MASTER'S THESIS

Master of Science in Advanced Mathematics and Mathematical Engineering

Title: Solutions of linear equations in random and pseudorandom sets

Author: Miquel Ortega Sánchez-Colomer

Advisor: Oriol Serra Albó

Department: Mathematics

Academic year: 2021-2022

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Abstract

A common theme in modern combinatorics consists in proving sparse analogues of results known in the dense setting. We review some of these for linear systems of equations. We first prove sparse analogues for random sets of Szemerédi's theorem and Rado's theorem via the hypergraph container method. Finally, we prove a sparse analogue for quasirandom sets of Roth's theorem via the regularity method.

Contents

Introduction

A common theme in modern combinatorics consists in proving sparse analogues of results known in the dense setting. This usually means that the ambient space — say, the integers or the complete graph — where a theorem holds is replaced by some sparse subset of it. Of course, this sparse subset must behave in some sense like the whole set in order for the theorem to hold.

The easiest way to do so is replacing the ambient space by a random subset. For example, Szemerédi's theorem [21] says that a subset of $[n] = \{1, \ldots, n\}$ with no k-term arithmetic progressions must have size $o(n)$. More formally, it says the following.

Theorem 0.1 (Szemerédi's theorem)**.** *For every $\delta > 0$ and positive integer k there exists n_0 such that the following holds for $n \geq n_0$. For every subset $X \subseteq [n]$ with $|X| \geq \delta n$, the subset X contains an arithmetic progression of length k.*

Suppose we replace the role of $[n]$ by $[n]_p$, the random binomial subset where every element of $[n]$ is taken with probability $p = p(n)$. Then the theorem should say something like the following.

Theorem 0.2. *For every $\delta > 0$ and positive integer k there exists n_0 such that the following holds for $n \geq n_0$ with high probability. For every subset $X \subseteq [n]_p$ with $|X| \geq \delta pn$, the subset X contains an arithmetic progression of length k.*

This was indeed proved simultaneously by Conlon and Gowers [5] and Schacht [19], under the necessary hypothesis that $p \geq Cn^{-1/(k-1)}$ for large enough C. Note, though, that its expected density p is allowed to satisfy $p = o(1)$, which is what we mean when we say a subset is sparse.

The goal of this thesis is to review some of these sparse analogues for results on linear equations and linear systems of equations over the integers. It is divided in two chapters:

- The first chapter concerns results such as the previous example, where the integers are replaced by a random subset $[n]_p$. In order to prove these results, we use the hypergraph container method developed by Saxton and Thomason [18] and Balogh, Morris and Samotij [1]. The first section of the chapter is devoted to a brief introduction to this technique. The rest of the chapter is devoted to proving sparse analogues of Szemerédi's theorem (in fact, a generalization of it) and of Rado's theorem.

- The second chapter studies instead the behaviour when the sparse subset is a quasirandom subset. These kind of subsets were first studied by

Thomason ([22], [23]) and Chung, Graham and Wilson ([4],[3]). They are deterministic sets that have many random-like properties. The possibility to establish sparse analogues to results over the integers is no exception, and in this case we are able to reach a proof of a sparse analogue to Roth's theorem, which is Szemerédi's theorem for the case of 3-term arithmetic progression. We prove this analogue via the regularity method following Conlon, Fox and Zhao [6].

Chapter 1

Arithmetic results in sparse random sets

The goal of this chapter is proving analogues in sparse random sets of some results on linear systems of equations over the integers. We do this via the hypergraph container method, to which we devote the first section. The second section reviews some necessary concepts from linear algebra and states the standard results over the integers. The last two sections finish out the proof of the desired results. The chapter mainly follows parts of [2], [14], [19] and [20].

1.1 The method of hypergraph containers

Recall that hypergraphs are an extension of graphs where edges are allowed to span across several vertices. More formally, we define them as follows.

Definition 1.1. A hypergraph $\mathcal{H}$ is a pair $\mathcal{H} = (V, E)$ with V an arbitrary set of *vertices* of size $v(\mathcal{H}) = |V|$ and $E \subseteq \mathcal{P}(V) \setminus \varnothing$ a set of *edges* or *hyperedges* with size $e(\mathcal{H}) = |E|$. An r-uniform hypergraph is a hypergraph where all edges have size r.

We can encode many kinds of objects and information, such as arithmetic properties, in hypergraphs. For example, we can define a hypergraph that has $[n] = \{1, \ldots, n\}$ as its vertices and contains the edge $\{x, y, z\}$ if and only if $x + z = 2y$, that is, if x, y and z form an arithmetic progression. In this setting, sets with no arithmetic progressions are sets that contain no edges.

Definition 1.2. An *independent set* of a hypergraph is a set of vertices that contains no edge. We write $\mathcal{I}(\mathcal{H})$ for the family of all indendent sets of $\mathcal{H}$.

So, in the previous example, if one wanted to estimate the number of sets with no arithmetic progression, one could do it from the perspective of counting independent sets in a given hypergraph.

It turns out that controlling independent sets directly is not easy. However, adding some hypothesis on the edge distribution of a given hypergraph, independent sets exhibit certain clustering, and one can control slightly larger sets that contain few edges and include all independent sets. This is what we call the method of hypergraph containers.

This method was developed simultaneously by Saxton and Thomason [18] and Balogh, Morris and Samotij [1] as a technique to study several combinatorial questions related with independent sets in hypergraphs. In rather informal terms, it allows one to find a *small* — in a suitable sense — collection of sets with few edges that contain all independent sets of a hypergraph with evenly distributed edges. This collection of *containers* thus encapsulates the clustering of independent sets that occurs in such hypergraphs, and usually allows one to study properties of finite structures with forbidden substructures (the independent sets). In this section, we see a toy example of how the method is used, following [2], and finally state the full version of the hypergraph container theorem that we use in later sections.

In the examples of this section, we study two properties of triangle-free graphs, that is, graphs that do not contain a copy of K_3. To do so, we use a simpler version of the hypergraph container lemma, which has somewhat simplified conclusions and applies only to 3-uniform hypergraphs. In order to state it, we need some notation.

Definition 1.3. Given a hypergraph $\mathcal{H}$ and a subset of its vertices $X \subseteq V(\mathcal{H})$, the degree of X is the number of edges containing it, and we write it as

$$\deg_{\mathcal{H}}(X) = \left|\{Y \in E(\mathcal{H})\colon X \subseteq Y\}\right|.$$

Definition 1.4. Given a hypergraph $\mathcal{H}$, its maximum l-degree is

$$\Delta_l(\mathcal{H}) = \max_{\substack{X \subseteq V(\mathcal{H}),\\ |X|=l}} \deg_{\mathcal{H}}(X).$$

We can then write the hypothesis on the edge distribution of the hypergraph as a bound on its maximum 1 and 2-degrees.

Theorem 1.5. *For every $c > 0$, there exists $\delta > 0$ such that the following holds. Let $\mathcal{H}$ be a 3-uniform hypergraph with average degree $d \geq \delta^{-1}$ and suppose that*

$$\Delta_1(\mathcal{H}) \leq cd \quad \text{and} \quad \Delta_2(\mathcal{H}) \leq c\sqrt{d}.$$

Then there exists a collection $\mathcal{C}$ of subsets of $V(\mathcal{H})$ with

$$|\mathcal{C}| \leq \binom{v(\mathcal{H})}{v(\mathcal{H})/\sqrt{d}}$$

such that

1. *for every $I \in \mathcal{I}(\mathcal{H})$, there exists $C \in \mathcal{C}$ such that $I \subset C$,*
2. *$|C| \leq (1-\delta)v(\mathcal{H})$ for every $C \in \mathcal{C}$.*

This version of the hypergraph container lemma is taken from [2, Theorem 2.1]. The bounds on the maximum degrees are the necessary conditions on the edge distribution of the hypergraph, and the family $\mathcal{C}$ is the family of containers we referred to. The first conclusion tells us that every independent set belongs to a container, and the second conclusion bounds the size of every container. While this bound may seem small, it can be used to control the number of iterations of the lemma when applying it repeatedly. We do so for the hypergraph of triangles in K_n. By this we mean a hypergraph $\mathcal{H}$ such that:

- Its $\binom{n}{2}$ vertices are the *edges* of K_n, meaning that
$$V(\mathcal{H}) = E(K_n).$$
- Its hyperedges are the triangles of K_n, so that
$$E(\mathcal{H}) = \left\{\{e_1, e_2, e_3\} \in E(K_n)^3 \mid e_i \sim e_{i+1}, e_i \neq e_j \text{ for } i \neq j\right\},$$
where $e_i \sim e_{i+1}$ if the edges are incident and the sum in indices wraps around.

Iterating Theorem 1.5 yields the following result, where one must read the collection $\mathcal{G}$ as the collection of containers, and the conclusion on the small number of triangles corresponds to a small number of hyperedges.

Theorem 1.6. *For every $\varepsilon > 0$ there exists $C > 0$ such that the following holds. For every $n \in \mathbb{N}$, there exists a collection $\mathcal{G}$ of graphs on n vertices, with*
$$|\mathcal{G}| \leq n^{Cn^{3/2}},$$
such that

1. *Each $G \in \mathcal{G}$ contains fewer than εn^3 triangles.*
2. *Each triangle-free graph on n vertices is contained in some $G \in \mathcal{G}$.*

Proof. Let $\mathcal{H}$ be the hypergraph of triangles in K_n we just defined. Note that, for a given edge $e \in K_n$, its degree $\deg_{\mathcal{H}}(e)$ in $\mathcal{H}$ is the number of triangles in K_n that contain an edge e, so that $\Delta_1(\mathcal{H}) = n - 2 \leq n$. Since there is at most one triangle that contains two given edges, $\Delta_2(\mathcal{H}) \leq 1$.

Applying Theorem 1.5 with $c = 1/\varepsilon$ and translating sets of vertices in the hypergraph to edge-induced subgraphs in K_n, we obtain a family of graphs on n vertices $\mathcal{G}_1$ such that

1. Every triangle-free graph is a subgraph of some $G \in \mathcal{G}_1$.
2. Every $G \in \mathcal{G}_1$ satisfies $e(G) \leq (1 - \delta)e(K_n)$.

We now apply the same argument while there is a container with more than εn^3 edges to break it up into smaller containers. Suppose $\mathcal{G}_i$ contains a graph $G \in \mathcal{G}_i$ with $e(G) > \varepsilon n^3$. Let $\mathcal{H}[G] \subset \mathcal{H}$ be the subhypergraph of triangles belonging to G. Note that
$$\Delta_1(\mathcal{H}[G]) \leq \Delta_1(\mathcal{H}) \leq n = c\varepsilon n \leq cd(\mathcal{H}[G]),$$
since the average degree of $\mathcal{H}[G]$ satisfies $d(\mathcal{H}[G]) \geq \varepsilon 3n^3/\binom{n}{2} \geq 6\varepsilon n$. Applying Theorem 1.5 with $c = 1/\varepsilon$ to $\mathcal{H}[G]$ we obtain a family $\mathcal{G}'$ such that any triangle-free graph in G is contained in $G' \in \mathcal{G}'$, as long as $d(\mathcal{H}[G]) \geq 1/\delta$, which is true if we assume that $6\varepsilon n \geq 1/\delta$. We set $\mathcal{G}_{i+1} = (\mathcal{G}_i \setminus G) \cup \mathcal{G}'$ and apply this repeatedly until there is no graph $G \in \mathcal{G}_i$ with $e(G) > \varepsilon n^3$.

Note that every time a graph is broken up, the size of the containers shrinks by a factor $(1 - \delta)$, so that a container may be broken up at most $\log_{(1-\delta)} \varepsilon$

times, since afterwards it will have less than $\varepsilon e(K_n)$ edges and thus less than εn^3 triangles. Furthermore, the hypergraph container lemma guarantees that at every step a container is broken into at most $\binom{n^2}{n^2/\sqrt{6\varepsilon n}}$ smaller containers. Therefore, when the iteration ends, the total number of containers is at most

$$\binom{n^2}{n^{3/2}/\sqrt{6\varepsilon}}^{\log_{1-\delta}\varepsilon} \leq n^{Cn^{3/2}}.$$

Finally, the cases where it does not hold that $6\varepsilon n \geq 1/\delta$, so that we can't apply the hypergraph container lemma, can be covered by taking C large enough and taking all independent sets as containers. □

With this theorem in hand, we may prove two relatively simple applications. The first one is a proof of Mantel's theorem for sparse random graphs. Mantel's theorem states the following.

Theorem 1.7 (Mantel's theorem)**.** *The maximum number of edges in an n-vertex triangle-free graph is $\lfloor n^2/4 \rfloor$.*

Frankl and Rödl proved in [8] an analogue for sparse random graphs, of which we first recall the definition.

Definition 1.8. We write $G(n,p)$ for the binomial random graph with probability p, i.e. the subset of the complete graph where every edge is taken with probability p. In other words, it is a random variable taking values on all subgraphs of K_n satisfying

$$P(G(n,p) = G) = p^{e(G)}(1-p)^{\binom{n}{2}-e(G)}$$

for all $G \subseteq K_n$.

Note too that we say an event X holds asimptotically almost surely if $P(X) \to 1$ when n tends to infinity. Then the result of Frankl and Rödl says the following.

Theorem 1.9. *For every $\alpha > 0$, there exists $C > 0$ such that the following holds. For $p \geq C/\sqrt{n}$ then $G(n,p)$ asimptotically almost surely satisfies that every subgraph $G \subset G(n,p)$ with*

$$e(G) \geq \left(\frac{1}{2} + \alpha\right) p \binom{n}{2}$$

contains a triangle.

One of the main ingredients for applications of the hypergraph container lemma is a supersaturation result. This means a result that lower bounds the number of appearances of a given substructure when the extremal threshold is surpassed. In the case that concerns us, the one of triangles in regular graphs, it can be directly proved applying Mantel's theorem to induced subgraphs of a fixed size.

Lemma 1.10. *For every $\delta > 0$ there exists $\varepsilon > 0$ such that the following holds. If G is a graph on n vertices satisfying*

$$e(G) \geq \left(\frac{1}{4} + \delta\right) n^2,$$

then it contains εn^3 triangles.

Proof. Let h be a constant we fix later on and $V = V(G)$ the vertices of G. In that case, when $n \geq h$ we have that

$$\binom{n-2}{h-2}\left(\frac{1}{4}+\delta\right)n^2 \leq \binom{n-2}{h-2}e(G) = \sum_{\substack{S\subset V,\\ |S|=h}} e(G[S]).$$

Let us split the previous sum between those induced subgraphs on which we may apply Mantel's theorem, that is, those that have more than $h^2/4$ edges, and those where we can't. Writing L for the number of induced subgraphs of size h with more than $h^2/4$ edges, we obtain that

$$\binom{n-2}{h-2}\left(\frac{1}{4}+\delta\right)n^2 \leq \frac{h^2}{4}\binom{n}{h} + \frac{Lh^2}{2}.$$

Let $\mathrm{t}(G)$ count the number of triangles in G. Note that by a double counting argument and Mantel's theorem we have that $L \leq \mathrm{t}(G)\binom{n}{h-3} \leq \mathrm{t}(G)n^{h-3}$, therefore

$$\mathrm{t}(G) \geq \frac{2\binom{n-2}{h-2}n^2}{h^2 n^{h-3}}\left(\delta + \frac{1}{4} - \frac{1}{4}\frac{(n-1)h}{n(h-1)}\right).$$

Letting h be a large enough constant, the conclusion of the statement is true for ε small enough and $n \geq n_0$.

The case of small n can be handled by taking ε small enough and applying the regular Mantel's theorem. □

We may now prove Theorem 1.9 up to a logarithmic factor in the hypothesis, that is, assuming that $p \geq C \log n/\sqrt{n}$.

Proof. Let ε be the constant given by the supersaturation lemma (Lemma 1.10) with $\delta = \alpha/2$. Apply Theorem 1.6 with the obtained ε. Using the supersaturation result we know that every container $G \in \mathcal{G}$ given by the hypergraph container lemma satisfies

$$e(G) < \left(\frac{1}{4}+\frac{\alpha}{2}\right)n^2.$$

Suppose that the conclusion we want to prove does not hold for a particular instance of $G(n,p)$, that is, suppose there exists $H \subset G(n,p)$ with $e(H) > (1/2+\alpha)p\binom{n}{2}$ and no triangle. Then, since it has no triangle, it holds that $H \subset G$ for some $G \in \mathcal{G}$. In particular, we have that

$$e(G(n,p)\cap G) \geq \left(\frac{1}{2}+\alpha\right)p\binom{n}{2}$$

for some $G \in \mathcal{G}$. Note that $e(G(n,p)\cap G) \sim \mathrm{Bin}(e(G),p)$. Applying the Chernoff bound to bound the tail of the binomial distribution, for any given $G \in \mathcal{G}$, we have that

$$\begin{aligned} P\left(e(G(n,p)\cap G) \geq \left(\frac{1}{2}+\alpha\right)p\binom{n}{2}\right) &\leq P\left(e(G(n,p)\cap G) \geq pe(G) + cn^2\right) \\ &\leq e^{-cn^2p} \end{aligned}$$

where $c = c(\alpha)$ is a small enough constant depending on α which changes in the second inequality.

Finally, we apply a union bound to bound the total probability of the existence of such H, which gives

$$P\left(\mathrm{ex}(G(n,p), K_3) \geq \left(\frac{1}{2}+\alpha\right) p\binom{n}{2}\right) \leq n^{Cn^{3/2}} e^{-cn^2 p}.$$

Finally, if $p \geq 2Cc^{-1} \log n/\sqrt{n}$, the previous bound tends to zero and we obtain the desired result. $\square$

Note that the log factor may be removed by a more careful usage of the hypergraph container method, using a more precise version of the hypergraph container lemma. We state this version at the end of the section and do this kind of analysis in later sections, but for the purposes of this section the straightforward application suffices.

Our second example also concerns triangle-free graphs, but in this case we prove a Ramsey result. This application of the hypergraph container lemma is due to Nenadov and Steger [12]. In later sections we prove also prove Ramsey results, and the corresponding proofs have a similar structure. In this case, we wish to study the threshold for the appearance of monochromatic triangles in random graphs. More formally, let us write

$$K_3 \to (G)_r$$

if any r-colouring of the edges of G (every partition of its edges into r disjoint sets) contains a monochromatic K_3 (three edges which form a K_3 belong to the same set in the partition). Rödl and Ruciński [15] proved the following.

Theorem 1.11. *For every $r \geq 2$ there exists C such that the following holds. If $p \geq C/\sqrt{n}$, then $K_3 \to (G(n,p))_r$ asimptotically almost surely.*

They went on to study this threshold for more general graphs, finally solving it in [16] for all graphs. The proof, however, is rather involved, and Nenadov and Steger provided a simpler proof, provided one is willing to use the hypergraph container lemma.

Let us first recall Ramsey's theorem for the complete graph.

Theorem 1.12. *For any graph G and number of colors r, there exists an n_0 such that the following holds. For any r-coloring of the edges of K_n with $n \geq n_0$, there exists at least one monochromatic copy of G.*

Again, one of the key ingredients to apply the hypergraph container method is a supersaturation version of this result, which follows from a standard averaging argument.

Lemma 1.13. *For any graph G, there exist n_0 and $c > 0$ such that the following holds. For any coloring of the edges of K_n with $n \geq n_0$ there exist at least $cn^{v(G)}$ monochromatic copies of G.*

Proof. Let h be the constant given by Ramsey's theorem with G and r, so that any r-coloring of the edges of a complete graph with more than h vertices contains a monochromatic copy of G. For a given $n \geq h$ and an r-coloring of the edges of K_n, consider all cliques of size h. By a double counting argument, writing ξ for the number of monochromatic copies of G, we have that

$$\binom{n}{h} \leq \xi \binom{n - v(G)}{h - vG},$$

from which the lemma follows. $\square$

From this supersaturation result, we may deduce that a collection of subgraphs where every subgraph has few monochromatic triangles cannot cover the complete graph, in the sense that it misses a significant amount of edges.

Lemma 1.14. *For every integer $r > 0$, there exist $\varepsilon > 0$, $\delta > 0$ and $n_0 > 0$ such that following holds for every $n \geq n_0$. Given $G_1, \dots, G_r$ subgraphs of K_n such that every G_i contains less than εn^3 triangles, then $e(K_n \setminus (G_1 \cup \dots \cup G_r)) \geq \delta n^2$.*

Proof. We may assume that the subgraphs are disjoint, since the hypothesis hold and the conclusion remains the same. We can then consider the subgraphs G_i as a coloring with an $r+1$-th color defined by $G_{r+1} = K_n \setminus (G_1 \cup \dots G_r)$. Applying the supersaturation version of Ramsey's theorem (Lemma 1.10), we obtain the existence of cn^3 monochromatic triangles. Then

$$cn^3 \leq \mathrm{t}(G_{r+1}) + r\varepsilon n^3 \leq e(G_{r+1})n + r\varepsilon n^3,$$

so taking $\delta = \varepsilon = c/(2r)$ guarantees the conclusion. $\square$

As we did before, we prove Theorem 1.11 under the additional hypothesis that $p \geq C \log n / \sqrt{n}$ for the sake of simplicity.

Proof. Let $\varepsilon > 0$ and $\delta > 0$ be as in the previous lemma (Lemma 1.14). Apply Theorem 1.6 with this ε so that we obtain a family of containers $\mathcal{G}$ with $|\mathcal{G}| < n^{C_1 n^{3/2}}$. Suppose $G(n,p)$ admits an r-coloring with no monochromatic triangles. In that case, every color must be contained in a container G_i, so that $G(n,p) \subseteq G_1 \cup \dots \cup G_r$. Note that

$$P\left(G(n,p) \subseteq \bigcup G_i\right) = P\left(G(n,p) \cap \left(K_n \setminus \bigcup G_i\right)\right) = \varnothing) = (1-p)^{e(K_n \setminus \bigcup G_i)}.$$

Since every G_i has at most εn^3 triangles, applying Lemma 1.14 we obtain that

$$P\left(G(n,p) \subseteq \bigcup G_i\right) \leq (1-p)^{\delta n^2} \leq e^{-\delta p n^2}.$$

Applying a union bound over all possible choices of r containers, we have that

$$P(G(n,p) \not\to (K_3)_r) \leq \left(n^{C_1 n^{3/2}}\right)^r e^{-\delta p n^2},$$

which tends to zero if we take C large enough in the hypothesis. $\square$

With this we end our two example applications of the hypergraph container method to illustrate how it is used. In later sections we need a stronger version of the theorem we have used until now. The full statement we use is taken from [1, Theorem 2.2], and first needs the following definition.

Definition 1.15. Let $\mathcal{H}$ be a uniform hypergraph, let $\mathcal{F}$ be an increasing family of subsets of $V(\mathcal{H})$ and let $\varepsilon \in (0,1]$. We say that $\mathcal{H}$ is $(\mathcal{F}, \varepsilon)$-*dense* if

$$e(\mathcal{H}[A]) \geq \varepsilon e(\mathcal{H})$$

for every $A \in \mathcal{F}$.

Then the theorem reads as follows.

Theorem 1.16 (Hypergraph Container Theorem)**.** *For every $k \in \mathbb{N}$ and all positive c and ε, there exists a positive constant C such that the following holds. Let $\mathcal{H}$ be a k-uniform hypergraph and let $\mathcal{F} \subseteq \mathcal{P}(V(\mathcal{H}))$ be an increasing family of sets such that $|F| \geq \varepsilon v(\mathcal{H})$ for all $F \in \mathcal{F}$. Suppose that $\mathcal{H}$ is $(\mathcal{F}, \varepsilon)$*-dense *and $p \in (0,1)$ is such that, for every $l \in [k]$,*

$$\Delta_l(\mathcal{H}) \leq cp^{l-1} \frac{e(\mathcal{H})}{v(\mathcal{H})} \tag{1.1}$$

for some absolute constant c. Then there exists a family $\mathcal{S} \subseteq \binom{V(\mathcal{H})}{\leq Cpv(\mathcal{H})}$ and functions $f \colon \mathcal{S} \to \mathcal{P}(V(\mathcal{H})) \setminus \mathcal{F}$ and $g \colon \mathcal{I}(\mathcal{H}) \to \mathcal{S}$ such that for every $I \in \mathcal{I}(\mathcal{H})$,

$$g(I) \subseteq I \qquad \text{and} \qquad I \setminus g(I) \subseteq f(g(I)).$$

The main difference of this version is that it applies to r-uniform graphs instead of 3-uniform ones, so that it allows us to work with larger forbidden substructures. Furthermore, instead of just including the independent set into a container, we associate each independent set to a container via a small *fingerprint* $g(I)$ (a subset of the independent set). This allows for somewhat finer control, which, for example, permits removing the log factor from the previous proofs. Moreover, the independent sets are mostly contained in sets outside of the family $\mathcal{F}$.

1.2 Partition and density regular matrices

In order to state and prove the results we are concerned with in this chapter, it is useful to introduce some notation and results from linear algebra. We devote this section to do so. For the most part of the section, we consider *homogeneous* linear systems of the form $A\boldsymbol{x} = 0$ with $A \in \mathcal{M}_{s \times m}(\mathbb{Z})$ an integer coefficient matrix with s rows and m columns. In some cases we may also include *non-homogeneous* systems of the form $A\boldsymbol{x} = \boldsymbol{b}$. We usually write vectors in boldface so that $\boldsymbol{v} = (v_1, \ldots, v_m)$ and matrices in capital letters.

To begin with, let us introduce some notation.

Notation 1.17. Given a matrix A with integer coefficients, we write

$$S(A, \boldsymbol{b}) = \{\boldsymbol{x} \in \mathbb{Z}^m \mid A\boldsymbol{x} = \boldsymbol{b}\}$$

for the set of solutions to the system $A\boldsymbol{x} = \boldsymbol{b}$, and set $S(A) = S(A, \boldsymbol{0})$.

Notation 1.18. Given a matrix A with integer coefficients, we write

$$S_0(A) = \{\boldsymbol{x} = (x_1, \ldots, x_m) \colon A\boldsymbol{x} = 0 \text{ and } x_i \neq x_j \text{ for all } i \neq j\}$$

for the set of *proper* solutions to the system $A\boldsymbol{x} = 0$.

Let us also prove an elementary bound on the number of integer solutions to a system, which we need later on.

Lemma 1.19. *For every matrix $A \in \mathcal{M}_{s \times m}(\mathbb{Z})$ and integer $n > 0$, we have that*

$$|S(A, \boldsymbol{b}) \cap [n]^m| \leq n^{m - \mathrm{rk}(A)}. \tag{1.2}$$

Proof. The difference between two elements in $S(A, \boldsymbol{b})$ is an element in $S(A)$, so we may reduce ourselves to the case $\boldsymbol{b} = \boldsymbol{0}$. One possible way to prove the result is to write A in Smith normal form. This tells us that there are invertible matrices $S \in \mathcal{M}_{s\times s}(\mathbb{Z})$ and $T \in \mathcal{M}_{m\times m}(\mathbb{Z})$ such that SAT is diagonal. In that case, the elements of $S(A)$ are in bijection with the solutions to

$$SAT\boldsymbol{x} = 0, \tag{1.3}$$

since we can multiply by S^{-1} at both sides and make the change of variables $\boldsymbol{y} = T\boldsymbol{x}$.

Finally, considering the equation (1.3) modulus n, we see that $A\boldsymbol{x} \equiv 0 \mod n$ has less than $n^{m-\mathrm{rk}(SAT)} = n^{m-\mathrm{rk}(A)}$ solutions in $\mathbb{Z}/(n\mathbb{Z})^m$, so there are at most $n^{m-\mathrm{rk}(A)}$ solutions in $[n]^m$. $\square$

The first class of matrices we introduce are *partition regular* matrices. A matrix is partition regular if, for any coloring of the integer numbers, it admits a monochromatic solution. More formally, we have the following.

Definition 1.20. A matrix A with integer coefficients is partition regular if, for every partition of the integers $C_1 \sqcup \cdots \sqcup C_r = \mathbb{Z}$, there exists a set C_i such that $C_i^m \cap S(A) \neq \varnothing$.

Rado's theorem [13] characterizes such matrices. To state it, we define the following condition.

Definition 1.21. A matrix A with integer coefficients and columns $\boldsymbol{c_1}, \ldots, \boldsymbol{c_n}$ satisfies the *columns condition* if there exists an ordered partition $C_1, \ldots, C_k$ of the column indices with sums $\boldsymbol{s_i} = \sum_{j\in C_i} \boldsymbol{c_j}$ such that

- The first sum is zero, i.e., $\boldsymbol{s_1} = \boldsymbol{0}$.
- Every sum can be written as a rational linear combination of the previous ones. That is, for $i \geq 2$, there exist $a_j \in \mathbb{Q}$ such that $\boldsymbol{s_i} = a_1\boldsymbol{s_1} + \cdots + a_{i-1}\boldsymbol{s_{i-1}}$.

Rado's theorem can then be stated as follows.

Theorem 1.22 (Rado)**.** *A matrix with integer coefficients is partition regular if and only if it satisfies the columns condition.*

In order to study sparse random analogues of Rado's theorem or Szeméredi's Theorem, Rödl and Ruciński [14] introduced a matrix parameter, which we write as m_A for a given matrix A, that determines the probability threshold on which properties we are concerned with — such as when random Rado's theorem holds — are satisfied. We introduce it now since its definition only requires linear algebra.

Notation 1.23. Given a matrix $A \in \mathcal{M}_{s\times m}(\mathbb{Z})$ and an index subset $Q \subseteq [m]$, we denote by A^Q the submatrix of A obtained by restricting to the columns indexed by the elements of Q.

Definition 1.24. For a given partition regular matrix $A \in \mathcal{M}_{s\times m}(\mathbb{Z})$ with full rank, let

$$m_A = \max_{\substack{Q\subseteq[m],\\ |Q|\geq 2}} \frac{|Q| - 1}{|Q| - 1 + \mathrm{rk}\left(A^{\overline{Q}}\right) - s}$$

The following lemma, taken from [14, Proposition 2.2] guarantees that the previous quantity is well-defined, and describes the case $|Q| = 1$, which is useful later on.

Lemma 1.25. *For every partition regular matrix $A \in \mathcal{M}_{s\times m}(\mathbb{Z})$ with full rank it holds that*

$$|Q| - 1 + \operatorname{rk}\left(A^{\overline{Q}}\right) - s \geq 1 \qquad \text{for } |Q| \geq 2, \tag{1.4}$$

while

$$\operatorname{rk}\left(A^{\overline{Q}}\right) = s \qquad \text{for } |Q| = 1. \tag{1.5}$$

Some issues with solutions to a linear system arise, particularly when codifying solutions in a hypergraph, if the solutions have repeated entries. In order to avoid this, we introduce the following class of matrices.

Definition 1.26. A matrix A is *irredundant* if $S_0(A) \neq \varnothing$.

If a given partition regular matrix A is not irredundant, one can find a matrix A' with the same family of solutions viewed as sets. See, e.g., [14, Section 1]. Hence, we will do all our analysis concerning irredundant matrices, but will not lose much generality.

Next, we introduce a subclass of partition regular and irredundant matrices. Recall Szemerédi's (Theorem 0.1), which tells us that positive upper density sets of the integers must contain arithmetic progressions of arbitrary size. Note that an arithmetic progression can be encoded as solutions to linear system. We define density regular matrices as precisely those where the conclusions of Szemerédi's theorem hold.

Definition 1.27. A matrix A is *density regular* if for every $\delta > 0$ there exists n_0 such that the following holds. Every subset $X \subseteq [n]$ with size $|X| \geq \delta n$ contains solutions to $A\boldsymbol{x} = 0$, that is, $S_0(A) \cap X^m \neq \varnothing$.

In [7] Frankl, Graham and Rödl proved that density regular matrices are precisely irredundant partition regular matrices A such that $(1 \dots 1) \in S(A)$. With this, we end the linear algebra prerequisites. However, we will still see a result that best fits into this section.

In the following sections, we use the hypergraph container method to prove our results. In particular, we need to encode the solutions to $A\boldsymbol{x} = 0$ in a hypergraph to which we may apply the hypergraph container theorem. We do so in the straightforward fashion, taking all solutions as edges. Concretely, we study the hypergraph $\mathcal{H}_n$ with vertices $V(\mathcal{H}_n) = [n]$ and edge multiset

$$E(\mathcal{H}_n) = \{\{x_1, \dots, x_m\} : (x_1, \dots, x_m) \in S_0(A) \cap [n]^m\},$$

so that $e(\mathcal{H}_n) = |S_0(A) \cap [n]^m|$.

In order to apply the hypergraph container theorem, we need to ensure this hypergraph satisfies the bound (1.6) on the maximum l-degree. That is the purpose of the following lemma.

Lemma 1.28. *Let $\mathcal{H}_n$ be the hypergraph of solutions to $A\boldsymbol{x} = 0$ for a given partition regular, full rank, irredundant matrix $A \in \mathcal{M}_{s\times m}(\mathbb{Z})$. Then we have that*

$$\Delta_l(\mathcal{H}_n) \leq m! \binom{m}{l} \max_{Q \subseteq [m], |Q| = l} n^{m - l - \operatorname{rk}\left(A^{\overline{Q}}\right)}. \tag{1.6}$$

Proof. Given $x_1, \ldots, x_l \in [n]$ distinct values, the number of edges containing this set is bounded by

$$\left|\left\{(\boldsymbol{y}, \pi) \mid \boldsymbol{y} \in [n]^m \text{ with } A\boldsymbol{y} = 0 \text{ and } \pi \colon [l] \to [m] \text{ such that } y_{\pi(i)} = x_i\right\}\right|.$$

There are at most $m!\binom{m}{l}$ possible π. It follows that

$$\Delta_l(\mathcal{H}_n) \leq m!\binom{m}{l} \max_{\substack{x_1,\ldots,x_l \\ Q \subseteq [m], \\ |Q|=l}} \left|\left\{\boldsymbol{y} \in [n]^{m-l} \mid A^Q \boldsymbol{x} + A^{\overline{Q}} \boldsymbol{y} = 0\right\}\right|.$$

Finally, using (1.2) to bound the number of solutions $\boldsymbol{y}$ to $A^{\overline{Q}}\boldsymbol{y} = -A\boldsymbol{x}$, we have that

$$\Delta_l(\mathcal{H}_n) \leq m!\binom{m}{l} \max_{\substack{Q \subseteq [m], \\ |Q|=l}} n^{m-l-\mathrm{rk}\left(A^{\overline{Q}}\right)},$$

as desired. ◻

With this bound in hand, we prove a neater version of the container theorem applied to our particular context.

Corollary 1.29. *For a given partition regular, full rank, irredundant matrix $A \in \mathcal{M}_{s\times m}(\mathbb{Z})$ and all positive ε, there exists a constant C such that the following holds. There is a family of sets $F_1, \ldots, F_t \subseteq \binom{[n]}{\leq Cn^{1-1/m_A}}$ and $C_1, \ldots, C_t \subseteq [n]$ such that, for any set S satisfying $S^m \cap S_0(A) = \varnothing$, there exists j with*

$$F_j \subseteq S \qquad S \setminus F_j \subseteq C_j.$$

Moreover, the sets C_i satisfy $|C_i^m \cap S_0(A)| \leq \varepsilon |S_0(A) \cap [n]^m|$.

Proof. We wish to apply Theorem 1.16 with $p = n^{-1/m_A}$. We do so directly with the family of sets with *many* solutions, that is, we apply it to

$$\mathcal{F} = \{S \in \mathcal{P}([n]) \colon |S_0(A) \cap S^m| \geq \varepsilon |S_0(A) \cap [n]^m|\}.$$

Note that this is an increasing family and $(\mathcal{F}, \varepsilon)$-dense by definition. As for the bound on the maximum l-degree, using (1.6), when $l \geq 2$ we have that

$$\begin{aligned}
\Delta_l(\mathcal{H}_n) \leq& c \max_{\substack{Q \subseteq [m], \\ |Q|=l}} n^{m-l-\mathrm{rk}\left(A^{\overline{Q}}\right)} \\
\leq& cn^{m-s-1} \max_{\substack{Q \subseteq [m], \\ |Q|=l}} n^{-|Q|+1-\mathrm{rk}\left(A^{\overline{Q}}\right)+s} \\
\leq& cn^{m-s-1}(n^{-1/m_A})^{l-1} \\
\leq& cp^{l-1}\frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}.
\end{aligned}$$

where the last inequality follows from A having full rank so that $e(\mathcal{H}_n) \geq c_0 n^{m-s}$ and c is an appropriately large constant that may change through the inequalities. When $|Q| = l = 1$, it holds that $\mathrm{rk}(A^{\overline{Q}}) = \mathrm{rk}(A)$ by (1.5), and applying again the bound (1.6) we have that

$$\Delta_1(\mathcal{H}_n) \leq cn^{m-1-s} \leq c\frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)}.$$ ◻

1.3 Szemerédi's Theorem

In this section we prove an analogue of Szemerédi's Theorem in the random sparse setting. In fact, we prove somewhat more, studying not only arithmetic progressions but solutions to any linear system $A\boldsymbol{x} = 0$ with A a density regular matrix.

This result was originally proved both by Schacht [19] and by Conlon and Gowers [5]. However, we carry out a proof using the hypergraph container method, as Balogh, Morris and Samotij [1] did — for the case of arithmetic progressions — in their original paper on the method. See also [17], where Rué, Serra and Vena use the hypergraph container method to solve the problem in more generality, counting configurations in groups. Before stating the exact theorem we prove, let us introduce some notation.

Notation 1.30. Given a matrix A with integer coefficients, a set of integers S and a positive integer $r \in \mathbb{N}$, we say that

$$S \to (A)_\varepsilon$$

if every subset $X \subset S$ with $|X| \geq \varepsilon|S|$ admits a proper solution $A\boldsymbol{x} = 0$, that is, it holds that $S_0(A) \cap X^m \neq \varnothing$.

Definition 1.31. We denote by $[n]_p$ the binomial random subset of $[n] = \{1, \dots, n\}$ where each element of $[n]$ is included independently with probability $p = p(n)$. In other words, it is a random variable taking values on all subsets of $[n]$ satisfying

$$P([n]_p = X) = p^{|X|}(1-p)^{n-|X|}$$

for all $X \subseteq [n]$.

Our goal then is to prove the following theorem.

Theorem 1.32. *For any density regular matrix A with integer coefficients and full rank, and $\varepsilon \in (0, 1/2)$, there exist constants $c = c(A, r)$ and $C = C(A, r)$ such that the following statement holds.*

- *If $p(n) \leq cn^{-1/m_A}$, then*

$$\lim_{n\to\infty} P([n]_p \to (A)_\varepsilon) = 0.$$

- *If $p(n) \geq Cn^{-1/m_A}$, then*

$$\lim_{n\to\infty} P([n]_p \to (A)_\varepsilon) = 1.$$

Note that the condition $\varepsilon < 1/2$ is due to the 0-statement. Our proof of the 1-statement does not require this property.

The first ingredient we need is a supersaturation result, as in our example applications of the hypergraph container method. In this case, Frankl, Graham and Rödl [7, Theorem 2] proved the following.

Theorem 1.33. *For any full rank, density regular matrix $A \in \mathcal{M}_{s\times m}(\mathbb{Z})$ and $\delta > 0$ there exist constants $c = c(A, \delta)$ and $n_0 = n_0(A, \delta)$ such that the following holds. For any $n \geq n_0$ and $X \subseteq [n]$ with $|X| > \delta n$, the set X must contain at least cn^{m-s} proper solutions, that is, we have that*

$$|X^m \cap S_0(A)| \geq cn^{m-s}.$$

We can then prove the 1-statement of the sparse analogue to this theorem.

Proof of the 1*-statement of Theorem 1.32.* Let $\delta = \delta(\varepsilon) > 0$ be a constant to be fixed later on. Let $F_1, \ldots, F_t$ and $C_1, \ldots, C_t$ be the sets and M the constant obtained in Corollary 1.29 setting $\varepsilon_{1.29} = c_{1.33}(A, \delta)$ to be the constant obtained in Theorem 1.33 with the current δ. Then, since the sets C_i have less than $c_{1.33} n^{m-s}$ proper solutions, we can conclude that they have size $|C_i| \leq \delta n$ on account of supersaturation.

Suppose that $[n]_p \not\to (A)_\varepsilon$. Then there exists a subset $X \subset [n]_p$ with $|X| = \varepsilon |[n]_p|$ and $|X^m \cap S_0(A)| = \varnothing$. Therefore, on account of Corollary 1.29, there exists an i such that $F_i \subseteq X$ and $X \subseteq C_i \cup F_i$. Suppose too that $|[n]_p| \geq pn/2$, so that

$$|[n]_p \cap C_i| \geq \frac{\varepsilon pn}{2} - |F_i|.$$

Note that we can take C large enough to ensure that

$$|F_i| \leq M n^{1-1/m_A} \leq \varepsilon pn/4. \tag{1.7}$$

Therefore, taking a union bound over all possible subsets and setting $m = \varepsilon pn/2$, then using that $\binom{n}{k} \leq (en/k)^k$, we have that

$$\begin{aligned}
P([n]_p \not\to (A)_\varepsilon \wedge |[n]_p| \geq pn/2) &\leq \sum_{1 \leq i \leq t} P(F_i \subseteq [n]_p) \sum_{\substack{Y \subseteq C_i, \\ |Y| = m - |F_i|}} P(Y \subseteq [n]_p) \\
&\leq \sum_{1 \leq i \leq t} p^{|F_i|} \binom{\delta n}{m - |F_i|} p^{m - |F_i|} \\
&\leq \sum_{1 \leq i \leq t} p^{|F_i|} \left(\frac{2e\delta pn}{m} \right)^{m/2}
\end{aligned}$$

Grouping terms by the size of the fingerprint $|F_i|$ and substituting the value of m, we have that

$$\begin{aligned}
P([n]_p \not\to (A)_\varepsilon \wedge |[n]_p| \geq pn/2) &\leq \sum_{k=1}^{Mn^{1-1/m_A}} \binom{n}{k} p^k \left(\frac{4e\delta}{\varepsilon} \right)^{\varepsilon pn/4} \\
&\leq \sum_{k=1}^{Mn^{1-1/m_A}} \left(\frac{epn}{k} \right)^k \left(\frac{4e\delta}{\varepsilon} \right)^{\varepsilon pn/4}.
\end{aligned}$$

Note that $\left(\left(\frac{a}{x} \right)^x \right)' = \left(\frac{a}{x} \right)^x (\log(a/x) - 1)$, so that, for C large enough, we can ensure that the summands are increasing and bound the sum by the largest term. Therefore, using (1.7) we have that

$$P([n]_p \not\to (A)_\varepsilon \wedge |[n]_p| \geq pn/2) \leq \varepsilon pn \left(\frac{eC}{M} \right)^{\varepsilon pn/4} \left(\frac{4e\delta}{\varepsilon} \right)^{\varepsilon pn/4},$$

which tends to zero if we take δ small enough. The case that $|[n]_p| < pn/2$ can also be exponentially bounded by Chernoff's inequality. $\square$

1.4 Rado's Theorem

Our goal in this section is proving an analogue of Rado's theorem for sparse random subsets of the integers. This result was partially proved by Rödl and Ruciński [14] and completely solved by Friedgut, Rödl and Schacht [9]. However, we follow Spiegel's [20] proof, which is quite less involved if one is willing to use the method of hypergraph containers, and is based on the solution of Nenadov and Steger [12] of which we gave a sample in Section 1.1.

Before stating the corresponding result, we introduce the following notation.

Notation 1.34. Given a matrix A with integer coefficients, a set of integers S and a positive integer $r \in \mathbb{N}$, we say that

$$S \to (A)_r$$

if every partition $S = T_1 \sqcup \cdots \sqcup T_r$ of size r admits a proper monochromatic solution to $A\boldsymbol{x} = 0$, that is, there exists T_i such that $T_i \cap S_0(A) \neq \varnothing$.

Theorem 1.35. *For any irredundant, partition regular matrix A with integer coefficients and full rank, and a given number of colors r, there exist constants $c = c(A, r)$ and $C = C(A, r)$ such that the following statement holds.*

- *If $p(n) \leq cn^{-1/m_A}$, then*

$$\lim_{n\to\infty} P([n]_p \to (A)_r) = 0.$$

- *If $p(n) \geq Cn^{-1/m_A}$, then*

$$\lim_{n\to\infty} P([n]_p \to (A)_r) = 1.$$

The corresponding items of Theorem 1.35 are respectively called the 0-statement and the 1-statement of the theorem. In this section we limit ourselves to proving the second one. See [14] for a proof of the 0-statement.

Frankl, Graham and Rödl [7] proved a supersaturation version of Rado's theorem that reads as follows.

Theorem 1.36. *Given a matrix $A \in \mathcal{M}_{s\times m}(\mathbb{Z})$ of rank s satisfying the columns condition and a number of colors r, there exists a constant $c = c(A, r)$ such that the following holds. For any r-coloring of $[n]$, there exist at least cN^{m-s} monochromatic solutions.*

The following corollary is a consequence of the supersaturation version of Rado's theorem. It tells us that if a collection of subsets contains *few* monochromatic solutions, then it cannot cover the whole interval (i.e. its complement has linear size).

Corollary 1.37. *Given a partition regular, full rank, irredundant matrix $A \in \mathcal{M}_{s\times m}(\mathbb{Z})$, there exist $\varepsilon = \varepsilon(A, r)$ and $\delta = \delta(A, r)$ such that the following statement holds. Given a collection of subsets $T_1, \ldots, T_r \subset [n]$ satisfying $|S_0(A) \cap T_i^m| \leq \varepsilon |S_0(A) \cap [n]^m|$, then $|[n] \setminus \bigcup T_i| \geq \delta n$.*

Proof. We may assume that the subsets T_i are disjoint, since removing duplicate elements respects both the hypothesis and the conclusion. Let $T_{r+1} = [n] \setminus \bigcup T_i$, so that $T_1, \dots, T_{r+1}$ form a coloring of $[n]$. Applying the supersaturation version of Rado (Theorem 1.36), we obtain c such that there are at least cn^{m-s} monochromatic proper solutions. Therefore,

$$cn^{m-s} \leq \sum_{i=1}^{r+1} |S_0(A) \cap T_i^m| \leq r\varepsilon n^{m-s} + |S_0(A) \cap T_{r+1}^m|$$

Taking $\varepsilon = c/(2r)$ and reordering terms, we obtain that

$$|S_0(A) \cap T_{r+1}^m| \geq \frac{c}{2} n^{m-s}.$$

Furthermore, applying the bound (1.6) and the property (1.5) we obtain that $|S_0(A) \cap T_{r+1}^m| \leq C|T_{r+1}|n^{m-s-1}$, from which we deduce the desired conclusion. □

With this result in hand, we are ready to prove the 1-statement of the sparse analogue to Rado's theorem. Note the similarities between this proof and the one we did for Theorem 1.11, although in this case a finer analysis allows us to remove the log factor.

Proof of the 1*-statement of Theorem 1.35.* We only prove the 1-statement. Let $F_1, \dots, F_t$ and $C_1, \dots, C_t$ be the sets and M the constant obtained in Corollary 1.29 with ε small enough so that Corollary 1.37 holds. Consider a partition of the random set $[n]_p = T_1 \sqcup \dots \sqcup T_r$. Then for each i there exists j_i such that $F_{j_i} \subseteq T_i$ and $T_i \setminus F_{j_i} \subseteq C_{j_i}$. In particular, $[n]_p \subseteq \bigcup F_{j_i} \cup C_{j_i}$. Therefore, writing $B_{\boldsymbol{j}} = \bigcup F_{j_i} \cup C_{j_i}$, with $\boldsymbol{j} = (j_1, \dots, j_r)$, we have that

$$P([n]_p \not\to (A)_r) \leq \sum_{\boldsymbol{j} \in [t]^r} P\left(\bigcup F_{j_i} \subseteq [n]_p \text{ and } [n]_p \subseteq B_{\boldsymbol{j}}\right).$$

Note that $[n]_p \subseteq B_{\boldsymbol{j}}$ if and only if $[n] \setminus B_{\boldsymbol{j}} \cap [n]_p = \varnothing$, so that both events are independent, and then

$$P([n]_p \not\to (A)_r) \leq \sum_{\boldsymbol{j} \in [t]^r} p^{|\bigcup F_{j_i}|}(1-p)^{n-|B_{\boldsymbol{j}}|} \leq (1-p)^{\delta n} \sum_{\boldsymbol{j} \in [t]^r} p^{|\bigcup F_{j_i}|}$$

for n large enough, since in that case $n - |B_{\boldsymbol{j}}| \geq n - |\bigcup C_{j_i}| - |\bigcup F_{j_i}|$, and $n - |\bigcup C_{j_i}| \geq \delta' n$ by Corollary 1.37 and $|\bigcup F_{j_i}| \leq rMn^{1-1/m_A}$ by Corollary 1.29, so that $n - |B_{\boldsymbol{j}}| \geq \delta n$ for small enough δ.

We can bound the last sum by grouping terms according to the size of $|\bigcup F_{j_i}|$, which is at most rMn^{1-1/m_A}. The number of possible terms of size k can be bounded by choosing which k elements belong in $|\bigcup F_{j_i}|$ and to which fingerprint they belong. This gives

$$P([n]_p \not\to (A)_r) \leq e^{-\delta np} \sum_{k=0}^{rMn^{1-1/m_A}} \binom{n}{k} 2^{rk} p^k \leq e^{-\delta np} \left(1 + \sum_{k=1}^{rMnp/C} \left(\frac{2^r npe}{k}\right)^k\right),$$

where we used that $\binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$ if $k \geq 1$. Note that $\left(\left(\frac{a}{x}\right)^x\right)' = \left(\frac{a}{x}\right)^x (\log(a/x) - 1)$, so that, for C large enough, we can ensure that the summands are increasing and bound the sum by the largest term. Therefore,

$$P([n]_p \not\to (A)_r) \leq e^{-\delta np}\left(1 + \frac{rMnp}{C}\left(\frac{2^r Ce}{rM}\right)^{rMnp/C}\right) \leq e^{-\delta np}(2e^{\delta np/2}) = o(1),$$

by taking C large enough. $\square$

Chapter 2

Arithmetic results in sparse quasirandom sets

In this chapter we present results in the same spirit as the ones of the previous chapter, but do it in the setting of quasirandom graphs and quasirandom sets. We say a graph is quasirandom if its edge distribution has similar properties to the ones one would expect of a random graph, in a sense we make precise later on, and a set is quasirandom if its Cayley graph is. Therefore, our goal is to prove sparse analogues to theorems in the integers for quasirandom subsets. Most of this chapter follows Conlon, Fox and Zhao's paper [6] on a sparse counting lemma and its applications.

First, we introduce the notion of quasirandom graphs and sets, and then state a sparse regularity lemma and the sparse counting lemma from Conlon, Fox and Zhao [6]. With these tools in hand, we then prove a sparse removal lemma, which allows one to establish the analogue to the result on density regular matrices from Chapter 1 for single equations.

2.1 Quasirandom graphs and sets

As we have seen in Chapter 1, the random binomial graph $G(n,p)$ and the random subset $[n]_p$ are very well-behaved, and have many of the properties the complete graph or the whole set $[n]$ enjoy. Indeed, we saw that many theorems in the complete setting hold even for the sparse setting, with $p = o(1)$. In order to determine when a deterministic construction of a graph had this kind of behaviour, Thomason developed the notion of jumbled graphs in his papers [22] and [23]. Following [6], we define a slight variation of Thomason's definition, and say a graph G is (p, α)-*jumbled* if

$$|e(X,Y) - p|X||Y|| \leq \alpha\sqrt{|X||Y|}.$$

for all subsets $X, Y \subseteq V(G)$, so that the edges between any pair of subsets are approximately what one would expect in a random graph. Using Chernoff-type bounds, one can see that the random graph $G(n,p)$ is $\left(p, O\left(\sqrt{pn}\right)\right)$-jumbled with high probability.

Additionally, Chung, Graham, and Wilson [4] studied what are commonly called quasirandom graphs, proving that many expected properties of random graphs are equivalent. Some of the most important ones are the following.

Notation 2.1. Given a graph G and a subgraph $H \subseteq G$, by $G(H)$ we mean the number of labeled occurrences of H in G.

Theorem 2.2. *For a fixed $0 < p < 1$ and a family of graphs Γ_n, the following properties are equivalent.*

1. *For all graphs H of fixed size,*

$$\Gamma_n(H) = (1 + o(1))n^{v(H)}p^{e(H)}.$$

2. *The previous property holds for C_4. That is,*

$$\Gamma_n(C_4) = (1 + o(1))(pn)^4.$$

3. *The graphs Γ_n are $(p, o(n))$ jumbled.*

4. *The graphs satisfy $e(\Gamma_n) \geq (1 + o(1))n^2/4$ and the largest eigenvalues of their adjacency matrices satisfy $\lambda_1(\Gamma_n) = (1 + o(1))n/2$ and $\lambda_2(\Gamma_n) = o(n)$.*

Note that all properties are what would one expect from the random graph $G(n, p)$, and that the original statement in [4] has a slightly different language but the content of the properties is essentially the same. The second property is perhaps the most surprising, which tells us that having the expected count of squares is enough to guarantee the expected count of all other graphs and the corresponding jumbledness.

Since all the properties are equivalent and can be verified deterministically, any graph that satisfies one of them may be called a quasirandom graph. However, some of these properties are not equivalent in the sparse setting, that is, when $p = o(1)$, where the analysis is somewhat more delicate. Therefore, we will work with (p, α)-jumbled graphs and sets for adequate values of α.

Let us more formally define a (p, α)-jumbled set.

Definition 2.3. Given a group G and a set X, the Cayley graph of S has vertex set G and all edges of the form (x, xs) with $s \in S$.

Definition 2.4. A subset X of a group G is (p, α)-jumbled set if its Cayley graph is.

Chung and Graham [3] also discuss the notion of quasirandom sets in abelian groups, obtaining similar results to Theorem 2.2 where they give several equivalent conditions for quasirandomness. One of the conditions is precisely the quasirandomness of the Cayley graph, which is the main property Conlon, Fox, and Zhao take for their definition of jumbledness.

2.2 A sparse removal lemma

The arithmetic statement we prove is a sparse removal lemma for groups. This will allow us to prove a relative Roth's theorem. The usual version of the theorem states the following.

Theorem 2.5 (Roth's theorem)**.** *For every $\delta > 0$ there exists n_0 such that the following holds for $n \geq n_0$. If a set $X \subseteq [n]$ satisfies $|X| \geq \delta n$, then it contains an arithmetic progression.*

In other words, if X has no arithmetic progressions of length 3, then $|X| = o(n)$. Note that this is Szemerédi's theorem (Theorem 0.1) in the case $k = 3$.

Our goal is to prove a relative Roth's theorem, that is, the analogue of Roth's theorem substituting $[n]$ by a quasirandom set as the ambient set. We prove something slightly more general, a sparse group removal lemma. In order to prove the removal lemma, we will use the regularity method. This method uses two main ingredients, a regularity lemma and a counting lemma. We use sparse versions of these results.

The regularity lemma splits a graph into different pieces in such a way that the edges between pieces are evenly distributed. In order to quantify this distribution, we define the discrepancy condition.

Definition 2.6. A bipartite graph with vertex set $X \times Y$ satisfies the discrepancy condition $\mathrm{DISC}(q, p, \varepsilon)$ if

$$|e(U, V) - q|U||V|| \leq \varepsilon p|U||V|$$

for all $U \subseteq X$ and $V \subseteq Y$.

If we set $p = 1$ we obtain the usual discrepancy in Szemerédi's regularity lemma, that is, the one in the dense setting. One must think about the parameter p as the density of the ambient graph, and q as density of a given subgraph. Therefore, one is interested in having error terms of the order of magnitude of the ambient graph.

We state a version of the sparse regularity lemma already tailored to our application. It is equivalent to [6, Lemma 8.1] adapted to multipartite graphs. Note that the reference for the sparse regularity lemma [10] already contemplates multipartite graphs, so it is not hard to adapt the proof.

Notation 2.7. Given a graph Γ and vertex subsets X_i, X_j, by $(X_i, X_j)_\Gamma$ we mean the restriction of Γ to the edges betweeen X_i and X_j.

Theorem 2.8. *For every $\varepsilon, \alpha > 0$ and positive integer m, there exists $c > 0$ and a positive integer M such that the following holds. Let Γ be a multipartite graph with vertex sets $X_1, \ldots, X_r$ of size $|X_i| = n$ such that $(X_i, X_{i+1})_\Gamma$ is a (p, cpn)-jumbled graph and $(X_i, X_j)_\Gamma = \varnothing$ for $j \neq i + 1$. Then any subgraph G of Γ is such that there is a subgraph G' of G with $e(G') \geq e(G) - 4\alpha e(\Gamma)$ and an equitable partition of the vertex set into k pieces $V_1, V_2, \ldots, V_k$ with $m \leq k \leq M$ such that the following conditions hold.*

1. *The partition is a refinement of $X_1, \ldots, X_r$.*

2. *Every non-empty subgraph $(V_i, V_j)_{G'}$ has density $q_{ij} \geq \alpha p$ and satisfies $\mathrm{DISC}(q_{ij}, p, \varepsilon)$.*

We only state the sparse counting lemma in the case of cycles and with only a lower bound on the count, which is all we need to prove the removal lemma. This counting lemma is proven in [6, Proposition 7.2]. We introduce some notation first.

Notation 2.9. Given a set of vertex subsets $X_1, \dots X_m$ and a graph H with m vertices, we write $p(H)$ for the expected number of copies of H in the multipartite random graph $X_1, \dots, X_m$ with density p. That is,

$$p(H) = \sum_{ab \in e(H)} p|X_a||X_b|.$$

For the statement of the theorem and the rest of the section, let us define $k_3 = 3$, $k_4 = 2$ and, for $l \geq 5$, define

$$k_l = \begin{cases} 1 + \frac{1}{l-3} \text{ when } l \text{ is odd.} \\ 1 + \frac{1}{l-4} \text{ when } l \text{ is even.} \end{cases}$$

These will be the required exponents on the jumblednes hypothesis. Note too that sums in indices in the following statement are taken modulo l.

Theorem 2.10. *Let C_l be a cycle of length $l \geq 3$. Let Γ a graph with vertex subsets $X_1, \dots, X_l$ satisfying the following conditions with $c > 0$, $\varepsilon > 0$, and G a subgraph of Γ:*

- *The bipartite graph $(X_i, X_{i+1})_\Gamma$ is $(p, cp^{k_l}\sqrt{|X_a||X_b|})$-jumbled.*
- *The bipartite graph $(X_i, X_{i+1})_G$ satisfies $\mathrm{DISC}(q, p, \varepsilon)$, with $0 \leq q \leq p$.*

Then

$$G(C_l) \geq q(C_l) - \theta p(C_l),$$

with $\theta \leq 100(\varepsilon^{1/2l} + lc^{2/3})$.

From the sparse counting lemma and the sparse regularity lemma one can deduce the graph removal lemma, which in the case of cycles reads as follows.

Theorem 2.11. *For every $l \geq 3$ and $\varepsilon > 0$ there exist $c > 0$ and $\delta > 0$ such that the following holds. Let Γ be an l-multipartite graph with $X_1, \dots, X_l$ vertex subsets of size n such that there are only edges between X_i and X_{i+1} and $(X_i, X_{i+1})_\Gamma$ are $(p, cp^{k_m}n)$-jumbled for all i. Then any subgraph with less than $\delta p^l n^l$ copies of C_l can be made C_l-free by removing less than εpn^2 edges.*

Proof. Let G be the subgraph from the statement. Apply Theorem 2.8 with α, $\varepsilon_{2.8}$ and m to be fixed later on. We then obtain a subgraph $G' \subseteq G$ which we claim contains no C_l. Suppose that it does. Then there exist $V_1, \dots, V_l$ with $V_i \in X_i$, with $(V_i, V_{i+1})_{G'}$ non-empty and therefore with $\mathrm{DISC}(q_i, p, \varepsilon_{2.8})$ and $q_i \geq \alpha p$ on account of Theorem 2.8.

Then, applying the sparse counting lemma (Theorem 2.10), we obtain that

$$G(C_l) \geq G'(C_l) \geq \prod_{i=1}^{l} q_i|V_i| - \theta p^l \prod_{i=1}^{l} |V_i| \geq (\alpha^l - \theta)p^l \left(\frac{n}{M}\right)^l,$$

with $\theta \leq 100(\varepsilon_{2.8}^{1/2l} + lc^{2/3})$. Therefore, ensuring that $\varepsilon_{2.8}$ and c are small enough, we have that $G(C_l) \geq \frac{\alpha^l}{2M^l}(pn)^l$. If we then fix $\delta = \frac{\alpha^l}{4M^l}$ we reach a contradiction with the original hypothesis on the copies of C_l. Therefore G' is C_l-free.

Finally, let us see that we can take G' without removing too many edges. Indeed, by Theorem 2.8 we know that

$$e(G') \geq e(G) - 4\alpha e(\Gamma) \geq e(G) - 8\alpha lpn^2,$$

so taking $\alpha \leq \varepsilon/8l$ guarantees the needed conclusion. □

2.3 A group removal lemma and Roth's theorem

With these results in hand, we may prove the group removal lemma. This reproduces the combinatorial proof of the group removal lemma done by Král', Serra and Vena [11] in the sparse setting.

Theorem 2.12. *For every $m \geq 3$ and $\varepsilon > 0$ there exist $c, \delta > 0$ such that the following holds. Let G be a finite group of size n and $Y_1, \dots, Y_m \subseteq G$ be $(p, cp^{k_m}n)$-jumbled subsets. Suppose that $X_1, \dots, X_m$ are subsets $X_i \subseteq Y_i$ with less than $\delta|Y_1|\cdots|Y_m|/n$ solutions to the equation $x_1 \cdots x_m = 1$ satisfying $x_i \in X_i$. Then one can remove less than $\varepsilon|Y_i|$ elements from every subset X_i and obtain solution-free subsets.*

Proof. Let us build an m-partite graph Γ with vertex set $G \times [m]$ and the Cayley graph of every Y_i as edges, so that $(g, i) \sim (gy_i, i+1)$ for all $g \in G$ and $y_i \in Y_i$. Writing $\Gamma_i = G \times \{i\}$ for the i-th piece of the graph, we have that the subgraphs $(\Gamma_i, \Gamma_{i+1})_\Gamma$ are $(p, cp^{k_m}n)$-jumbled by definition. Since they are at least (p, cpn)-jumbled, by taking $c < 1/2$ we can ensure that

$$\left|n|Y_i| - pn^2\right| = \left|e(\Gamma_i, \Gamma_{i+1}) - pn^2\right| \leq pn^2/2,$$

so that $pn/2 \leq |Y_i| \leq 2pn$ for all i.

Let H the subgraph given by edges $(g, i) \sim (gx_i, i+1)$ for all $x_i \in X_i$. Note that paths of length m in H starting in Γ_1 are of the form $(g, 1) \sim (gx_1, 2) \sim \cdots \sim (gx_1 \dots x_{m-1}, m) \sim (gx_1 \dots x_m, 1)$, so that copies of C_l in H correspond exactly to an element $g \in G$ and a solution to $x_1 \cdots x_m = 1$ with $x_i \in X_i$. Hence, by hypothesis we have less than $\delta|Y_1|\cdots|Y_m| \leq 2^m\delta p^m n^m$ copies of C_l in H.

Applying the cycle removal lemma (Theorem 2.11) with $\varepsilon_{2.11}$ to be fixed later on, and ensuring that $\delta < \delta_{2.11}/(2^m)$ and c is small enough to apply the theorem, we obtain a collection of edges F with $|F| \leq \varepsilon_{2.11}pn^2$ such that, after deleting them, the graph is C_l-free. However, we want to remove elements of X_i instead of edges of H. To do so, we choose the most *popular* terms, concretely those $x_i \in X_i$ such that at least n/m edges $(g, i) \sim (gx_i, i+1)$ belong to F.

Suppose that there exists a solution $x_1 \cdots x_m = 1$ with $x_i \in X_i$ after deleting these elements. Then there are n different disjoint cycles in H given by $(g, 1) \sim (gx_1, 2) \sim \cdots \sim (gx_1 \cdots x_{m-1}, m)$ for all $g \in G$, and all of them must contain an edge in F, so that there would be an x_i with at least n/m edges $(g, i) \sim g(x_i, i+1)$ in F, a contradiction. Hence, after deleting these elements, there are no solutions left to the equation. Finally, note that we have erased at most $m\varepsilon_{2.11}pn$ elements, so setting $\varepsilon_{2.11} = \varepsilon/(2m)$ finishes the proof. □

From the group removal lemma a relative Roth's theorem follows. We state it in a slightly more general form, that is, for arbitrary translation-invariant equations instead of $x + z = 2y$.

Corollary 2.13 (Relative Roth's theorem)**.** *For any $m \geq 3$ and $\delta > 0$ there exists a positive integer n_0, and constants $C > 0$ and $c > 0$ such that the following holds for $n \geq n_0$. Let Y be a $(p, cp^{k_m} n)$-jumbled subset of $\mathbb{Z}_n$ with $p^{m-1} \geq Cn^{-(m-2)}$. Let $a_1 x_1 + \cdots + a_m x_m = 0$ be an equation with coefficients $a_i \in \mathbb{Z}_n$ such that $a_i \neq 0$, $\gcd(a_i, n) = 1$ and $a_1 + \cdots + a_m = 0$. If $X \subseteq Y$ satisfies $|X| \geq \delta |Y|$, then there is a non-trivial solution to the equation with $x_i \in X$ for all i, that is, a solution with $x_i \neq x_j$ for some i and j.*

Proof. Since $\gcd(a_i, |G|) = 1$, multiplication by a_i is a bijection and the sets $Y_i = a_i Y$ are $(p, cp^{k_m} n)$-jumbled. Suppose that the equation has only trivial solutions, which, on account of jumbledness of Y, may be bounded by

$$|X| \leq |Y| \leq 2pn = \leq 2 \frac{p^m n^{m-1}}{p^{m-1} n^{m-2}} \leq \frac{2^{m+1} |Y|^m}{Cn}$$

solutions if $c < 1/2$.

Hence, for C large enough we may apply the sparse group removal lemma (Theorem 2.12) to the sets $X_i = a_i X$ with $\varepsilon = \delta/(m+1)$, and removing a total of at most $m\varepsilon |Y| \leq m/(m+1)|X|$ elements from X we should remove all solutions to the equation. However, there still remain $|X|/(m+1)$ elements in $|X|$ which provide a trivial solution, a contradiction. The contradiction comes from asssuming that there are only trivial solutions. ∎

Bibliography

[1] József Balogh, Robert Morris, and Wojciech Samotij. "Independent sets in hypergraphs". In: *J. Amer. Math. Soc.* 28.3 (2015), pp. 669–709. ISSN: 0894-0347. DOI: 10.1090/S0894-0347-2014-00816-X.

[2] József Balogh, Robert Morris, and Wojciech Samotij. "The method of hypergraph containers". In: *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures*. World Sci. Publ., Hackensack, NJ, 2018, pp. 3059–3092.

[3] F. R. K. Chung and R. L. Graham. "Quasi-random subsets of Z_n". In: *J. Combin. Theory Ser. A* 61.1 (1992), pp. 64–86. ISSN: 0097-3165. DOI: 10.1016/0097-3165(92)90053-W.

[4] F. R. K. Chung, R. L. Graham, and R. M. Wilson. "Quasi-random graphs". In: *Combinatorica* 9.4 (1989), pp. 345–362. ISSN: 0209-9683. DOI: 10.1007/BF02125347.

[5] D. Conlon and W. T. Gowers. "Combinatorial theorems in sparse random sets". In: *Ann. of Math. (2)* 184.2 (2016), pp. 367–454. ISSN: 0003-486X. DOI: 10.4007/annals.2016.184.2.2.

[6] David Conlon, Jacob Fox, and Yufei Zhao. "Extremal results in sparse pseudorandom graphs". English. In: *Adv. Math.* 256 (2014), pp. 206–290. ISSN: 0001-8708. DOI: 10.1016/j.aim.2013.12.004.

[7] P. Frankl, R. L. Graham, and V. Rödl. "Quantitative theorems for regular systems of equations". In: *J. Combin. Theory Ser. A* 47.2 (1988), pp. 246–261. ISSN: 0097-3165. DOI: 10.1016/0097-3165(88)90020-9.

[8] P. Frankl and V. Rödl. "Large triangle-free subgraphs in graphs without K_4". In: *Graphs Combin.* 2.2 (1986), pp. 135–144. ISSN: 0911-0119. DOI: 10.1007/BF01788087.

[9] Ehud Friedgut, Vojtěch Rödl, and Mathias Schacht. "Ramsey properties of random discrete structures". In: *Random Structures Algorithms* 37.4 (2010), pp. 407–436. ISSN: 1042-9832. DOI: 10.1002/rsa.20352.

[10] Y. Kohayakawa. "Szemerédi's regularity lemma for sparse graphs". In: *Foundations of computational mathematics (Rio de Janeiro, 1997)*. Springer, Berlin, 1997, pp. 216–230.

[11] Daniel Král, Oriol Serra, and Lluís Vena. "A combinatorial proof of the removal lemma for groups". In: *J. Combin. Theory Ser. A* 116.4 (2009), pp. 971–978. ISSN: 0097-3165. DOI: 10.1016/j.jcta.2008.12.003.

[12] Rajko Nenadov and Angelika Steger. “A short proof of the random Ramsey theorem”. In: *Combin. Probab. Comput.* 25.1 (2016), pp. 130–144. ISSN: 0963-5483. DOI: 10.1017/S0963548314000832.

[13] Richard Rado. “Studien zur Kombinatorik”. In: *Mathematische Zeitschrift* 36 (1933), pp. 424–470.

[14] Vojtech Rödl and Andrzej Ruciński. “Rado partition theorem for random subsets of integers”. In: *Proc. London Math. Soc. (3)* 74.3 (1997), pp. 481–502. ISSN: 0024-6115. DOI: 10.1112/S0024611597000178.

[15] Vojtěch Rödl and Andrzej Ruciński. “Random graphs with monochromatic triangles in every edge coloring”. In: *Random Structures Algorithms* 5.2 (1994), pp. 253–270. ISSN: 1042-9832. DOI: 10.1002/rsa.3240050202.

[16] Vojtěch Rödl and Andrzej Ruciński. “Threshold functions for Ramsey properties”. In: *J. Amer. Math. Soc.* 8.4 (1995), pp. 917–942. ISSN: 0894-0347. DOI: 10.2307/2152833.

[17] Juanjo Rué, Oriol Serra, and Lluis Vena. “Counting configuration-free sets in groups”. In: *European J. Combin.* 66 (2017), pp. 281–307. ISSN: 0195-6698. DOI: 10.1016/j.ejc.2017.06.027.

[18] David Saxton and Andrew Thomason. “Hypergraph containers”. In: *Invent. Math.* 201.3 (2015), pp. 925–992. ISSN: 0020-9910. DOI: 10.1007/s00222-014-0562-8.

[19] Mathias Schacht. “Extremal results for random discrete structures”. In: *Ann. of Math. (2)* 184.2 (2016), pp. 333–365. ISSN: 0003-486X. DOI: 10.4007/annals.2016.184.2.1.

[20] Christoph Spiegel. “A note on sparse supersaturation and extremal results for linear homogeneous systems”. In: *Electron. J. Combin.* 24.3 (2017), Paper No. 3.38, 19.

[21] E. Szemerédi. “On sets of integers containing no k elements in arithmetic progression”. In: *Acta Arith.* 27 (1975), pp. 199–245. ISSN: 0065-1036. DOI: 10.4064/aa-27-1-199-245.

[22] Andrew Thomason. “Pseudorandom graphs”. In: *Random graphs ’85 (Poznań, 1985)*. Vol. 144. North-Holland Math. Stud. North-Holland, Amsterdam, 1987, pp. 307–331.

[23] Andrew Thomason. “Random graphs, strongly regular graphs and pseudorandom graphs”. In: *Surveys in combinatorics 1987 (New Cross, 1987)*. Vol. 123. London Math. Soc. Lecture Note Ser. Cambridge Univ. Press, Cambridge, 1987, pp. 173–195.