup\,\tilde{T}_s\,\dot\cup\,\tilde{T}_{s+1}$. By Lemma 13 we have $|\mathcal{S}_0(A)\cap\tilde{T}_1^m| + \dots + |\mathcal{S}_0(A)\cap\tilde{T}_{s+1}^m| \geqslant \zeta\,|\mathcal{S}_0(A)\cap[n]^m|$ and since by assumption $|\mathcal{S}_0(A)\cap\tilde{T}_i^m| \leqslant |\mathcal{S}_0(A)\cap T_i^m| \leqslant \zeta/2s\,|\mathcal{S}_0(A)\cap[n]^m|$ for all $i\in\{1,\dots,s\}$, we have $|\mathcal{S}_0(A)\cap([n]\setminus(T_1\cup\dots\cup T_s))^m| \geqslant \zeta/2\,|\mathcal{S}_0(A)\cap[n]^m|$. Observe that by Lemma 19 every element in $[n]$ is contained in at most $m\,n^{m-\mathrm{rk}(A)-1}$ solutions and by Lemma 7 there exists $c_0 = c_0(A) > 0$ such that $|\mathcal{S}_0(A)\cap[n]^m| \geqslant c_0\,n^{m-\mathrm{rk}(A)}$ for n large enough, so that the results follows for $\delta = \zeta c_0/2$. □

The second ingredient is stated in the following corollary that is obtained by applying the Hypergraph Container Theorem to the hyperpgraph of solutions.

Corollary 21. *For a given partition regular matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ and $\epsilon > 0$ there exist $t = t(n)$ sets $T_1,\dots,T_t \in \binom{[n]}{\leqslant c_0\,n^{1-1/m_1(A)}}$ for some $c_0 > 0$ as well as sets $C_1,\dots,C_t \subseteq [n]$ such that*

$$|\mathcal{S}_0(A)\cap C_i^m| \leqslant \epsilon\,|\mathcal{S}_0(A)\cap[n]^m|. \tag{19}$$

Furthermore, for every set $T\subseteq[n]$ satisfying $\mathcal{S}_0(A)\cap T^m = \varnothing$ there exists $1\leqslant i\leqslant t$ such that

$$T_i \subseteq T \subseteq C_i. \tag{20}$$

Proof. Let $\mathcal{H}_n$ again be the hypergraph with vertex set $V(\mathcal{H}_n) = [n]$ and edge multiset

$$E(\mathcal{H}_n) = \big\{\big\{\{x_1,\dots,x_m\} : (x_1,\dots,x_m)\in\mathcal{S}_0(A)\cap[n]^m\big\}\big\}.$$

We have previously observed that there exists a $c > 0$ such that for $1\leqslant\ell\leqslant m$ we have $\Delta_\ell \leqslant c\,p(n)^{\ell-1}\,e(\mathcal{H}_n)/v(\mathcal{H}_n)$ for $p = p(n) = n^{-1/m_1(A)}$. We observe that $\mathcal{H}_n$ is trivially $(\mathcal{F},\epsilon)$-dense for $\mathcal{F} = \{T\subseteq[n] : |\mathcal{S}_0(A)\cap T^m| \geqslant \epsilon\,|\mathcal{S}_0(A)\cap[n]^m|\}$. Applying Theorem 17 gives the desired statement. □

We are now ready to give a short proof of the 1-statement in Theorem 1 following the ideas of Nenadov and Steger [16]. Let $\epsilon,\delta > 0$ be as in Corollary 20 and let $t = t(n)$, c_0, $S_1,\dots,S_t$ and $C_1,\dots,C_t$ be as in Corollary 21. Let $C = C(A,s)$ be large enough such that

$$\left(1 + \ln\left(\frac{2^s}{sc_0}\right) + \ln(C)\right)\frac{sc_0}{C} < \frac{1}{2}.$$

Observe now that for a partition of the random set $T_1 \dot\cup \ldots \dot\cup T_s = [n]_p$ satisfying $\mathcal{S}_0(A) \cap T_i^m = \varnothing$ for all $i \in \{1,\ldots,s\}$ there exist $j_1,\ldots,j_s \in \{1,\ldots,t\}$ so that $S_{j_i} \subseteq T_i \subseteq C_{j_i}$ for all $i \in \{1,\ldots,s\}$. Since $T_i \subseteq [n]_p$ for $1 \leqslant i \leqslant s$ and $[n] \setminus (C_1 \cup \cdots \cup C_s) \cap [n]_p = \varnothing$ we can bound the probability of $[n]_p$ not fulfilling the partition property by

$$\mathbb{P}\left([n]_p \not\to_s A\right) \leqslant \sum_{j_1,\ldots,j_s \in \{1,\ldots,t\}} \mathbb{P}\left(S_{j_1},\ldots,S_{j_s} \subseteq [n]_p \,\wedge\, [n] \setminus (C_{j_1},\ldots,C_{j_s}) \cap [n]_p = \varnothing\right).$$

Observe that the two events $S_{j_1},\ldots,S_{j_s} \subseteq [n]_p$ and $[n] \setminus (C_{j_1},\ldots,C_{j_s}) \cap [n]_p = \varnothing$ are independent, so that we have

$$\mathbb{P}\left([n]_p \not\to_s A\right) \leqslant \sum_{j_1,\ldots,j_s \in \{1,\ldots,t\}} p^{\left|\bigcup_{j=1}^s S_j\right|} (1-p)^{\left|[n]\setminus(C_{j_1},\ldots,C_{j_s})\right|}.$$

We bound this by choosing $k = |\bigcup_{j=1}^s S_j| \leqslant sc_0 n^{1-1/m_1(A)}$, then picking k elements and lastly deciding for each element in this selection in which of the S_i it is contained, so that we have

$$\mathbb{P}\left([n]_p \not\to_s A\right) \leqslant (1-p)^{\delta n} \sum_{k=0}^{sc_0 n^{1-1/m_1(A)}} \binom{n}{k} (2^s)^k p^k \leqslant e^{-\delta np} \left(1 + \sum_{k=1}^{sc_0 C^{-1} np} \left(\frac{e2^s\, np}{k}\right)^k\right).$$

Lastly we note that for $c > 0$ the function $f(x) = (c/x)^x$ is increasing for $0 < x \leqslant c/e$ since $d/dx\, f(x) = (c/x)^x (\log(c/x) - 1)$. We have chosen C large enough so that for n large enough we have

$$\mathbb{P}\left([n]_p \not\to_s A\right) \leqslant e^{-\delta np} \left(1 + (sc_0 C^{-1} np) \left(\frac{e2^s\, np}{sc_0 C^{-1} np}\right)^{sc_0 C^{-1} np}\right) \leqslant e^{-\delta np}\, e^{\delta np/2} = o(1).$$

As desired it follows that $[n]_p \to_s A$ asymptotically almost surely for $p \geqslant C\, n^{-1/m_1(A)}$.

5 Proof of Theorem 3 – A Rado-type 0-statement

Let $Q \subseteq [m]$ be a set of column indices satisfying $|Q| \geqslant 2$ such that $(|Q|-1)/(|Q| - r_Q - 1) = m_1(A)$. Due to Lemma 11 we can replace A with $A[Q]$ if necessary in order to guarantee that $(m-1)/(m - \mathrm{rk}(A) - 1) = m_1(A)$. Due to Lemma 8 we know that

$$\mathbb{P}\left([n]_p \to_s^\star A\right) \leqslant \mathbb{P}\left(\bigcup_{p \in \mathcal{P}(A)} \left([n]_p \to_s A_p\right)\right) \leqslant \sum_{p \in \mathcal{P}(A)} \mathbb{P}\left([n]_p \to_s A_p\right). \tag{21}$$

Let us bound the individual probabilities $\mathbb{P}\left([n]_p \to_s A_p\right)$ for each $p \in \mathcal{P}(A)$. For $|p| = m$, that is $p = \{\{1\},\ldots,\{m\}\}$, we know due to Rödl and Ruciński's Theorem 1 that there exists a $c = c(A,s)$ such that $\lim_{n\to\infty} \mathbb{P}\left([n]_p \to_s A\right) = 0$ for $p = p(n) \leqslant c\, n^{-1/m_1(A)}$. For

$|p| < m$ we consider two separate cases. If A_p is not partition regular, then $[n] \not\to_s A$ and therefore trivially $\lim_{n\to\infty} \mathbb{P}\left([n]_p \to_s A_p\right) = 0$. If A_p is partition regular, then

$$m_1(A_p) \geqslant \frac{|p|-1}{|p|-\mathrm{rk}(A)-1} > \frac{m-1}{m-\mathrm{rk}(A)-1} = m_1(A)$$

so that $n^{-1/m_1(A)} = o\big(n^{-1/m_1(A_p)}\big)$ and therefore again by Theorem 1 for $p = p(n) \leqslant c\, n^{-1/m_1(A)}$ we have $\lim_{n\to\infty} \mathbb{P}\left([n]_p \to_s A_p\right) = 0$. The desired statement follows due to Equation (21).

6 Proof of Theorem 4 – A Szémeredi-type 0-statement

Due to Lemma 8 we know that

$$\mathbb{P}\left([n]_p \to^\star_\epsilon A\right) \leqslant \mathbb{P}\left(\bigcup_{p\in\mathcal{P}(A)} \Big([n]_p \to_\epsilon A_p\Big)\right) \leqslant \sum_{p\in\mathcal{P}(A)} \mathbb{P}\left([n]_p \to_\epsilon A_p\right). \tag{22}$$

We will therefore analyze the individual probabilities $\mathbb{P}\left([n]_p \to_\epsilon A_p\right)$ for each $p \in \mathcal{P}(A)$. The constant $c = c(A,\epsilon)$ will be define later in Equation (27). We start by first stating the following proposition, which restricts the statement of Theorem 4 to proper solutions. Before we give its proof, we will show that the 0-statement of Theorem 4 follows easily from it.

Proposition 22. *For every $\epsilon > 0$, $r, m \in \mathbb{N}$ such that $m \geqslant 3$ and matrix $A \in \mathcal{M}_{r\times m}(\mathbb{Z})$ there exists a constant $c = c(A,\epsilon)$ such that the following holds. If A is irredundant, positive and abundant, then we have $\lim_{n\to\infty} \mathbb{P}\left([n]_p \to_\epsilon A\right) = 0$ if $p(n) \leqslant c\, n^{-1/m_1(A)}$.*

For $|p| < m$ we now observe that A_p again clearly is irredundant and positive since p indicates the repeated entries of an actual solution in $\mathcal{S}_1(A)$. If A_p is not abundant, then Proposition 12 states that $\lim_{n\to\infty} \mathbb{P}\left([n]_p \to_\epsilon A_p\right) = 0$ for $p = p(n) \leqslant c\, n^{-1/m_1(A)} = o(1)$ independent of the constant c. If A_p is abundant, then we can apply Proposition 22 to it. If we assume as in the proof of Theorem 3 that $m_1(A) = (m-1)/(m-\mathrm{rk}(A)-1)$, then we again have $n^{-1/m_1(A)} = o\big(n^{-1/m_1(A_p)}\big)$ and therefore $\lim_{n\to\infty} \mathbb{P}\left([n]_p \to_s A_p\right) = 0$ for $p = p(n) \leqslant c\, n^{-1/m_1(A)}$ independent of c. Lastly, let $|p| = m$, that is $p = \{\{1\},\dots,\{m\}\}$ and therefore $A_p = A$. Proposition 22 applies to A and therefore we obtain the desired statement with $c = c(A,\epsilon)$ as given by Proposition 22. The desired statement now follows due to Equation (22).

Proof of Proposition 22. Observe that the expected number of elements in $[n]_p$ is

$$\mathbb{E}(|[n]_p|) = np. \tag{23}$$

We also note that due to Lemma 7 there exists $c_0 = c_0(A[Q])$ such that

$$\mathbb{E}\big(\big|\,\mathcal{S}_0(A[Q]) \cap [n]_p^{|Q|}\big|\big) \geqslant c_0\, n^{|Q|-r_Q}\, p^{|Q|} \quad \text{for} \quad \varnothing \neq Q \subseteq [m] \tag{24}$$

and due to Equation (9) we also have for $\mathbf{b} \in \mathbb{Z}^r$ and $\varnothing \neq Q \subseteq [m]$ that

$$\mathbb{E}\big(|\, \mathcal{S}_0(A[Q], \mathbf{b}) \cap [n]_p^{|Q|}|\big) \leqslant n^{|Q|-r_Q}\, p^{|Q|} \quad \text{for} \quad \varnothing \neq Q \subseteq [m]. \tag{25}$$

Following the alteration method as used for example by Schacht [23], we make three case distinctions. For this, we define the *maximum density* of A to be

$$m(A) = \max_{\varnothing \neq Q \subseteq [m]} \frac{|Q|}{|Q| - r_Q}. \tag{26}$$

Note the difference to the previously defined maximum 1-density. The constant $c = c(A, \epsilon)$ will be stated later in context in Equation (27). Note that we need to cover the whole range of $0 \leqslant p(n) \leqslant c\, n^{-1/m_1(A)}$ since we are not dealing with a monotone property.

Case 1. Assume that $p \ll n^{-1/m(A)}$. Let $\varnothing \neq Q_1 \subseteq [m]$ be a set of column indices such that $|Q_1|/(|Q_1| - r_{Q_1}) = m(A)$. By Equation (25) we now have

$$\lim_{n\to\infty} \mathbb{E}\big(|\, \mathcal{S}_0(A[Q_1]) \cap [n]_p^m|\big) \leqslant \lim_{n\to\infty} n^{|Q_1|-r_{Q_1}}\, p^{|Q_1|} = 0.$$

Markov's Inequality and Lemma 11 therefore give us $\lim_{n\to\infty} \mathbb{P}\left(|\, \mathcal{S}_0(A) \cap [n]_p^m| \neq 0\right) = 0$, see also Rué et al. [20]. It clearly follows that we also have $\lim_{n\to\infty} \mathbb{P}\left([n]_p \to_\epsilon A\right) = 0$ for any $\epsilon > 0$ if $p = p(n) \ll n^{-1/m(A)}$.

Case 2. Assume that $n^{-1} \ll p \ll n^{-1/m_1(A)}$. Let $Q_2 \subseteq [m]$ be a set of column indices satisfying $|Q_2| \geqslant 2$ such that $(|Q_2| - 1)/(|Q_2| - r_{Q_2} - 1) = m_1(A)$ and $|Q_2|$ is as small as possible. Since $np \to \infty$ we have $\lim_{n\to\infty} \mathbb{P}\left(|[n]_p| \geqslant np/2\right) = 1$ due to Chernoff's bound. The expected number of solutions in $[n]_p$ now is asymptotically smaller than the number of elements since by Equation (25) we have

$$\mathbb{E}\big(|\, \mathcal{S}_0(A[Q_2]) \cap [n]_p^m|\big) \leqslant m^m\, n^{|Q_2|-r_{Q_2}}\, p^{|Q_2|} = m^m\, np\, \left(n^{1/m_1(A)}\, p\right)^{|Q_2|-1} = o(np/2).$$

It follows by Markov's Inequality that for any subset of $[n]_p$ of positive density $\epsilon > 0$ we can remove one element per solution contained in this subset so that the resulting set is free of solutions while asymptotically almost surely still having positive density ϵ in $[n]_p$. Lemma 11 therefore gives us that we have $\lim_{n\to\infty} \mathbb{P}\left([n]_p \to_\epsilon^\star A\right) = 0$ for any $\epsilon > 0$ if $n^{-1} \ll p = p(n) \ll n^{-1/m_1(A)}$.

Case 3. Lastly, assume that $n^{-1/m(A)} \ll p \leqslant cn^{-1/m_1(A)}$, where $c = c(A, \epsilon)$ will be given in Equation (27). Due to Chernoff we again have $|[n]_p| \geqslant np/2$ asymptotically almost surely. We now observe that due to Lemma 11 we can replace A with $A[Q_2]$ if necessary in order to guarantee that $(m-1)/(m-\mathrm{rk}(A)-1) = m_1(A)$ as well as $(|Q|-1)/(|Q|-r_Q-1) < m_1(A)$ for any $Q \subsetneq [m]$. We have previously observed that $A[Q_2]$ is again irredundant, positive and abundant. Let $X = (X_n)_{n\in\mathbb{N}}$ denote the sequence of random variables counting the number of proper solutions in $[n]_p$, that is $X_n = |\, \mathcal{S}_0(A) \cap [n]_p^m|$ for $n \in \mathbb{N}$. For

$$c = c(A, \epsilon) = \left(\frac{1-\epsilon}{4}\right)^{1/(m-1)} \tag{27}$$

it follows by Equation (25) that

$$\mathbb{E}(X_n) \leqslant n^{m-\mathrm{rk}(A)}\, p^m \leqslant np\left(n^{1/m_1(A)}\, p\right)^{m-1} \leqslant (1-\epsilon)\, np/4.$$

For a given vector $\mathbf{x} = (x_1, \ldots, x_m)$ we let $s(\mathbf{x}) = \{x_1, \ldots, x_m\}$ denote the set of its entries. Using this we can now estimate the variance of X_n by

$$\begin{aligned}
\mathrm{Var}(X_n) - \mathbb{E}(X_n) &\leqslant \sum_{\substack{\mathbf{x},\mathbf{y}\in\mathcal{S}_0(A)\\ s(\mathbf{x})\cap s(\mathbf{y})\neq\varnothing}} p^{|s(\mathbf{x})|+|s(\mathbf{y})|-|s(\mathbf{x})\cap s(\mathbf{y})|} \\
&= \sum_{\mathbf{x}\in\mathcal{S}_0(A)} p^{|s(\mathbf{x})|}\left(\sum_{\varnothing\neq Q\subsetneq[m]}\left[\sum_{\substack{\mathbf{y}\in\mathcal{S}_0(A)\\ s(\mathbf{x})\cap s(\mathbf{y})=s(\mathbf{y}^{\bar{Q}})}} p^{|s(\mathbf{y})|-|s(\mathbf{x})\cap s(\mathbf{y})|}\right] + \sum_{\substack{\mathbf{y}\in\mathcal{S}_0(A)\\ s(\mathbf{x})\subseteq s(\mathbf{y})}} 1\right) \\
&\leqslant \sum_{\mathbf{x}\in\mathcal{S}_0(A)} p^m\left(\sum_{\varnothing\neq Q\subsetneq[m]} m^{|\bar{Q}|}\max_{\mathbf{b}\in\mathbb{Z}^r}\left(\sum_{\mathbf{y}'\in\mathcal{S}_0(A[Q],\mathbf{b})} p^{|Q|}\right) + m^m\right) \\
&\leqslant m^m \sum_{\mathbf{x}\in\mathcal{S}_0(A)} p^m\left(\sum_{\varnothing\neq Q\subsetneq[m]} n^{|Q|-r_Q}p^{|Q|} + 1\right) \\
&= O\Big(n^{m-\mathrm{rk}(A)}p^m \max_{\varnothing\neq Q\subsetneq[m]}\left(n^{|Q|-r_Q}p^{|Q|}\right)\Big).
\end{aligned}$$

We observe that due to Equation (24) and the assumption on $m_1(A)$ we now have

$$\mathrm{Var}(X_n) = o(\mathbb{E}(X)^2).$$

Chebyshev's inequality therefore gives us $\mathbb{P}\left(|X - \mathbb{E}(X)\,| \geqslant \mathbb{E}(X)\right) = o(1)$ so that

$$\left|\,\mathcal{S}_0(A) \cap [n]_p^m\right| \leqslant 2\,\mathbb{E}(X) = (1-\epsilon)\, np/2$$

asymptotically almost surely. It follows that, given a set of density ϵ, we can remove one element from $[n]_p$ for each solution in $\mathcal{S}_0(A) \cap [n]_p^m$ and asymptotically almost surely still be left with a set of density ϵ, so that $\lim_{n\to\infty}\mathbb{P}\left([n]_p \to_\epsilon A\right) = 0$ for any $\epsilon > 0$ if $n^{-1/m(A)} \ll p = p(n) \leqslant c\, n^{-1/m_1(A)}$ where $c = c(A,\epsilon)$ as given in Equation (27). □

Acknowledgements

I would like to thank Lluís Vena for his tremendous help and input regarding the removal lemma and the proof of the supersaturation results. I would also like to thank the anonymous referee for many helpful comments and suggestions. Furthermore, I would like to thank my advisor Juanjo Rué for his input and supervision.

References

[1] József Balogh, Robert Morris, and Wojciech Samotij. Independent sets in hypergraphs. *Journal of the American Mathematical Society*, 28:669–709, 2015.

[2] David Conlon and William Timothy Gowers. Combinatorial theorems in sparse random sets. *Annals of Mathematics*, (184(2)):367–454, 2016.

[3] Paul Erdős and Miklós Simonovits. A limit theorem in graph theory. In *Studia Sci. Math. Hung*, pages 51–57, 1965.

[4] Paul Erdős, Arthur Harold Stone, et al. On the structure of linear graphs. *Bull. Amer. Math. Soc*, 52(1087-1091):4, 1946.

[5] Paul Erdős and Pál Turán. On a problem of Sidon in additive number theory, and on some related problems. *J. London Math. Soc.*, 1(4):212–215, 1941.

[6] Peter Frankl, Ronald Lewis Graham, and Vojtech Rödl. Quantitative theorems for regular systems of equations. *Journal of Combinatorial Theory, Series A*, 47(2):246–261, 1988.

[7] Ehud Friedgut, Vojtech Rödl, and Mathias Schacht. Ramsey properties of random discrete structures. *Random Structures & Algorithms*, 37(4):407–436, 2010.

[8] Ben Green. A Szemerédi-type regularity lemma in abelian groups, with applications. *Geometric & Functional Analysis GAFA*, 15(2):340–376, 2005.

[9] Robert Hancock, Katherine Staden, and Andrew Treglown. Independent sets in hypergraphs and Ramsey properties of graphs and the integers. Available online at arXiv:1701.04754.

[10] Robert Hancock and Andrew Treglown. On solution-free sets of integers. Available online at arXiv:1607.08399. To appear in *European Journal of Combinatorics*.

[11] Robert Hancock and Andrew Treglown. On solution-free sets of integers II. Available online at arXiv:1611.08498. To appear in *Acta Arithmetica*.

[12] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random Graphs*. Wiley-Interscience series in Discrete Mathematics and Optimization. John Wiley & Sons, Wiley-Interscience, New York, 2000.

[13] Svante Janson and Andrzej Ruciński. Upper tails for counting objects in randomly induced subhypergraphs and rooted random graphs. *Arkiv für Matematik*, 49(1):79–96, 2011.

[14] Daniel Král', Oriol Serra, and Lluís Vena. A removal lemma for systems of linear equations over finite fields. *Israel Journal of Mathematics*, 187(1):193–207, 2012.

[15] Christopher Kusch, Juanjo Rué, Christoph Spiegel, and Tibor Sabó. Random strategies are nearly optimal for generalized van der Waerden games. In preparation.

[16] Rajko Nenadov and Angelika Steger. A short proof of the random Ramsey theorem. *Combinatorics, Probability and Computing*, 25(01):130–144, 2016.

[17] Richard Rado. Studien zur Kombinatorik. *Mathematische Zeitschrift*, 36(1):424–470, 1933.

[18] Vojtech Rödl and Andrzej Ruciński. Rado partition theorem for random subsets of integers. *Proc. London Math. Soc. (3)*, 74(3):481–502, 1997.

[19] Klaus Friedrich Roth. On certain sets of integers. *Journal of the London Mathematical Society*, 1(1):104–109, 1953.

[20] Juanjo Rué, Christoph Spiegel, and Ana Zumalacárregui. Threshold functions and poisson convergence for systems of equations in random sets. *Mathematische Zeitschrift*, May 2017.

[21] Imre Z. Ruzsa. Solving a linear equation in a set of integers I. *Acta Arith.*, 65(3):259–282, 1993.

[22] David Saxton and Andrew Thomason. Hypergraph containers. *Inventiones mathematicae*, 201(3):925–992, 2015.

[23] Matthias Schacht. Extremal results for random discrete structures. *Annals of Mathematics*, (184(2)):333–365, 2016.

[24] Asaf Shapira. A proof of Green's conjecture regarding the removal properties of sets of linear equations. *Journal of the London Mathematical Society*, 81(2):355–373, 2010.

[25] Endre Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975. Collection of articles in memory of Juriĭ Vladimirovič Linnik.

[26] Bartel Leendert van der Waerden. Beweis einer Baudetschen Vermutung. *Nieuw Arch. Wisk*, 15(2):212–216, 1927.

[27] Panayiotis Varnavides. On certain sets of positive density. *Journal of the London Mathematical Society. Second Series*, 1(3):358–360, 1959.