Probabilistic methods in Ramsey theory and additive combinatorics

Gabriel Dahia Fernandes

Instituto de Matemática Pura e Aplicada
Rio de Janeiro, Brasil

Supervised by Marcelo Campos & Robert Morris

2026

Abstract

This thesis presents results obtained using probabilistic methods in the areas of graph Ramsey theory and additive combinatorics. In Ramsey theory, we resolve a well-known conjecture of Erdős by proving an exponential upper bound on the induced Ramsey number of an arbitrary graph. In additive combinatorics, we obtain tight bounds for the typical independence number of a sparse random Cayley graph, extending a result of Green and Morris, and prove an essentially optimal result on the length of arithmetic progressions contained in sumsets of subsets of random sparse sets, significantly improving a result of Hamel and Łaba.

The main result in Chapter 2, which was obtained jointly with Aragão, Campos, Filipe and Marciano, is an exponential upper bound for the induced Ramsey number $\mathrm{r}_{\mathrm{ind}}(H;r)$ of a graph H: the minimum number N such that there exists a graph with N vertices for which all r-colourings of its edges contain a monochromatic induced copy of H. That is, we show the existence of a constant $C > 0$ such that, for every $r \in \mathbb{N}$ and every graph H on k vertices, these numbers satisfy

$$\mathrm{r}_{\mathrm{ind}}(H;r) \leqslant r^{Crk}.$$

When $r = 2$, this resolves a conjecture of Erdős from 1975, and for $r > 2$, it answers a question of Conlon, Fox and Sudakov in a strong form. The main technical tool in this proof is a novel strengthening of the celebrated method of hypergraph containers of Balogh, Morris and Samotij, and Saxton and Thomason, which applies to sets in which the edges are not "well-distributed".

Chapter 3 studies Cayley sum graphs G_S, where $S \subseteq \mathbb{Z}_n$ and G_S is defined to have vertex set $\mathbb{Z}_n$ and an edge between two distinct vertices $x, y \in \mathbb{Z}_n$ if $x + y \in S$. Green and Morris established in 2016 that, if S is a uniformly random subset of $\mathbb{Z}_n$, then

$$\alpha(G_S) = \big(1 + o(1)\big)\alpha\big(\mathbb{G}(n, 1/2)\big)$$

with high probability, where $\mathbb{G}(n, 1/2)$ is a uniform random graph with n vertices. In Chapter 3, which is based on joint work with Campos and Marciano, we give the first extension of their result to the more challenging setting where S is a sparse p-random subset of $\mathbb{Z}_n$, establishing that, for all $(\log n)^{-1/80} \leqslant p \leqslant 1/2$, we have

$$\alpha(G_S) = \big(1 + o(1)\big)\alpha\big(\mathbb{G}(n, p)\big)$$

with high probability. The main technical component in the proof is a result in discrete convex

geometry: if $A \subseteq \mathbb{R}^d$ has $\operatorname{rank}(A) = d$, then there exists $T \subseteq A$ with $|T| = O_\gamma(d)$ such that

$$|A + T| \geqslant (1 - \gamma)\frac{d+1}{2}|A|.$$

Apart from the $(1-\gamma)/2$ term, this result generalises a classical lemma of Freĭman, which implies that $|A + A| \geqslant (d+1)|A| - \binom{d+1}{2}$ under the same conditions.

The final result of this thesis, proved in work with Campos and Kohayakawa and presented here in Chapter 4, is that, for every $\varepsilon > 0$, if S is a p-random subset of $[n]$ with $p \geqslant n^{-1/2+\varepsilon}$, then, with high probability, for any A, $B \subseteq S$ such that $|A| \geqslant \alpha|S|$ and $|B| \geqslant \beta|S|$, $A + B$ contains an arithmetic progression of length at least

$$\exp(c(\alpha\beta \log n)^{1/2} - \log\log n)$$

where $c > 0$ is an absolute constant. Our condition on p is best possible up to the ε term, and the length of the progressions that we are able to guarantee matches, up to the constant c, the best value currently known in the case $p = 1$, a result due to Green. Hamel and Łaba in 2008 proved a weaker version of this result, which required $A = B$ and $p \geqslant n^{-1/140}$, and also obtained shorter progressions. To prove our result, we employ the asymmetric container theorem for sumsets of Campos, Coulson, Serra and Wötzel, and extend the "structure from almost periodicity" results of Croot, Łaba and Sisask to a "robust" setting.

Resumo

Esta tese apresenta resultados obtidos usando métodos probabilísticos nas áreas de teoria de Ramsey em grafos e combinatória aditiva. Em teoria de Ramsey, resolvemos uma famosa conjectura de Erdős, provando uma cota superior exponencial para o número de Ramsey induzido de um grafo arbitrário. Em combinatória aditiva, obtemos cotas justas para o número de independência típico de um grafo de Cayley aleatório esparso, estendendo um resultado de Green e Morris, e provamos um resultado essencialmente ótimo sobre o comprimento de progressões aritméticas contidas em conjuntos soma de subconjuntos de conjuntos aleatórios esparsos, melhorando significativamente um resultado de Hamel e Łaba.

O resultado principal no Capítulo 2, obtido em conjunto com Aragão, Campos, Filipe e Marciano, é uma cota superior exponencial para o número de Ramsey induzido $\mathrm{r}_{\mathrm{ind}}(H;r)$ de um grafo H: o menor número N tal que existe um grafo com N vértices para o qual todas as r-colorações de suas arestas contêm uma cópia induzida monocromática de H. Isto é, mostramos a existência de uma constante $C > 0$ tal que, para todo $r \in \mathbb{N}$ e todo grafo H com k vértices, esses números satisfazem

$$\mathrm{r}_{\mathrm{ind}}(H;r) \leqslant r^{Crk}.$$

Quando $r = 2$, isso resolve uma conjectura de Erdős de 1975, e, para $r > 2$, responde em forma forte a uma pergunta de Conlon, Fox e Sudakov. A principal ferramenta técnica nesta prova é um novo fortalecimento do célebre método dos contêineres em hipergrafos de Balogh, Morris and Samotij, e Saxton and Thomason, que se aplica a conjuntos em que as arestas não são "bem-distribuídas".

O Capítulo 3 estuda grafos de soma de Cayley G_S, onde $S \subseteq \mathbb{Z}_n$ e G_S é definido por ter como conjunto de vértices $\mathbb{Z}_n$ e uma aresta entre dois vértices distintos $x, y \in \mathbb{Z}_n$ se $x + y \in S$. Green and Morris estabeleceram originalmente que, se S é um subconjunto aleatório uniforme de $\mathbb{Z}_n$, então

$$\alpha(G_S) = \big(1 + o(1)\big)\alpha\big(\mathbb{G}(n, 1/2)\big)$$

com alta probabilidade, onde $\mathbb{G}(n, 1/2)$ é um grafo aleatório uniforme com n vértices. No Capítulo 3, baseado em um trabalho conjunto com Campos e Marciano, apresentamos a primeira extensão do resultado deles para o cenário mais desafiador onde S é um subconjunto p-aleatório de $\mathbb{Z}_n$ estabelecendo que, para todo $(\log n)^{-1/80} \leqslant p \leqslant 1/2$, vale

$$\alpha(G_S) = \big(1 + o(1)\big)\alpha\big(\mathbb{G}(n, p)\big)$$

com alta probabilidade. O principal componente técnico nessa demonstração é um resultado em

geometria convexa discreta: se $A \subseteq \mathbb{R}^d$ tem $\operatorname{rank}(A) = d$, então existe $T \subseteq A$ com $|T| = O_\gamma(d)$ tal que

$$|A + T| \geqslant (1 - \gamma)\frac{d+1}{2}|A|.$$

À parte do termo $(1-\gamma)/2$, esse resultado generaliza um lema clássico de Freĭman, que implica que $|A + A| \geqslant (d+1)|A| - \binom{d+1}{2}$ sob as mesmas condições.

O último resultado desta tese, provado em um trabalho em conjunto com Campos e Kohayakawa e apresentado aqui no Capítulo 4, é que, para todo $\varepsilon > 0$, se S é um subconjunto p-aleatório de $[n]$ com $p \geqslant n^{-1/2+\varepsilon}$, então, com alta probabilidade, para quaisquer A, $B \subseteq S$ tais que $|A| \geqslant \alpha|S|$ e $|B| \geqslant \beta|S|$, $A + B$ contém uma progressão aritmética de comprimento pelo menos

$$\exp(c(\alpha\beta \log n)^{1/2} - \log\log n)$$

onde $c > 0$ é uma constante absoluta. Nossa condição sobre p é ótima exceto o termo ε, e o comprimento das progressões que conseguimos garantir coincide, possivelmente com outra constante c, com o melhor valor atualmente conhecido no caso $p = 1$, originalmente provado por Green. Hamel and Łaba provaram em 2008 uma versão mais fraca deste resultado, que exigia $A = B$ e $p \geqslant n^{-1/140}$, e também obtinha progressões mais curtas. Para provar esse resultado, empregamos o teorema assimétrico dos contêineres para conjuntos soma de Campos, Coulson, Serra and Wötzel, e estendemos os resultados de "estrutura a partir de quase-periodicidade" de Croot, Łaba and Sisask para valerem de forma "robusta".

Contents

Chapter 1

Introduction

In 1916, Schur [119] proved that Fermat's last theorem fails in the group $\mathbb{Z}_p$ for all sufficiently large primes p, that is, there exist solutions to the congruence $x^m + y^m \equiv z^m \pmod{p}$ for all $m \in \mathbb{N}$. The proof of this theorem relies on a key lemma, now known as Schur's theorem, which states that any finite colouring of the positive integers contains a monochromatic set $\{x, y, z\}$ such that $x + y = z$. A decade later, van der Waerden [128] proved an analogue of Schur's theorem for patterns of the form

$$\{a, a + d, \ldots, a + (k - 1)d\},$$

that is, he showed that every finite colouring of $\mathbb{N}$ contains arbitrarily long arithmetic progressions.

These two results, by Schur and van der Waerden, belong to an area called additive (or arithmetic) combinatorics: the (combinatorial) study of objects that support the addition operation, usually relying on a combination of tools from graph theory, analysis, linear algebra and ergodic theory. The above theorems concern two of the central objects studied in the area: k-term arithmetic progressions (or k-APs) and the sumset

$$A + B = \{a + b : a \in A, b \in B\}.$$

Indeed, we can rephrase Schur's theorem as saying that every finite colouring of $\mathbb{N}$ contains a colour class that is not *sum-free*, meaning that $(A + A) \cap A \neq \emptyset$.

This phenomenon, whereby some structure is inevitable in finite partitions of some collection, is not unique to additive objects. Ramsey [107] proved, in 1930, that for every finite graph H and every $r \in \mathbb{N}$, there exists $N \in \mathbb{N}$ such that every r-colouring of the edges of K_N contains a monochromatic copy of H, a result which significantly generalises Schur's theorem, an easy consequence of the case $H = K_3$. This was later rediscovered by Erdős and Szekeres [54], who in 1935 started the systematic study of Ramsey theory and its applications.

In this thesis we will explore another beautiful connection between Ramsey theory and additive combinatorics: the use of *probabilistic* techniques to prove results in these areas. In particular, we will study properties of random graphs or sets of integers, and how determining carefully chosen properties of these objects can be used to resolve extremal problems.

One of the earliest examples of this approach is from 1947, when Erdős [49] proved an exponential lower bound for $\mathrm{R}(k) = \mathrm{R}(k,k)$. In his proof, he showed that a random colouring of the edges of K_N avoids monochromatic copies of K_k by establishing that the expected number of the latter is less than 1 when $N = 2^{k/2}$. If we then interpret the value of this expectation as an average over all colourings, we immediately obtain the existence of a single colouring where there are no such cliques, and therefore every k-clique receives at least two colours. Combining this with the Erdős–Szekeres result, one obtains the bounds

$$2^{k/2} \leqslant \mathrm{R}(k) \leqslant 4^k, \tag{1.1}$$

which remained virtually unchanged until very recently (see Section 1.1).

Another famous early proof using the probabilistic method is the following result of Erdős [50], proved in 1965: every set $A \subseteq \mathbb{Z}$ contains a sum-free subset S satisfying

$$|S| \geqslant |A|/3.$$

The simplicity of this proof, which considers the set $\lambda A \bmod p$, where $\lambda \in \mathbb{Z}_p$ is chosen randomly, and p is a suitably large prime, stands in contrast with the fact that one cannot obtain a proportion larger than $1/3$ asymptotically, as was shown in a famous paper of Eberhard, Green and Manners [48]. Another surprising aspect of this result is that the first non-constant additive improvement over the $1/3$ bound was only obtained very recently, in a remarkable breakthrough due to Bedert [17].

Presented only with these two early results of Erdős using the probabilistic method, one could initially be tempted to dismiss this technique as a trick that works elegantly and remarkably well in a few simple cases, but is not suited to the complex nature of modern mathematics. The main argument against this stance is the existence of numerous, complex results whose only known proof relies on the probabilistic method. For instance, this proof technique plays an important role in the record for the longest gaps between consecutive primes [55] and the densest sphere packings [31, 84], and it is also responsible for the resolution of Bourgain's slicing conjecture [85], the Kadison–Singer problem [93], the existence conjecture in design theory [82, 65] and Littlewood's conjecture about flat polynomials [12].

In this thesis, we give another example of an intricate proof that relies on the probabilistic method, and, more importantly, solves a longstanding conjecture in graph Ramsey theory. To do so, we will prove that almost all graphs with not too few vertices have this desired Ramsey property, and therefore establish that a typical large graph suffices for our proof. This latter interpretation of our first result has the same flavour as the results on additive combinatorics that we prove later on, which concern properties of random sparse sets of integers that hold with high probability[1]. We therefore give a general introduction to each of these two areas in their respective sections (Sections 1.1 and 1.2), and then specialise to the particular subfields where we obtain our results.

[1]Throughout, "with high probability" means with probability tending to 1 as the relevant parameter tends to infinity.

1.1 Ramsey theory

Denoting by $\mathrm{R}(\ell_1, \dots, \ell_r)$ the minimum N such that every r-colouring of the edges of K_N contains a monochromatic copy of K_{ℓ_i} in colour i for some $i \in [r]$, Erdős and Szekeres [54] proved an upper bound for $\mathrm{R}(k; r) := \mathrm{R}(k, \dots, k)$ by establishing the relation

$$\mathrm{R}(\ell_1, \dots, \ell_r) \leqslant \sum_{i=1}^{r} \mathrm{R}(\ell_1, \dots, \ell_i - 1, \dots, \ell_r). \tag{1.2}$$

To prove it, using the case $r = 2$ for simplicity and letting $\ell_1 = \ell$ and $\ell_2 = k$, we take an arbitrary vertex v in any 2-colouring of K_N, where N is equal to the right-hand side of (1.2). Observe that by our choice of N, either v is connected to $\mathrm{R}(\ell - 1, k)$ vertices by edges coloured red, or to $\mathrm{R}(\ell, k - 1)$ vertices by edges coloured blue. In the former case, by definition of $\mathrm{R}(\ell - 1, k)$, the neighbourhood of v contains either a red $K_{\ell-1}$, which v itself completes to a red K_ℓ, or a blue K_k. As the second case is exactly analogous, we conclude that indeed $\mathrm{R}(\ell, k) \leqslant N$, and generalising to arbitrary r, we obtain

$$\mathrm{R}(k; r) \leqslant r^{rk}. \tag{1.3}$$

While many works obtained subexponential improvements in the $r = 2$ case [69, 126, 39, 113], no improvement for $r > 2$ was known until very recently. This changed after Campos, Griffiths, Morris and Sahasrabudhe [30] obtained an exponential improvement for $\mathrm{R}(k)$, and their proof was simplified and generalised to the multicolour setting by Balister, Bollobás, Campos, Griffiths, Hurley, Morris, Sahasrabudhe and Tiba [11]. The main result of [11] is that

$$\mathrm{R}(k; r) \leqslant e^{-\delta_r k} r^{rk} \tag{1.4}$$

when r is fixed and k is sufficiently large, where $\delta_r > 0$ depends only on r. The key step in this shorter and more general proof is a geometric-flavoured lemma that replaces many of the delicate computations in the previous proof of the exponential improvement.

For the lower bound, a simple and elegant construction due to Abbott [1] using blow-ups established that

$$\mathrm{R}(k; r + s) \geqslant (\mathrm{R}(k; r) - 1) \cdot (\mathrm{R}(k; s) - 1) + 1,$$

which combined with the lower bound of Erdős for $r = 2$ yields

$$\mathrm{R}(k; r) \geqslant 2^{crk}$$

for an absolute constant $c > 0$, the value of which has since then been improved for $r > 2$. The first such construction was obtained by Conlon and Ferber [41], whose work inspired several further improvements [130, 116, 33, 10], in each case by replacing the uniform random colouring as the base of the recursion by a more sophisticated random construction. By contrast, the lower bound when $r = 2$ has only been improved by a factor of 2, and that improvement is obtained by a standard application of the Lovász Local Lemma [121].

Even closer to the topics of this thesis is the Ramsey number for arbitrary graphs H,

$$\mathrm{r}(H;r) = \min\left\{N : K_N \underset{r}{\to} H\right\}$$

where $G \underset{r}{\to} H$ denotes that in every r-colouring of $E(G)$ there exists a monochromatic copy of H. This number and $\mathrm{r}(H) := \mathrm{r}(H;2)$ satisfy $\mathrm{R}(k) = \mathrm{r}(K_k)$ and $\mathrm{R}(k;r) = \mathrm{r}(K_k;r)$, and one of the main interests in studying this variant is that we now know much better bounds for $\mathrm{r}(H)$ for certain classes of graph $H \neq K_k$.

A remarkable, and very general, result of Chvatál, Rödl, Szemerédi and Trotter [37] established that

$$\mathrm{r}(H) \leqslant C_\Delta k \tag{1.5}$$

for every graph H with k vertices and maximum degree at most Δ, for a value $C_\Delta > 0$ that depends only on Δ. This was one of the earliest applications of Szemerédi's regularity method, and so the dependency of C_Δ on Δ is of tower-type, that is, the upper bound on C_Δ is only $\log^* C_\Delta \leqslant \Delta$. Although such growth is inevitable when using the regularity method [66], Graham, Rödl and Ruciński [70] were able to vastly improve this dependency, proving that

$$C_\Delta \leqslant 2^{C\Delta(\log\Delta)^2} \tag{1.6}$$

is admissible in (1.5) for an absolute constant $C > 0$.

To avoid the use of the regularity lemma, Graham, Rödl and Ruciński established that in every two colouring of K_N, either there is a large subset of the vertices that define a very dense subgraph in one of the colour classes, or there is a large subset of vertices in which the edges in the complementary colour class are well-distributed. Another way to phrase this dichotomy, which proved very influential, is that if both colour classes have comparable densities in every large subset of vertices, then the subgraph defined by one of those colours shares some properties with a random graph of comparable density, or, for short, this colour class is a *pseudorandom* graph (a notion that is related to the foundational work of Thomason [125]).

Bounded-degree graphs H are also one of the main classes studied in another variant of $\mathrm{r}(H)$, called the size-Ramsey number and defined by

$$\hat{\mathrm{r}}(H) = \min\left\{e(G) : G \underset{2}{\to} H\right\}.$$

Like many results in Ramsey theory, the strongest lower bounds for $\hat{\mathrm{r}}(H)$ when H is a bounded-degree graph [110, 127] are proved probabilistically, but with a different perspective: one chooses H at random, typically as a random regular graph, and shows that if G has fewer edges than the desired lower bound, then with positive probability we can colour $E(G)$ to avoid a monochromatic copy of H. Interestingly, the best known upper bounds for $\hat{\mathrm{r}}(H)$ for H with bounded maximum degree are also obtained using the probabilistic method [16, 88] but choosing instead the host graph G at random and showing that, with positive probability, such a sparse random graph already forces a monochromatic copy of H in every two-colouring of its edges.

There are many other questions in Ramsey theory for which the best-known bounds use random constructions, such as the (very close to sharp) lower bounds for $\mathrm{R}(3,k)$ [83, 32, 77] and,

slightly further away from the upper bound, $\mathrm{R}(4,k)$ [94], and the existence of K_4-free graphs for which every subgraph with slightly more than half of the edges contains a triangle [59], where in fact a binomial random graph with the correct sparsity is optimal. In the next section we will discuss in greater detail another such problem: the induced Ramsey number of a graph H.

1.1.1 Induced Ramsey number

A natural and well-studied variant of $\mathrm{r}(H)$ is the setting of induced subgraphs, where we say that a graph H is an induced subgraph of another graph G if $V(H) \subseteq V(G)$ and its edge set satisfies $E(H) = \{uv \in E(G) : u, v \in V(H)\}$.

The induced Ramsey property, which we write as $G \xrightarrow{\text{ind}} H$, means that for any red/blue colouring of the edges of G, there exists an induced monochromatic copy of H (that is, a copy of H which is induced in G, and all of its edges have the same colour). We then define

$$\mathrm{r}_{\text{ind}}(H) = \min\left\{v(G) : G \xrightarrow{\text{ind}} H\right\},$$

which satisfies $\mathrm{r}_{\text{ind}}(K_k) = \mathrm{R}(k)$ for every $k \in \mathbb{N}$, since every copy of K_k in a graph G is also an induced subgraph of G. For general graphs H, however, Erdős [52] remarked that "the existence of [the induced Ramsey number] is not at all obvious."

Deuber [47], Erdős, Hajnal and Pósa [53] and Rödl [108] independently established in the 1970s that $\mathrm{r}_{\text{ind}}(H)$ is finite for every graph H. While none of these works provide an explicit dependency on k, the number of vertices of H, Erdős [52] later remarked that the best bound that one can deduce from the proofs in [47, 53, 108] is of the form

$$\mathrm{r}_{\text{ind}}(H) \leqslant 2^{2^{k^{1+o(1)}}},$$

where $o(1)$ denotes a function $f(k)$ that satisfies $f(k) \to 0$ as $k \to \infty$. Nevertheless, Erdős [51, 52] conjectured, first implicitly in 1975 and then explicitly in 1984, that the function $\mathrm{r}_{\text{ind}}(H)$ should grow at most exponentially as a function of $v(H) = k$ for every H. Note that, if true, this would be the best possible bound for $\mathrm{r}_{\text{ind}}(H)$, up to the basis of the exponent, since we have $\mathrm{r}_{\text{ind}}(K_k) = \mathrm{R}(k) \geqslant 2^{k/2}$ by (1.1).

Using the techniques of Rödl [108], one can prove the conjecture of Erdős for bipartite H. Łuczak and Rödl [92] also established that the conjecture holds whenever H has bounded degree Δ, where in fact the much stronger, polynomial bound $\mathrm{r}_{\text{ind}}(H) \leqslant k^{O_\Delta(1)}$ is true, for a constant $O_\Delta(1)$ that depends only on Δ. This can be seen as a weaker analogue of (1.5), the celebrated result of Chvatál, Rödl, Szemerédi and Trotter [37], to the induced setting.

The problem is much harder for general non-bipartite H, however, and the next significant advance was not obtained until almost 25 years later, by Kohayakawa, Prömel and Rödl [86]. By taking G (the host graph) to be a random graph built using projective planes, they showed, among other results, that

$$\mathrm{r}_{\text{ind}}(H) \leqslant k^{O(k \log k)} \tag{1.7}$$

for every graph H with k vertices, where, in (1.7) and throughout, $f(x) = O(g(x))$ means that there exists a constant $C > 0$ such that $f(x) \leqslant Cg(x)$ for all $x > 0$. Their host graph G is random

and is built using projective planes, but, despite this randomness, their embedding method is deterministic and relies only on pseudorandom properties of G. The two main properties that they use are that every sufficiently large pair of vertex subsets has edge density very close to its average value, and that collections of large disjoint subsets of vertices in G induce large blow-ups of H.

Fox and Sudakov [57] gave a different proof of the bound in (1.7) using a different but still deterministic embedding procedure. Their embedding technique has a weaker requirement on the properties of G; in particular, it requires only that the density of edges between large sets is close to the average density of the graph. As there are explicit constructions of such graphs, e.g. the Paley graph, this result proves that an explicit graph attains the bound (1.7). The approach in [57] is in fact even more general, and applies to Ramsey-type theorems in the setting where one fixed H is forbidden as an induced subgraph. This became influential for other notorious conjectures, like the Erdős–Hajnal problem (see e.g. [25, 103, 104]).

A few years later, Conlon, Fox and Sudakov [43] removed a factor of $\log k$ from the exponent in (1.7) and showed, also using an explicit graph, that

$$\mathrm{r}_{\mathrm{ind}}(H) \leqslant k^{O(k)}. \tag{1.8}$$

In order to prove (1.8), the authors of [43] developed a general method for proving Ramsey-type theorems using pseudorandom properties of the host graph G. For example, their method also allowed them to improve (1.6), the bound on C_Δ in the celebrated result of Graham, Rödl and Ruciński [70] on the Ramsey numbers of bounded degree graphs, to

$$\mathrm{r}(H) \leqslant 2^{O(\Delta \log \Delta)} k,$$

which is a $\log \Delta$ factor in the exponent away from the known lower bound. Briefly, the approach of Conlon, Fox and Sudakov [43] uses the same dichotomy as in [70] (which implicitly also appeared in [86]), but relies on a more symmetric approach to this duality.

Prior to the work presented in Chapter 2, (1.8) was the best known bound for $\mathrm{r}_{\mathrm{ind}}(H)$ for general graphs H. In that chapter, we present an affirmative resolution to the conjecture of Erdős about the growth of the induced Ramsey number $\mathrm{r}_{\mathrm{ind}}(H)$ for all graphs H.

Theorem 1.1. *There exists a constant $C > 0$ such that*

$$\mathrm{r}_{\mathrm{ind}}(H) \leqslant 2^{Ck}$$

for every graph H with k vertices.

We will also consider the induced Ramsey number for r-colourings. Define $\mathrm{r}_{\mathrm{ind}}(H; r)$ to be the *r-colour induced Ramsey number* of a graph H; that is, the minimum number of vertices of a graph G such that every r-colouring $\chi\colon E(G) \to [r]$ of the edges of G contains an induced monochromatic copy of H. The techniques used in [43] and [86] do not work in this more general setting, and provide no bounds when $r \geqslant 3$, but Fox and Sudakov [58] introduced a different

approach in 2009, which can be used to show that

$$\mathrm{r}_{\mathrm{ind}}(H;r) \leqslant r^{O(rk^2)}, \tag{1.9}$$

even though the authors of [58] only state the weaker bound $\mathrm{r}_{\mathrm{ind}}(H;r) \leqslant r^{O(rk^3)}$, because their focus in that paper was on fixed k and large r. Their proof can easily be modified to obtain (1.9), and that bound was explicitly obtained by Balogh and Samotij [15] using different techniques (see below).

The large gap between (1.9) and the known bounds for the $r = 2$ case motivated Conlon, Fox and Sudakov [44, Problem 3.5] to ask if one could show that, for fixed $r \in \mathbb{N}$,

$$\mathrm{r}_{\mathrm{ind}}(H;r) \leqslant 2^{k^{1+o(1)}}.$$

Our solution to the conjecture of Erdős, presented here in Chapter 2, works directly in this more general setting.

Theorem 1.2. *There exists a constant $C > 0$ such that*

$$\mathrm{r}_{\mathrm{ind}}(H;r) \leqslant r^{Crk} \tag{1.10}$$

for every $r \geqslant 2$ and every graph H with k vertices.

We remark that, up to the value of the constant C, the bound (1.10) matches (1.3), the classical upper bound of Erdős and Szekeres [54] on the r-colour Ramsey numbers $\mathrm{R}(k;r)$, which was only recently improved by a small exponential factor in [11] (see (1.4)). Our method will moreover imply the stronger statement that there exists a graph G with $N = r^{Crk}$ vertices such that every r-colouring of the edges of G contains an induced monochromatic copy of *every* graph H on k vertices. In fact, we will show that almost every graph G with N vertices has this property.

A key component of our approach is a novel strengthening of the method of hypergraph containers. This method, which was introduced in 2015 by Balogh, Morris and Samotij [13] and Saxton and Thomason [117], is a flexible technique for controlling the probability that a random set avoids some forbidden substructure (see the surveys [14, 114]). Roughly speaking, the basic container lemma implies that the sets that avoid these substructures are "clustered", in the sense that they can be covered by a relatively small number of sets that contain only a few copies of the forbidden substructures.

The container method has been used by several different sets of authors to prove Ramsey theoretic properties of random graphs. For example, it was used by Nenadov and Steger [100] to give a simpler proof of the remarkable random Ramsey theorem of Rödl and Ruciński [109]. This result determines the threshold $p = p(n)$ for which every r-colouring of the edges of $\mathbb{G}(n,p)$ contains a monochromatic copy of any fixed graph H (save a few degenerate exceptions) with high probability, which we abbreviate simply to $\mathbb{G}(n,p) \overset{}{\underset{r}{\to}} H$.

Later, Mousset, Nenadov and Samotij [97] built on the techniques in [100] to obtain the threshold for the 1-statement of the asymmetric variant $\mathbb{G}(n,p) \to (H_1, H_2)$ for every fixed pair of graphs H_1 and H_2, and the matching threshold for the 0-statement was recently determ-

ined in a series of works [24, 90, 36] using different techniques, thus settling what was known as the Kohayakawa–Kreuter conjecture. The container method is also an important part of the argument that establishes that the thresholds for these Ramsey properties are moreover sharp [62].

More relevant to this chapter are the works of Conlon, Dellamonica, La Fleur, Rödl and Schacht [40] and Balogh and Samotij [15], where they used the method of hypergraph containers to bound the induced Ramsey numbers of graphs and hypergraphs. In particular, the authors of [40] significantly improved the best-known upper bound for the induced Ramsey number of an arbitrary s-uniform hypergraph with k vertices when $s \geqslant 3$. Their result was then improved when $s = 2$ (that is, for graphs) in [15], where the authors introduced a new container lemma with a dramatically improved dependency on the size of the forbidden structure, and applied it to give a bound of the form

$$\mathrm{r}_{\mathrm{ind}}(H; r) \leqslant r^{O(rk^2)}. \tag{1.11}$$

However, further improvements to the dependency of the container lemma would not advance the bound beyond (1.11), and we must therefore take a different approach.

Our extension of the container method shows that not only sets that avoid these substructures are clustered, but that such a phenomenon is also observed for sets in which these substructures are not "well-distributed," for a precise definition of this notion. Even though we believe that this definition and the strengthening of the method of hypergraph containers are both of independent interest, we postpone their statements to Chapter 2, because they are technical and would be hard to motivate here.

1.2 Additive combinatorics

A strengthening of van der Waerden's theorem, due to Szemerédi, is perhaps the most celebrated result in additive combinatorics. It states that every subset of $\mathbb{N}$ with positive upper density contains a k-AP. Despite the simplicity of its statement, the various different proofs of this theorem have led to many important developments in modern mathematics: the regularity method in graph theory [123], the correspondence principle of Furstenberg [63, 64] in ergodic theory, and the higher-order Fourier analysis of Gowers [67].

Another seminal result in additive combinatorics was originally proven by Freĭman [60] and much later, using different methods that proved very influential, by Ruzsa [111]. Recall that, for two sets A and B in an Abelian group, their sumset is defined as

$$A + B = \{a + b : a \in A,\, b \in B\}.$$

It is easy to see that when $A \subseteq \mathbb{Z}$ is an arithmetic progression, we have $|A + A| = 2|A| - 1$, and that this is the minimum possible size of a sumset. Taking the sumset of APs with distinct common differences leads to the definition of a generalised arithmetic progression (GAP), a set of the form

$$P = \left\{ a_0 + \sum_{i=1}^{d} n_i a_i : 0 \leqslant n_i < \ell_i \text{ for all } i \in [d] \right\},$$

where d is called the dimension of the GAP, $a_0 \in \mathbb{Z}$ is the starting point of the GAP, the integers $a_1, \dots, a_d$ are called the differences, and $\ell_1, \dots, \ell_d$ are called the side-lengths of the GAP.

In particular, if $P \subseteq \mathbb{Z}$ is a d-dimensional GAP, then $|P + P| \leqslant 2^d|P|$, and hence every A which is a dense subset of a d-dimensional GAP, with constant d, has constant *doubling* $\sigma[A] := |A+A|/|A|$. The Freĭman–Ruzsa theorem gives a complete characterisation of sets with bounded doubling in essentially the same terms: if $A \subseteq \mathbb{Z}$ has constant doubling σ, then there exists a d-dimensional GAP $P \subseteq \mathbb{Z}$ such that $A \subseteq P$, $|P| \leqslant C|A|$ and $d \leqslant C$ for a constant $C > 0$ that depends only on σ.

In this thesis, the first object in additive combinatorics that we study, in Chapter 3, is the Cayley *sum* graph, a variant of Cayley graphs that is related to sumsets because a subset A of its vertices is independent exactly when $A \hat{+} A \subseteq S^{\mathsf{c}}$, where

$$A \hat{+} A = \{a_1 + a_2 : a_1, a_2 \in A, a_1 \neq a_2\}$$

and S is the generator of the graph. We introduce this subject and the main result of chapter, originally obtained in joint work with Campos and Marciano [29], in the next section.

In Chapter 4, we study the length of the longest AP inside a sumset $A+B$ when $A, B \subseteq [n]_p$, where $[n]_p$ denotes a p-random subset of $[n]$, i.e. a subset in which each element of $[n]$ is included independently and with probability p, and both A and B are not too sparse. Before that, we connect this problem with previous work, primarily the deterministic precursor of our result, and with other random analogues of classical results in additive combinatorics, in Section 1.2.2.

1.2.1 Independence number of random Cayley graphs

For a finite Abelian group Γ, we define the Cayley sum graph G_S of a set $S \subseteq \Gamma$ to have Γ as its vertices and an edge between two distinct $x, y \in \Gamma$ if and only if $x + y \in S$. Green [72, Question 2] asked the following general question: to what extent does a random Cayley sum graph G_S, obtained via a uniform random choice of $S \subseteq \Gamma$, simulate a random graph on $n = |\Gamma|$ vertices?

Focusing on the group $\mathbb{Z}_n$, for a large prime $n \in \mathbb{N}$, and denoting by G_p the random Cayley sum graph G_S when S is a p-random subset of $\mathbb{Z}_n$, Green [72] showed that with high probability the size of the largest clique in $G_{1/2}$ is the same as that of $\mathbb{G}(n, 1/2)$ up to a constant factor. Since a 1/2-random subset is distributed as its complement, [72] establishes that both the clique and independence number of $G_{1/2}$ are within a constant factor of those of $\mathbb{G}(n, 1/2)$, and therefore the random Cayley sum graph in $\mathbb{Z}_n$ yields another random construction for an exponential (with a smaller base in the exponent) lower bound for $\mathrm{R}(k)$. This improved on previous work of Alon and Orlitsky [6], who, motivated by questions in information theory, proved the existence of Cayley graphs that are within one logarithmic factor of $\mathrm{R}(k)$.

Interestingly, Green's result is an almost direct consequence of theorems counting the number of sets $A \subseteq \mathbb{Z}_n$ such that the size of A is given and $|A + A| \leqslant m$. To obtain this counting result, he refined a previous theorem of Konyagin and Lev [89], which shows that equivalence classes of Freĭman isomorphisms (see Chapter 3 for a definition) can be efficiently counted using linear algebra. Green deduces his bound on the typical size of $\alpha(G_{1/2})$ by taking a union bound over all candidate sets, thus bounding the probability that any of them is independent in $G_{1/2}$.

Later, Green and Morris [73] improved the bounds on both the clique number (and, consequently, also of the independence number) of $G_{1/2}$ to show that, in fact,

$$\alpha(G_{1/2}) = \big(1+o(1)\big)\alpha\big(\mathbb{G}(n,1/2)\big) \tag{1.12}$$

with high probability, which in particular gives an alternative proof of $\mathrm{R}(k) \geqslant 2^{(1-o(1))k/2}$, because, recall, $o(1)$ denotes a function $f(n)$ that satisfies $f(n) \to 0$ as $n \to \infty$. They obtained these results by sharpening Green's count of sets A with given doubling $\sigma[A] = |A+A|/|A|$. When $\sigma[A] = O(1)$ they used an arithmetic regularity lemma to obtain a tighter count on the number of those sets, and, when $\sigma[A] \geqslant c|A|$ for a constant $c > 0$, or equivalently $\sigma[A] \gtrsim |A|$ (a notation that we will use throughout), they refined Green's original estimate by leveraging the isoperimetric inequality on $\mathbb{Z}^d$ and an elegant graph theoretic argument.

In Chapter 3, we present the first extension of (1.12), the bound of Green and Morris on $\alpha\big(\mathbb{G}(n,p)\big)$, to the more challenging setting when $p = o(1)$. That is, Theorem 1.3 is the first asymptotically sharp bound on the independence number of G_p for any group in the sparse setting.

Theorem 1.3. *Let n be a prime number and let $p = p(n)$ be such that $(\log n)^{-1/80} \leqslant p \leqslant 1/2$. The random Cayley sum graph G_p of $\mathbb{Z}_n$ satisfies*

$$\alpha(G_p) = \big(2+o(1)\big)\log_{\frac{1}{1-p}} n \tag{1.13}$$

with high probability as $n \to \infty$.

If $p = o(1)$, then this bound is asymptotically equal to $(2+o(1))(\log n)/p$; it also matches $\alpha\big(\mathbb{G}(n,p)\big)$ in the corresponding range of p [19, 79]. The lower bound in (1.13) follows from a second moment computation together with a simple combinatorial argument (see Section 3.9 in Chapter 3). The proof of the upper bound, on the other hand, is much more challenging: we take a union bound over sets A that are candidates to be independent in G_p when $|A|^c < \sigma[A]$ for a small constant $c > 0$, relying on a sufficiently precise count of such sets (that crucially relies on their additive structure), but such an approach is not sufficient when $\sigma[A] \leqslant |A|^c$.

Indeed, if we try to take a union bound over the sets in $\mathcal{A}' = \{A \subseteq \mathbb{Z}_n : |A| = k,\ \sigma[A] \leqslant |A|^c\}$,

$$\mathbb{P}(\exists\, A \in \mathcal{A}' : A \hat{+} A \subseteq S^{\mathrm{c}}) \leqslant \sum_{A\in\mathcal{A}'} (1-p)^{|A\hat{+}A|}, \tag{1.14}$$

we are faced with the following problem: for every $m \in \{2k, \dots, k^{1+c}\}$, every k-subset A of an interval of length $\lfloor m/2 \rfloor$ satisfies $\sigma[A] \leqslant m/k$, and there are at least $\binom{\lfloor m/2 \rfloor}{k}$ such sets. It follows that the right-hand side of (1.14) is at least

$$\sum_{m=2k}^{k^{1+c}} \binom{\lfloor m/2 \rfloor}{k} (1-p)^m \geqslant \binom{2k}{k}(1-p)^{4k} \to \infty,$$

as $n \to \infty$, whenever $p = o(1)$ and $k \to \infty$. The conclusion is that any approach that uses the union bound over all sets with small doubling cannot give the optimal upper bound on the independence number for p smaller than some explicit constant, like $1/5$.

To overcome this obstacle, we show that each of those sets contains a much smaller subset F – we call it a "fingerprint" of A – so that, after determining that F possesses some special properties, it suffices to count the fingerprints to deduce Theorem 1.3. Even though this idea is inspired by the application of the asymmetric container method of Morris, Samotij and Saxton [96] to the problem of counting sets with small sumset by Campos [26][2], we emphasize that this fingerprint strategy is not a novelty of our work. For instance, Green [72] used a related approach to handle sets A with $\sigma[A] \leqslant 3$, and this was the main tool in the proof of the upper bound $\alpha(G_p) \lesssim p^{-2}(\log n)^2$ given by Alon [5] for arbitrary Abelian groups Γ of order n. What is genuinely new about the fingerprints in Chapter 3 is using this strategy to obtain a sharp bound on the typical independence number of random Cayley sum graphs.

The special property that we require of each fingerprint $F \subseteq A$ is having a sufficiently large sumset. Ideally, we would like $F + F$ to be as large as $A + A$; this obviously imposes a lower bound of $|F| \geqslant |A + A|^{1/2}$. Bollobás, Leader and Tiba studied the question of finding such sets for general Abelian groups before our work: they obtain $F \subseteq A$ satisfying $|F + F| \approx |A + A|$ when A has bounded doubling [21, Theorem 2]. However, their result does not serve our needs because it does not handle the case of $\sigma[A]$ being $|A|^c$ for fixed, small $c > 0$.

We pursue a different approach to obtain our collection of fingerprints, one that is able to handle sets with doubling that is polynomial in their size. This approach yields fingerprints such that

$$|F| \approx |A + A|^{1/2} \log |A|$$

which is only a logarithmic factor away from any bound that can achieve $|F + F| \approx |A + A|$, simply because $|F + F| \leqslant |F|^2$. We will not, however, be able to ensure that $|F + F| \approx |A + A|$ in complete generality, instead obtaining that the size of the fingerprint sumset $F + F$ is as large as what is ensured by Freĭman's lemma for $|A + A|$, which is just enough to obtain the sharp bound on $\alpha(G_p)$.

The key step in the proof of Theorem 3.1, where we determine the existence of such fingerprints, is a result stating that we can find some small $T \subseteq A$ such that $A + T$ almost attains the bound in Freĭman's lemma. Proving this theorem is the most difficult part of our proof, and we consider it to be of independent interest. We will write $\operatorname{rank}(A)$ for the minimum dimension of an affine subspace that contains a set $A \subseteq \mathbb{R}^d$.

Theorem 1.4. *Let $d, r \in \mathbb{N}$ and $\gamma > r^{-1/3}$. If $A \subseteq \mathbb{R}^d$ is a finite set with $\operatorname{rank}(A) \geqslant r$, then there exists $T \subseteq A$ such that $|T| \leqslant 4(r+1)/\gamma$ and*

$$|A + T| \geqslant (1 - 5\gamma)\frac{(r+1)|A|}{2}.$$

The proof of this theorem is technical; however, we will also show that if we were satisfied with $1/6$ being the leading constant instead of $(1 - 5\gamma)/2$, then a much simpler approach would suffice. This weaker version would also be enough to prove the upper bound in Theorem 1.3 with 6 being the leading constant instead of 2.

[2]Indeed, we could also use asymmetric containers as a tool to prove Theorem 1.3, but we prefer this simpler approach as it results in better bounds and a more self-contained proof.

1.2.2 Additive structure in subsets of sparse random sets

One equivalent way to state Szemerédi's theorem from 1975 is that for all $k \in \mathbb{N}$ and $\delta > 0$, if $n \geqslant n_0(\delta, k)$, then every $A \subseteq [n]$ with $|A| \geqslant \delta n$ contains a k-AP. The best quantitative dependency of n_0 on δ and k is due to Gowers [67], who showed that

$$\delta \geqslant \frac{1}{(\log\log n)^{2^{-2^{k+9}}}}$$

suffices. Recent work of Leng, Sah and Sawhney [91] improved this dependency in the regime where k is fixed and n is sufficiently large to

$$\delta \geqslant \frac{1}{\exp((\log\log n)^{c_k})}$$

for some constant $c_k > 0$.

A classical result of Green [71], improving on previous work of Bourgain [23], shows that much longer APs can be found if one looks into $A + A$ instead of directly in A. Concretely, he proved that if $A, B \subseteq [n]$ are such that $|A| \geqslant \alpha n$ and $|B| \geqslant \beta n$, then the sumset $A + B$ contains an AP of length at least

$$\exp(c(\alpha\beta \log n)^{1/2} - \log\log n) \tag{1.15}$$

for some absolute constant $c > 0$. A construction due to Ruzsa [112] shows that the $1/2$ in the exponent of (1.15) cannot be improved to any constant larger than $2/3$.

Another direction is to ask for APs when A has density δ inside $[n]_p$, a p-random subset of $[n]$, rather than inside $[n]$, and to determine the threshold in $p = p(n)$ that ensures such sets A still have the property proved by Szemerédi. This line of work was initiated by Kohayakawa, Łuczak and Rödl [87], who first proved that Roth's theorem, the case 3-APs of Szemerédi's theorem, holds in the random setting whenever $p \gg n^{-1/2}$. The threshold for all k was determined only later, in independent work of Conlon and Gowers [38] and Schacht [118], both of which established this result as a consequence of more general statements about the transference of extremal results in dense sets to sparse random sets.

Hamel and Łaba [75] considered the combination of these two lines of work, which is to take A and B to be subsets of $[n]_p$, instead of $[n]$, and to look for the longest AP in $A + B$. To state their result, let $L(S)$ denote the length of the longest AP contained in a set S.

Theorem 1.5 ([75, Theorem 1.4])**.** *There exists a constant $C > 0$ such that for every*

$$p = p(n) \geqslant n^{-1/140} \qquad \text{and} \qquad \frac{C\log\log n}{\sqrt{\log n}} \leqslant \alpha \leqslant 1,$$

the following holds with high probability. Every subset $A \subseteq [n]_p$ with $|A| \geqslant \alpha\big|[n]_p\big|$ satisfies

$$L(A + A) \geqslant \exp\left(\frac{\alpha^2 \log\log n}{C\log(1/\alpha)\big(\log\log\log n + \log(1/\alpha)\big)}\right). \tag{1.16}$$

Very recently, Alon and Pham [7] considered the above problem when $\alpha = 1$, which is equivalent to replacing the original dense sets A and B by random sets. Specifically, they

studied the behaviour of $L(A+A)$ when A is a p-random subset of $[n]$ for various functions $p = p(n)$. Miyazaki [95] previously obtained tight bounds for the length of the longest AP in many regimes of p using the Park–Pham theorem (previously known as the Kahn–Kalai conjecture) [105], but the typical value of this function when $p = n^{-1/2}$ was determined only by Alon and Pham, using Talagrand's inequality.

Returning to Theorem 1.5, there are two ways in which one could hope to improve upon this result. The first goal would be to weaken the hypothesis on p. Hamel and Łaba themselves remarked that the "natural threshold [for the power of $1/n$] would be $1/2$" but that determining its precise value seemed to require other types of arguments [75].

We could also hope to show that, under the same assumptions of Theorem 1.5, there is an AP in $A+A$ whose length is greater than what is given by (1.16). A natural goal here would be to match (1.15), the bound due to Green [71] for the case in which $p = 1$.

In Chapter 4, we present joint work with Campos and Kohayakawa [28] that accomplishes these two goals. That is, for every $\varepsilon > 0$, we show that $A + B$ contains APs essentially as long as (1.15) for all $p \geqslant n^{-1/2+\varepsilon}$.

Theorem 1.6. *There exist constants $c > 0$ and $C > 0$ such that, for any $0 < \beta \leqslant \alpha \leqslant 1$, $k \in \mathbb{N}$ and $p = p(n)$ satisfying*

$$k \leqslant \exp\left(c(\alpha\beta \log n)^{1/2} - \log\log n\right)$$

and

$$p \geqslant C\,\frac{k}{\beta}\left(\frac{(\log n)^3}{n}\right)^{1/2},$$

the following holds with high probability for $S \sim [n]_p$. Every pair of subsets $A, B \subseteq S$ with

$$|A| \geqslant \alpha|S| \qquad \text{and} \qquad |B| \geqslant \beta|S|$$

satisfies

$$L(A+B) \geqslant k.$$

While the approach of Hamel and Łaba to this problem is mainly based on analytical techniques, our proof is essentially combinatorial, with just a small amount of (elementary) Fourier analysis.

To prove Theorem 1.6, we use the method of hypergraph containers [13, 117], which is known to be useful for related problems. For instance, one can easily deduce the aforementioned random version of Szemerédi's theorem from its standard version using the container method.

Another example is the work of Nguyen [102], who proved an essentially optimal random analogue of the Furstenberg–Sárközy theorem regarding sets with no square difference, a problem that was also first considered by Hamel and Łaba. The main combinatorial lemma in Nguyen's work is a probabilistic version of the graph container method [114], which was later generalized to obtain the method of hypergraph containers.

The other technical requirement in the proof of Theorem 1.6 is the almost periodicity results of Croot, Łaba and Sisask [45]. One of the applications of their method is to give an alternative proof of Green's result about the length of the longest arithmetic progression in $A + B$ that requires only elementary Fourier analysis.

Chapter 2

An exponential upper bound for induced Ramsey numbers

The goal of this chapter, whose contents are adapted from joint work with Aragão, Campos, Filipe and Marciano [9], is to prove the following theorem.

Theorem 1.2. *There exists a constant $C > 0$ such that*

$$\mathrm{r}_{\mathrm{ind}}(H; r) \leqslant r^{Crk}$$

for every $r \geqslant 2$ and every graph H with k vertices.

Unlike the earlier approaches in [86, 43, 57], where the authors developed ingenious deterministic algorithms to embed H in a pseudorandom host graph G, we will instead use a relatively simple vertex-by-vertex embedding strategy in a truly random graph G. Specifically, we consider the Erdős–Rényi random graph $\mathbb{G}(N, 1/2)$, where each edge of K_N is included independently at random with probability $1/2$; equivalently, we can choose a (labelled) graph G on N vertices uniformly at random.

Our strategy will crucially exploit the fact that the edges of $\mathbb{G}(N, 1/2)$ are chosen randomly, rather than relying on pseudorandom properties that are also satisfied by $G \sim \mathbb{G}(N, 1/2)$. We will show that, with (extremely) high probability, every r-colouring of the edges of $G \sim \mathbb{G}(N, 1/2)$ contains an induced monochromatic copy of an arbitrary k-vertex graph H.

One way to approach the problem is via an Erdős–Szekeres-type induction. In other words, we might generalise to the setting in which we want to find an induced copy of H_i in colour i, apply the induction hypothesis inside a subset $U \subseteq V(G)$ of size δN to find an induced copy of H_i^- (that is, H_i minus a vertex) in colour i for some $i \in [r]$, and then attempt to use the randomness between U and the rest of the vertices to extend this copy of H_i^- to a copy of H_i.

There are several major obstacles to utilizing such a strategy. First, and most obviously, there may be no edges of colour i between U and $V(G) \setminus U$, in which case we have no chance of extending H_i^- to a copy of H_i in colour i. We can easily deal with this issue, however, by instead using the induction hypothesis to find an induced copy of $H_i^- \subseteq G[U]$ in colour i for *every* $i \in [r]$, and then considering the colour that is used most often between U and $V(G) \setminus U$.

A second (and more challenging) issue is that the colouring of the edges inside the set U

is allowed to depend on the (random) edges between U and $V(G) \setminus U$. We will deal with this problem by taking a union bound over all possible colourings of the edges of $G[U]$. However, this does not come for free: it requires us to prove an extremely strong bound on the probability of failure in each step of the induction.

In order to prove such a strong bound, we must strengthen our induction hypothesis. Indeed, if we only have one copy of H_i^-, the probability that no vertex of $V(G) \setminus U$ extends this fixed copy of H_i^- to an induced copy of H_i in G is essentially $(1-2^{-k})^N$, which is much too large to beat the roughly $r^{\delta^2 N^2}$ choices in our union bound. In order to reduce this failure probability, we will need to instead find many copies of H_i^- in $G[U]$ that are moreover "well-distributed" in a certain precise sense, which we define in Section 2.2 (see Definition 2.2). We say that a hypergraph is (p,σ)-*Janson* if its edges are well-distributed in this sense, since this definition resembles (and was inspired by) the condition in Janson's inequality.

There is still, however, one further (and even more critical) obstruction to this approach: we must also handle all colourings of the edges between U and $V(G) \setminus U$, and in this case we cannot do so simply by taking a union bound, since there are too many possible colourings. In order to deal with this more serious obstacle, we will use the method of hypergraph containers.

Our first significant departure from the earlier applications of the container method for Ramsey problems (cf. [15, 40, 97, 100]) is in how we apply this method. Previously, the authors embedded the whole of H in a single step, encoding the edge sets of induced copies of H using a hypergraph whose vertex set consists of r copies of $E(K_n)$. In contrast, we will have not just one, but roughly $r^{|U|^2}$ different hypergraphs, one for each of the possible colouring of $G[U]$. Our hypergraphs will encode the vertex sets, rather than the edge sets, of monochromatic induced copies of H_i^- in $G[U]$, and our aim will then be to find a (p,σ)-Janson collection of monochromatic induced copies of H_i in an arbitrary r-colouring of the edges between U and $V(G) \setminus U$.

Although our final goal is to find a (p,σ)-Janson collection of copies of H_i, applying the method of containers to find a single copy is already instructive. This will require introducing a correspondence between independent sets in our hypergraph and pairs of neighbourhoods $\mathrm{N}_G(v)$ and $\mathrm{N}_{G_i}(v)$ (where G_i is the subgraph of i-coloured edges) that do not extend any copy of H_i^-. In this simplified setting, applying a standard hypergraph container lemma is sufficient to conclude that the probability that a vertex does not extend any copies of H_i^- is exponentially small.

Changing from extending a single copy to finding a (p,σ)-Janson collection of copies of H_i introduces significant new difficulties, which further set our argument apart from those in [15, 40, 97, 100]. The general strategy will be to have a second induction on the Janson parameter σ: we will show that adding a vertex to U increases σ with very high probability, which results in a "richer" family of copies of H_i. In this way, after adding δN vertices, σ will be as large as we need. However, even the first step of this incrementing argument requires containers for sets where the hypergraph is not (p,σ)-Janson, rather than the classical container lemma from [13, 117] or the "efficient" container lemma proved in [15], see Theorem 2.21. For this purpose, our main tool will be a strengthening of the container method, proved recently by Campos and Samotij [34].

One of the main challenges of carrying out the above outline is that our induction hypothesis is a *global property* (being (p,σ)-Janson) of the hypergraph that encodes monochromatic induced copies of H_i in colour i, whereas previous container lemmas are only able to handle *local properties* of the forbidden substructures.

Our most important technical contribution is a novel method that allows one to prove container theorems that can handle such global properties that depend on local behaviour. The first step of this method is to represent the global property as edges, of potentially linear size, in a hypergraph. We then decompose the independent sets of this hypergraph into independent sets of a small collection of *container hypergraphs*. Together, the container hypergraphs encode all of the local obstructions to this global property, which effectively separates the local obstructions from the global obstructions. This separation leads to a more favourable setting: to deal with local obstructions, we have a vast array of tools at our disposal, including traditional container theorems.

We illustrate this method by proving a general container theorem for vertex-induced subhypergraphs that are not (p,σ)-Janson, Theorem 2.21, which we moreover expect to have further applications. The proof of this result starts with an application of the aforementioned result of Campos and Samotij [34] (see Theorem 2.23), yielding a decomposition of the sets that are not (p,σ)-Janson into a collection of independent sets in container hypergraphs. Roughly speaking, Theorem 2.23 states that each container hypergraph that does not have a "large local part" has the following property: a random independent set I with $q|V|$ vertices looks very similar to a binomial random subset V_q of the vertices of the hypergraph. More precisely, all sets L that are not edges of the container hypergraph satisfy

$$\mathbb{P}\big(L \subseteq I\big) \approx \mathbb{P}\big(L \subseteq V_q\big). \tag{2.1}$$

From (2.1), we will be able to deduce that the random independent set I is (p,σ)-Janson with positive probability, contradicting the fact that independent sets are not (p,σ)-Janson. It then follows that all container hypergraphs have a large local part, and we can apply another container theorem (Theorem 2.15, also proved in [34]) to the local part of each container hypergraph to finish the proof.

2.1 Organization of the chapter

The next section provides key definitions for the rest of the chapter and an important reduction. In particular, we define two events, one encoding the induction hypothesis and the other encoding the failure to contain a "well-distributed" collection of induced copies of H_i. We then state our key probabilistic lemma, Lemma 2.8, which says that the probability of these two events happening simultaneously is very small, and prove that Theorem 1.2 follows from it by induction. The bulk of this chapter is then concerned with proving Lemma 2.8, first proving intermediate results that we require.

We will have two expository sections that contain many of the ideas that we use later, but in simpler settings. In Section 2.3, this simpler setting corresponds to showing that the probability that a vertex extends no copy of H_i^- in a fixed colour $i \in [r]$ is exponentially small. This section

shows how to map neighbourhoods that fail to extend a copy of H_i^- to independent sets in a certain hypergraph $\mathcal{H}$, which will illustrate how the method of hypergraph containers can be applied to this problem. It is also where we explain how our induction hypothesis, the event $\mathcal{E}(\mathbf{s})$ defined in Definition 2.6 of Section 2.2, yields a supersaturation result for our application(s) of containers.

In Section 2.4, we present the proof of a container theorem for "sets that are not well-distributed" (see Theorem 2.21, and also Theorem 2.26, which implies it). While this result is not sufficiently strong to prove Lemma 2.8, its proof already presents the general argument that we will later adapt to prove Theorem 2.35, the container theorem that we do use in the proof of Lemma 2.8. We see that argument as a general method to reduce a problem about sets avoiding a global property to a simpler problem, where the sets only need to avoid a local property. In this particular instance, we deal with this local property by using a standard hypergraph container lemma (Theorem 2.15).

Section 2.5 extends the proof in Section 2.3 to replace the use of containers for independent sets by the statement proved in Section 2.4. This yields a statement saying that one step of our inductive embedding procedure succeeds with very high probability. In fact, we will use Theorem 2.35, the container theorem tailored to our application, but we postpone its proof to the last section of the chapter, since it is the most technical part of the entire argument.

In Section 2.6, we prove Lemma 2.8 in two stages. The first and easier stage uses a double counting argument to deduce that if every sufficiently large subhypergraph of a hypergraph is well-distributed, then the entire hypergraph is also well-distributed, albeit with slightly changed parameters. The second stage proves Lemma 2.8 using the inductive argument.

The last section, Section 2.7, contains the proof of Theorem 2.35. As mentioned above, this proof follows the blueprint of Section 2.4, but with the additional details required to prove the main result of Section 2.5. The complications make the proof in Section 2.7 more technical, but it turns out the approach we developed in Section 2.4 is sufficiently flexible to deal with them.

2.2 Reduction to a key lemma

This section is devoted to reducing Theorem 1.2 to a key probabilistic lemma. To that end, we first introduce two important definitions, Definitions 2.1 and 2.2. The first will allow us to reason about induced copies of a graph H as edges in a hypergraph, while the other is the key notion of how "well-distributed" the edges of a hypergraph are. This notion will underpin our container theorem, so immediately afterwards, we note a few simple properties of this latter notion that will be useful later.

Using these two definitions, we will further define an event to represent our inductive assumption, Definition 2.6, and another to stand for the failure of completing an induction step, Definition 2.7. We follow this with the statement of the key probabilistic lemma, Lemma 2.8, which bounds the probability of these events happening simultaneously, and a proof that Theorem 1.2 can be deduced from Lemma 2.8. This section concludes with a summary for the rest of the chapter, where we provide the tools that we will use to prove the key lemma in Section 2.6.

2.2.1 Preliminaries

The first definition that we need is that of the hypergraph which encodes copies of some graph $F \subseteq G'$ that is also induced in G. We will use Definition 2.1 with G' being the subgraph defined by the edges in some colour $i \in [r]$, and G being the underlying (random) graph. In what follows and in the rest of this chapter, we will identify a hypergraph with its edge set.

Definition 2.1. Given graphs F, G' and G such that $G' \subseteq G$, define $\mathfrak{I}_{F,G',G}$ to be the hypergraph with $v(F)$ uniformity, vertex set $V(G)$, and

$$\mathfrak{I}_{F,G',G} = \big\{L \subseteq V(G) : F \cong G'[L] = G[L]\big\}.$$

That is, each hyperedge of $\mathfrak{I}_{F,G',G}$ corresponds to a copy of $F \subseteq G'$ that is induced in G.

As we will apply the container method to a hypergraph that is closely related to $\mathfrak{I}_{F,G',G}$, we emphasise that it has $V(G)$ for its vertex set. This is in contrast with using $E(G)$ for the vertices of the auxiliary hypergraph, the more common definition in applications of the container method.

We now formalise what it means for a hypergraph to be "well-distributed." A crucial aspect of this definition is that we let ourselves assign weights to the edges of the hypergraph using measures. We say that a measure ν is *supported* on a hypergraph $\mathcal{G}$ if ν is non-zero only on the edges of $\mathcal{G}$. A hypergraph $\mathcal{G}$ is (p,σ)-Janson when there exists a measure ν with certain properties that is supported on $\mathcal{G}$.

We call this property (p,σ)-Janson because, under some extra assumptions, one can apply Janson's inequality and conclude that a p-random subset of the hypergraph's vertices is an independent set with probability at most $\exp(-\sigma)$. Even though this is the original motivation for the definition, we will not require these extra assumptions or use Janson's inequality.

Definition 2.2. Let $\mathcal{G}$ be a hypergraph and $p > 0$. For measures $\nu : \mathcal{G} \to \mathbb{R}_{\geqslant 0}$, we define

$$e(\nu) = \sum_{E \in \mathcal{G}} \nu(E) \qquad \text{and} \qquad \Lambda_p(\nu) = \sum_{\substack{L \subseteq V(\mathcal{G}) \\ |L| \geqslant 2}} d_\nu(L)^2\, p^{-|L|}, \tag{2.2}$$

where

$$d_\nu(L) = \sum_{L \subseteq E \in \mathcal{G}} \nu(E)$$

is the degree of the set L in the measure ν. For every $\sigma > 0$, we say that $\mathcal{G}$ is (p,σ)-Janson if there exists a measure $\nu : \mathcal{G} \to \mathbb{R}_{\geqslant 0}$ such that

$$\Lambda_p(\nu) < \frac{e(\nu)^2}{\sigma}. \tag{2.3}$$

If $\sigma = 0$, then every hypergraph is (p,σ)-Janson.

We mention three very simple, but useful observations about this property, for future reference. The first is so simple that we omit its proof.

Observation 2.3. *If $\mathcal{G}$ is a (p,σ)-Janson hypergraph for $p>0$ and $\sigma \geqslant 0$, then for all $p' \geqslant p$ and $\sigma' \leqslant \sigma$, it is also (p',σ')-Janson.*

The next observation is that we can normalise any measure ν satisfying (2.3) to choose the value it takes in $e(\nu)$.

Observation 2.4. *Let $p,\sigma>0$, and let $\mathcal{G}$ be a hypergraph that is (p,σ)-Janson. For every $y>0$, there exists $\nu : \mathcal{G} \to \mathbb{R}_{\geqslant 0}$ such that*

$$e(\nu) = y \qquad \text{and} \qquad \Lambda_p(\nu) < \frac{y^2}{\sigma}.$$

Proof. As $\mathcal{G}$ is (p,σ)-Janson, there exists $\tilde{\nu} : \mathcal{G} \to \mathbb{R}_{\geqslant 0}$ such that $\Lambda_p(\tilde{\nu}) < e(\tilde{\nu})^2/\sigma$, which implies that $e(\tilde{\nu})>0$ by the non-negativity of Λ_p. One can now verify that taking $\nu = y\tilde{\nu}/e(\tilde{\nu})$ completes the proof. □

Finally, we observe being (p,σ)-Janson is monotone with respect to taking subhypergraphs.

Observation 2.5. *Let $\sigma \geqslant 0$ and $p>0$, and let $\mathcal{G}$ and $\mathcal{G}'$ be hypergraphs satisfying $\mathcal{G}' \subseteq \mathcal{G}$. If $\mathcal{G}'$ is (p,σ)-Janson, then $\mathcal{G}$ is also (p,σ)-Janson.*

Proof. The observation follows immediately from the fact that any ν supported on $\mathcal{G}'$ is also supported on $\mathcal{G}$, by assigning measure zero to every $E \in \mathcal{G} \setminus \mathcal{G}'$. □

It is important to remark that the notion of (p,σ)-Janson appeared implicitly in previous work of Saxton and Thomason [117] and Balogh and Samotij [15]. Indeed, the main condition for the efficient container lemma of [15] is equivalent to the hypergraph $\mathcal{H}$ in that statement being (p,σ)-Janson for $\sigma = C_k pv(\mathcal{H})$, where C_k depends only on the uniformity of $\mathcal{H}$.

Being (p,σ)-Janson is also equivalent to other notions of pseudorandomness for hypergraphs, notably the p-spread condition, which holds for a hypergraph $\mathcal{H}$ if, for every $L \subseteq V(\mathcal{H})$,

$$d(L) \leqslant p^{-|L|} e(\mathcal{H}) \tag{2.4}$$

where $d(L) = \{A \in \mathcal{H} : L \subseteq A\}$. This definition is used extensively in the literature, including in the recent advances in the sunflower conjecture [8] and the Park–Pham theorem (previously the Kahn–Kalai conjecture) [105].

To connect this notion to Definition 2.2, we will show, in the proof of Theorem 2.15, that if a hypergraph $\mathcal{G}$ is $(p, p\,v(\mathcal{G}))$-Janson, then there is no hypergraph $\mathcal{C}$ such that

$$\sum_{A\in\mathcal{C}} p^{|A|} \leqslant pv(\mathcal{G}) \qquad \text{and} \qquad \mathcal{G} \subseteq \langle \mathcal{C} \rangle,$$

where

$$\langle \mathcal{C} \rangle = \{L \subseteq V(\mathcal{C}) : \exists A \in \mathcal{C}, A \subseteq L\}.$$

On the other hand, the existence of such a hypergraph $\mathcal{C}$, called a p-cheap cover for $\mathcal{G}$, certifies, via LP duality, that $\mathcal{G}$ is not p-spread.

In view of these (near-)equivalences, we mention that the advantage of using the (p,σ)-Janson definition in our proof is that it is expressed through the quadratic quantity $\Lambda_p(\nu)$,

rather than through the collection of inequalities required to express p-spreadness, cf. (2.4). This clean algebraic expression makes it more amenable to conditional expectation arguments, such as those used in the proof of Theorem 2.26; this is the only reason for our use of the (p,σ)-Janson definition rather than p-spreadness.

2.2.2 The key lemma

As outlined in the beginning of the chapter, our proof will be by induction on k, and we will need a very strong bound on probability of each inductive step failing, in order to allow for a union bound over colourings of $G[U]$ for some set U of size δN. We will now define the event representing our induction hypothesis, Definition 2.6, and the event that represents the failure of completing an induction step, Definition 2.7. Once these are defined, we will state the key lemma, Lemma 2.8, and prove that our main theorem is a straightforward consequence of it.

To state our induction hypothesis, we introduce some notation. Given $\chi : E(G) \to [r]$ and a colour $i \in [r]$, we will denote by $G_i^{(\chi)}$ the subgraph consisting of the edges of G that are coloured i by χ. We will omit the dependency of $G_i^{(\chi)}$ on χ when the colouring is evident from context, writing simply G_i.

The induction assumption is that for every sufficiently large subset $W \subseteq V(G)$, and every collection of graphs $F_1, \dots, F_r$ such that

$$v(F_1), \dots, v(F_r) \leqslant k \qquad \text{and} \qquad \sum_{i=1}^{r} v(F_i) = rk - 1,$$

the following holds: in any r-colouring of $G[W]$, there is a colour $i \in [r]$ for which the copies of $F_i \subseteq G_i[W]$ are well-distributed. To be precise, and using the definitions that we have just introduced, we will have that in this colour $i \in [r]$, the hypergraph $\mathfrak{I}_{F_i,G_i,G}[W]$ is $(p, p|W|)$-Janson.

Definition 2.6. Given $p > 0$, $r \in \mathbb{N}$ and $\mathbf{s} = (s_i)_{i\in[r]} \in \mathbb{N}^r$, let $\mathcal{E}(\mathbf{s}) = \mathcal{E}(\mathbf{s}; p)$ be the family of graphs G with the following property. Fix $\delta = r^{-50}$. For all graphs $F_1, \dots, F_r$ satisfying

$$\sum_{i=1}^{r} v(F_i) = \sum_{i=1}^{r} s_i - 1 \qquad \text{and} \qquad v(F_i) \leqslant s_i \ \text{ for each } i \in [r],$$

for every $W \subseteq V(G)$ with

$$|W| \geqslant \frac{\delta}{8r}\, v(G),$$

and every colouring $\chi : E(G[W]) \to [r]$, there is $i \in [r]$ such that $\mathfrak{I}_{F_i,G_i,G}[W]$ is $(p, p|W|)$-Janson.

We now define the event that corresponds to the failure of the induction step, $\mathcal{B}(\mathbf{H})$.

Definition 2.7. Given $p > 0$ and a collection of graphs $H_1, \dots, H_r$, let $\mathbf{H} = (H_i)_{i\in[r]}$ and let $\mathcal{B}(\mathbf{H}) = \mathcal{B}(\mathbf{H}; p)$ be the family of graphs G with the following property. There exists a colouring $\chi : E(G) \to [r]$ such that, for every $i \in [r]$, the hypergraph $\mathfrak{I}_{H_i,G_i,G}$ is not $(p, p\, v(G))$-Janson.

To simplify the notation, whenever all the $(H_i)_{i\in[r]}$ are equal to a single graph H, we denote this by $\mathcal{B}(H; p, r)$, and omit the p to leave only $\mathcal{B}(H; r)$ when clear from context. Generalising

the notation used in the introduction, we write

$$G \xrightarrow[r]{\text{ind}} H$$

to denote that, for any r-colouring of the edges of G, there exists an (induced) monochromatic copy of H. Observe that if $\mathbb{P}\big(G \in \mathcal{B}(H;r)\big) < 1$, then there exists a graph G such that $G \xrightarrow[r]{\text{ind}} H$.

We can now state the result from which we will deduce Theorem 1.2.

Lemma 2.8. *There exists an absolute constant $C' > 0$ such that the following holds. For all $k, r \in \mathbb{N}$ with $r \geqslant 2$ and for all graphs $H_1, \dots, H_r$ with at most k vertices, if $N \in \mathbb{N}$ satisfies*

$$N \geqslant r^{C'(k+t)} \tag{2.5}$$

for $t = \sum_{i=1}^{r} v(H_i)$, then

$$\mathbb{P}\big(G \in \mathcal{B}(\mathbf{H};p) \cap \mathcal{E}(\mathbf{s};p)\big) \leqslant 2^{-\delta^2 N^2},$$

where $G \sim \mathbb{G}(N,1/2)$, $p = 1/(2^{25}k^2r^4)$, $\delta = r^{-50}$, $\mathbf{H} = (H_i)_{i\in[r]}$ and $\mathbf{s} = \big(v(H_i)\big)_{i\in[r]}$.

To prove that $\mathbb{P}\big(G \in \mathcal{B}(H;r)\big) < 1$, we will find, for each $G \in \mathcal{B}(H;r)$, a large set $W \subseteq V(G)$ and graphs $H_1, \dots, H_r$ such that $G[W] \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})$. Then, we will apply Lemma 2.8 to each $G[W]$ and take a union bound over choices of W and $H_1, \dots, H_r$ to complete the proof.

Proof that Lemma 2.8 implies Theorem 1.2. Throughout the proof, the value of p is always going to be $p = 1/(2^{25}k^2r^4)$, so we omit it in the definition of the events. Fix

$$N = r^{Crk} \tag{2.6}$$

with the goal of showing that there exists G on N vertices such that $G \xrightarrow[r]{\text{ind}} H$. As previously observed, it will suffice to prove that

$$\mathbb{P}\big(G \in \mathcal{B}(H;r)\big) < 1, \tag{2.7}$$

for $G \sim \mathbb{G}(N,1/2)$. The next claim will let us bound this probability with a union bound.

Claim 2.9. *For every $G \in \mathcal{B}(H;r)$, there are $W \subseteq V(G)$ and $H_1, \dots, H_r \subseteq K_k$ such that*

$$|W| \geqslant \left(\frac{\delta}{8r}\right)^{rk-t} N \tag{2.8}$$

for $t = \sum_{i=1}^{r} v(H_i)$, and

$$G[W] \in \mathcal{E}(\mathbf{s}) \cap \mathcal{B}(\mathbf{H}) \tag{2.9}$$

where $\delta = r^{-50}$, $\mathbf{H} = (H_i)_{i\in[r]}$ and $\mathbf{s} = \big(v(H_i)\big)_{i\in[r]}$.

Proof. Choosing $H_1, \dots, H_r = H$ and $W = V(G)$ trivially satisfies (2.8) and $G[W] \in \mathcal{B}(\mathbf{H})$ by $G \in \mathcal{B}(H;r)$, which is a choice of $\mathbf{H}$ and W that satisfies the assumptions in the claim. The existence of one such choice means that we can take $\mathbf{H} = (H_i)_{i\in[r]}$ and $W \subseteq V(G)$ to minimise $t = \sum_{i=1}^{r} v(H_i)$ for $H_1, \dots, H_r \subseteq K_k$ among the choices that satisfy $G[W] \in \mathcal{B}(\mathbf{H})$ and (2.8). We claim that this $\mathbf{H}$ and W also satisfies $G[W] \in \mathcal{E}(\mathbf{s})$, and therefore (2.9).

Indeed, if $G[W] \notin \mathcal{E}(\mathbf{s})$, then, by definition, there are graphs $H'_1, \dots, H'_r$ with

$$\sum_{i=1}^{r} v(H'_i) = t - 1, \tag{2.10}$$

a set $W' \subseteq W$ with

$$|W'| \geqslant \frac{\delta}{8r}\,|W|,$$

and a colouring $\chi : E(G[W']) \to [r]$ such that $\mathfrak{I}_{H'_i,G_i,G}[W']$ is not $(p, p|W'|)$-Janson for all $i \in [r]$. But then W' and $\mathbf{H}' = (H'_i)_{i\in[r]}$ also satisfy $G[W'] \in \mathcal{B}(\mathbf{H}')$ and

$$|W'| \geqslant \frac{\delta}{8r}\,|W| \geqslant \left(\frac{\delta}{8r}\right)^{rk-(t-1)} N,$$

since W satisfies (2.8). Therefore, W' would also satisfy (2.8), and the existence of W' and $\mathbf{H}'$ would, by (2.10), contradict the minimality of t in the original choice of W and $\mathbf{H}$, so $G[W] \in \mathcal{E}(\mathbf{s})$. ■

Applying Claim 2.9, we can take a union bound over the choices of $\mathbf{H}$ and $W \subseteq V(G)$ in that statement, obtaining as a result

$$\mathbb{P}\big(G \in \mathcal{B}(H; r)\big) \leqslant \sum_{\mathbf{H}} \sum_{\substack{W\subseteq V(G)\\ |W|\geqslant\delta^{2rk}N}} \mathbb{P}\big(G[W] \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s}_{\mathbf{H}})\big), \tag{2.11}$$

where $\mathbf{s}_{\mathbf{H}} = \big(v(H_i)\big)_{i\in[r]}$, and we have bounded $(\delta/8r)^{rk-t}N \geqslant \delta^{2rk}N$ for all $t \geqslant 0$ by $r \geqslant 2$ and the value of $\delta = r^{-50} \leqslant 1/(8r)$.

Now, with the goal of bounding each term in (2.11) by applying Lemma 2.8, fix $W \subseteq V(G)$ with $|W| \geqslant \delta^{2rk}N$ and $\mathbf{H} = (H_i)_{i\in[r]}$, a collection of graphs, each with at most k vertices. Observe that letting $G' = G[W]$, we have $G' \sim \mathbb{G}(N', 1/2)$ where $N' = |W|$, and also that $\mathbf{H}$ trivially satisfies

$$t = \sum_{i=1}^{r} v(H_i) \leqslant rk. \tag{2.12}$$

The only assumption in Lemma 2.8 that remains to be checked is that N' satisfies (2.5), which it does because

$$N' \geqslant \delta^{2rk}N \geqslant r^{-100rk}r^{Crk} \geqslant r^{2C'rk} \geqslant r^{C'(k+t)}$$

by our choice of $N = r^{Crk}$ in (2.6) and $\delta = r^{-50}$, by choosing $C \geqslant 2C' + 100$, and finally by (2.12).

Applying Lemma 2.8 to $G' = G[W]$ and $\mathbf{H}$, we conclude that

$$\mathbb{P}\big(G[W] \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s}_{\mathbf{H}})\big) \leqslant 2^{-\delta^2|W|^2} \leqslant 2^{-\delta^{5rk}N^2}, \tag{2.13}$$

for all $|W| \geqslant \delta^{2rk}N$. As W and $\mathbf{H}$ were arbitrary, (2.13) holds for all terms in (2.11).

Replacing (2.13) in (2.11), bounding the choices for $\mathbf{H}$ by $k^r 2^{r\binom{k}{2}}$ and for $W \subseteq V(G)$ by 2^N,

we have

$$\mathbb{P}\big(G \in \mathcal{B}(H;r)\big) \leqslant k^r 2^{r\binom{k}{2}} 2^N 2^{-\delta^{5rk}N^2} \leqslant 2^{-\delta^{5rk}N^2/2} < 1 \tag{2.14}$$

where the final inequalities hold since

$$\frac{\delta^{5rk}N^2}{2} > N + r\log k + r\binom{k}{2} > 0$$

by our choice of $N = r^{Crk}$ and the fact that $\delta = r^{-50}$, taking $C > 256$. We have established (2.7) and completed the proof. □

Remark 2.10. Observe that the bound in (2.14) is sufficiently small for us to take another union bound, this one over all graphs H with k vertices, for which there are at most $2^{\binom{k}{2}}$ choices. The conclusion is that, with high probability, if $G \sim \mathbb{G}(N, 1/2)$, then $G \xrightarrow[r]{\text{ind}} H$ for all such H.

2.3 Warm-up to the extension lemma

In this section, we give an essentially complete proof of a weak "extension lemma", Lemma 2.11. Very roughly, this result bounds the probability that a fixed vertex of G fails to complete any copy of a graph F. We will describe the setting for this result in Section 2.3.1, connecting it with the application of its stronger variant in the proof of Lemma 2.8.

Of central importance in the proof of Lemma 2.11 is a hypergraph container theorem, here stated as Theorem 2.15, which motivates many of the definitions that follow. The most significant of these is that of the auxiliary hypergraph $\mathcal{H}$, defined in Section 2.3.2; the other definitions will, above all, assist in reasoning about it.

As we will later see, each edge of this auxiliary hypergraph represents one way to extend, using a fixed vertex v, an (induced) copy of F^- to a copy of F. Observation 2.13 then fulfils a crucial requirement to apply the container method, establishing that independent sets in $\mathcal{H}$ correspond to pairs of graphs (G', G) in which the neighbourhood of v fails to extend every F^- in a fixed set U.

2.3.1 Extension lemmas

The setting of Lemma 2.11 is that $\tilde{G}'$ and $\tilde{G}$ are m-vertex graphs such that $\tilde{G}' \subseteq \tilde{G}$ and $\mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}[W]$ is $(p, 2pm)$-Janson for every $W \subseteq V(\tilde{G})$ with $|W| \geqslant m/16$. Let $U = V(\tilde{G})$ and let G be the random graph $\mathbb{G}(m+1, 1/2)$ conditioned on $G[U] = \tilde{G}$. The lemma bounds the probability that G contains a subgraph G' with the following properties:

(a) $G'[U] = \tilde{G}'$,

(b) the vertex $v \in V(G) \setminus U$ does not have small degree in G', and

(c) v does not extend any copy of $F^- \subseteq \tilde{G}'$ to an induced F in G using the edges in G'.

In our application, U will be a set of vertices for which the colouring $\chi : E(G) \to [r]$ is fixed, v will be a vertex not in U such that $d_{G_i}(v, U) \geqslant |U|/4r$ for some colour $i \in [r]$, and we

will set $\tilde{G} = G[U]$ and $\tilde{G}' = G_i[U]$. In this setting, we could use this lemma (when $r = 2$) to conclude that it is extremely likely that v extends some i-coloured copy of $F^- := H_i^-$ to a copy of $F = H_i$.

Lemma 2.11. *Let $m, s, k \in \mathbb{N}$ with $s < k \leqslant m$, $p = 2^{-20}k^{-2}$ and F be a graph on $s+1$ vertices. Further let $\tilde{G}'$ and $\tilde{G}$ be graphs on m vertices satisfying $\tilde{G}' \subseteq \tilde{G}$.*

If $\mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}[W]$ is $(p, 2pm)$-Janson for every $W \subseteq U = V(\tilde{G})$ with $|W| \geqslant m/16$, then

$$\mathbb{P}\left(\exists\, G' \subseteq G : \begin{array}{c} G'[U] = \tilde{G}',\ d_{G'}(v) \geqslant m/8 \\ \text{and}\ \ \mathfrak{I}_{F,G',G} = \emptyset \end{array} \,\middle|\, G[U] = \tilde{G}\right) \leqslant 2^{-m/32}, \tag{2.15}$$

where $G \sim \mathbb{G}(m+1, 1/2)$ and $V(G) = U \cup \{v\}$.

We refer to Lemma 2.11 as a weak extension lemma because it concerns $\{\mathfrak{I}_{F,G',G} = \emptyset\}$, whereas our main extension result, Lemma 2.32, is instead about $\{\mathfrak{I}_{F,G',G}$ is not (p,σ)-Janson$\}$. The proof of Lemma 2.11 presents most of the ideas that we will require in Section 2.5, while avoiding some technicalities from dealing with this more complicated property of $\mathfrak{I}_{F,G',G}$.

2.3.2 The auxiliary hypergraph

We want to bound the probability in (2.15) using the method of hypergraph containers. To accomplish that, the first step is to represent the graphs $G' \subseteq G$ such that $G'[U] = \tilde{G}'$ and $\mathfrak{I}_{F,G',G} = \emptyset$ as independent sets in an auxiliary hypergraph $\mathcal{H}$. This auxiliary hypergraph will be a function of $\tilde{G}'$ and $\tilde{G}$ only, which crucially means that $\mathcal{H}$ does not depend on the edges of G (nor on the edges of $G' \subseteq G$) between v and U.

Define $w \in V(F) \setminus V(F^-)$ and note that the only randomness in the event inside the probability in (2.15) comes from the edges of G between U and v, since we condition on the event $\{G[U] = \tilde{G}\}$. Our goal is to have each edge of $\mathcal{H}$ correspond to one way of extending a copy of F^- in $\tilde{G}'$, which is also induced in $\tilde{G}$, to a (still induced) copy of F, by adding the vertex v in the role of w. To be precise, for each edge $L \in \mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}$, fix a bijection $\phi_L : L \to V(F^-)$ such that $F^- \cong_{\phi_L} \tilde{G}'[L] = \tilde{G}[L]$. Now define $\mathcal{H}$ to be the hypergraph with vertex set $U \times \{0,1\}$ and edge set

$$\mathcal{H} = \Big\{E_L : L \in \mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}\Big\}, \tag{2.16}$$

where

$$E_L = \Big\{\Big(u, \mathbb{1}\big[\phi_L(u) \in \mathrm{N}_F(w)\big]\Big) : u \in L\Big\},$$

for every $L \in \mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}$. In words, the vertex set of $\mathcal{H}$ is two copies of U, corresponding to the neighbours and non-neighbours of w in F, and its edge set has the following property: if $u \in \mathrm{N}_{G'}(v)$ for every $(u,1) \in E_L$ and $u \notin \mathrm{N}_G(v)$ for every $(u,0) \in E_L$, then $L \cup \{v\} \in \mathfrak{I}_{F,G',G}$.

To formalise that independent sets in $\mathcal{H}$ correspond to pairs of graphs (G', G) of our interest, denote the non-neighbours of $v \notin U$ by

$$\mathrm{N}_G(v)^{\mathsf{c}} = U \setminus \mathrm{N}_G(v) \tag{2.17}$$

and, for each $A \subseteq V(\mathcal{H})$ and $i \in \{0,1\}$, let

$$A^{(i)} = \big\{u \in U : (u,i) \in A\big\}.$$

We repeat the following property, previously stated without these definitions, for future reference.

Observation 2.12. *Let G' and G be graphs satisfying $G' \subseteq G$, $G'[U] = \tilde{G}'$ and $G[U] = \tilde{G}$. If $L \in \mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}$ and $E = E_L \in \mathcal{H}$, then $L \cup \{v\} \in \mathfrak{I}_{F,G',G}$ (i.e. we can extend L to a copy of $F \subseteq G'$ induced in G using v) whenever*

$$E^{(0)} \subseteq \mathrm{N}_G(v)^{\mathsf{c}} \qquad \text{and} \qquad E^{(1)} \subseteq \mathrm{N}_{G'}(v).$$

The most useful consequence of this fact, which allows applying the method of hypergraph containers, is more conveniently stated with some additional notation. Let

$$\Gamma_G = \left\{G' \subseteq G : \begin{array}{c} G'[U] = \tilde{G}',\ d_{G'}(v) \geqslant m/8 \\ \text{and } \mathfrak{I}_{F,G',G} = \emptyset \end{array}\right\} \tag{2.18}$$

be the collection of G' whose existence is the event in the probability of (2.15). To index vertex subsets of $\mathcal{H}$ by graphs G' and G, further let

$$\iota(G',G) = \big\{(u,0) : u \in \mathrm{N}_G(v)^{\mathsf{c}}\big\} \cup \big\{(u,1) : u \in \mathrm{N}_{G'}(v)\big\}, \tag{2.19}$$

and note that if $I = \iota(G',G)$, then

$$I^{(0)} = \mathrm{N}_G(v)^{\mathsf{c}} \qquad \text{and} \qquad I^{(1)} = \mathrm{N}_{G'}(v). \tag{2.20}$$

To justify one of the premises in Observation 2.13, recall that we have conditioned the distribution of $G \sim \mathbb{G}(m+1,1/2)$ on $\big\{G[U] = \tilde{G}\big\}$ in (2.15).

Observation 2.13. *If $G[U] = \tilde{G}$ and $G' \in \Gamma_G$, then $\iota(G',G)$ is an independent set in $\mathcal{H}$.*

Proof. Let $E = E_L \in \mathcal{H}$, and suppose by contradiction that $E \subseteq I = \iota(G',G)$. It then follows from (2.20) that $E^{(1)} \subseteq \mathrm{N}_{G'}(v)$ and $E^{(0)} \subseteq \mathrm{N}_G(v)^{\mathsf{c}}$, so Observation 2.12 implies that $L \cup \{v\} \in \mathfrak{I}_{F,G',G}$. Therefore, the hypergraph $\mathfrak{I}_{F,G',G}$ is not empty, and we could not have $G' \in \Gamma_G$ by definition, (2.18). This contradiction means that $E \not\subseteq I$, which, as E was an arbitrary edge of $\mathcal{H}$, means that $I = \iota(G',G)$ is an independent set in $\mathcal{H}$. □

2.3.3 Containers

Having connected independent sets in $\mathcal{H}$ to pairs of graphs (G',G) such that $G' \in \Gamma_G$, we can proceed to the next step of the proof, which is applying a container theorem to $\mathcal{H}$. In this section, we use a simpler version of the container theorem than the one required in Section 2.5, since this avoids technical complications that arise in the proof of Lemma 2.32.

Even though a previous result of Balogh and Samotij [15, Theorem 2.1] would most likely suffice to prove Lemma 2.11, we rely on the following theorem of Campos and Samotij [34]. In

what follows, we say that a hypergraph $\mathcal{C}$ covers another hypergraph $\mathcal{G}$ when $\mathcal{G} \subseteq \langle \mathcal{C} \rangle$, where, recall,

$$\langle \mathcal{C} \rangle = \{L \subseteq V(\mathcal{C}) : \exists A \in \mathcal{C}, A \subseteq L\}$$

is the up-set of $\mathcal{C}$.

Theorem 2.14 (Campos and Samotij [34, Theorem A]). *Let $\mathcal{G}$ be an s-uniform hypergraph with n vertices. For every $0 < p' \leqslant 1/(8s^2)$, there exists a family $\mathcal{S} \subseteq 2^{V(\mathcal{G})}$ and functions*

$$\phi : \mathcal{I}(\mathcal{G}) \to \mathcal{S} \quad \text{and} \quad \psi : \mathcal{S} \to 2^{V(\mathcal{G})} \tag{2.21}$$

such that:

(A) *For each $I \in \mathcal{I}(\mathcal{G})$, we have $\phi(I) \subseteq I \subseteq \psi(\phi(I))$.*

(B) *Each $S \in \mathcal{S}$ has at most $8s^2p'n$ elements.*

(C) *For every $S \in \mathcal{S}$, letting $X = \psi(S)$, there exists a hypergraph $\mathcal{C}$ on X with*

$$w_{p'}(\mathcal{C}) \leqslant p'|X| \tag{2.22}$$

that covers $\mathcal{G}[X]$ and satisfies $|E| \geqslant 2$ for all $E \in \mathcal{C}$.

The variant of Theorem 2.14 that we use, Theorem 2.15 below, provides a small family of containers for the independent sets of $\mathcal{H}$, with each container inducing a subhypergraph that is not (p, σ)-Janson.

Theorem 2.15 (Campos and Samotij [34, modified Theorem A]). *Let $\mathcal{G}$ be an s-uniform hypergraph with n vertices. For all $0 < \zeta \leqslant 1$ and $0 < p \leqslant \zeta/(8s^2)$, there is a family $\mathcal{S} \subseteq 2^{V(\mathcal{G})}$ and functions*

$$\phi : \mathcal{I}(\mathcal{G}) \to \mathcal{S} \quad \text{and} \quad \psi : \mathcal{S} \to 2^{V(\mathcal{G})} \tag{2.23}$$

such that:

(i) *For each $I \in \mathcal{I}(\mathcal{G})$, we have $\phi(I) \subseteq I \subseteq \psi(\phi(I))$.*

(ii) *Each $S \in \mathcal{S}$ has at most $8s^2pn/\zeta$ elements.*

(iii) *For every $S \in \mathcal{S}$, letting $X = \psi(S)$, $\mathcal{G}[X]$ is not $(p, \zeta p|X|)$-Janson.*

For the proof of Theorem 2.15 from Theorem 2.14, recall from [34] that when $\mathcal{C}$ is a hypergraph and $0 < p < 1$, we denote by

$$w_p(\mathcal{C}) = \sum_{E \in \mathcal{C}} p^{|E|}$$

what is called the p-weight of $\mathcal{C}$. More generally, if $\vartheta : 2^V \to \mathbb{R}_{\geqslant 0}$ is a measure on the subsets of some vertex set V, we write

$$w_p(\vartheta) = \sum_{L \subseteq V} \vartheta(L) p^{|L|}$$

for its p-weight.

Proof of Theorem 2.15. We apply Theorem 2.14 to $\mathcal{G}$ with parameter $p' = p/\zeta \leqslant 1/(8s^2)$ and obtain ϕ, ψ and $\mathcal{S}$. Items (i) and (ii) in Theorem 2.15 are direct consequences of items (A) and (B) in Theorem 2.14, so it remains only to show that item (C) implies item (iii). Fix $S \in \mathcal{S}$ and the corresponding $X = \psi(S)$, let $\mathcal{C}$ be the cover of $\mathcal{G}[X]$ given by item (C) in Theorem 2.14. We will in fact prove the more general assertion in which $\mathcal{C}$ is replaced by a measure $\vartheta : 2^X \to [0,1]$ that is supported on sets of size at least 2, satisfies

$$\sum_{L \subseteq E} \vartheta(L) \geqslant 1 \qquad \text{for every } E \in \mathcal{G}[X], \tag{2.24}$$

and has p'-weight bounded by $p'|X|$. Indeed, the case provided by Theorem 2.14 is recovered by taking ϑ to be the indicator measure of $\mathcal{C}$. Notice that, by the assumed bound on the p'-weight of ϑ, we have

$$w_{p'}(\vartheta) \leqslant p'|X| = p|X|/\zeta. \tag{2.25}$$

We can use (2.25) and combine the assumption $\zeta \leqslant 1$ with the fact that ϑ is supported on sets of size at least 2 to bound the p-weight of ϑ from its p'-weight:

$$w_p(\vartheta) = \sum_{L \subseteq X} \vartheta(L) p^{|L|} \leqslant \zeta^2 \sum_{L \subseteq X} \vartheta(L)(p/\zeta)^{|L|} = \zeta^2 w_{p'}(\vartheta) \leqslant \zeta p|X|. \tag{2.26}$$

To show that $\mathcal{G}[X]$ is not $(p, \zeta p|X|)$-Janson, take any measure $\nu : \mathcal{G}[X] \to \mathbb{R}_{\geqslant 0}$ with the goal of establishing that

$$\Lambda_p(\nu) \geqslant \frac{e(\nu)^2}{\zeta p|X|}.$$

Observe first that

$$\Lambda_p(\nu) = \sum_{\substack{L \subseteq X, \\ |L| \geqslant 2}} d_\nu(L)^2 \, p^{-|L|} \geqslant \sum_{L \subseteq X} \vartheta(L) d_\nu(L)^2 p^{-|L|}, \tag{2.27}$$

since ϑ is supported on sets of size at least 2, we have $\vartheta(L) \leqslant 1$ for every $L \subseteq X$, and all the terms in the sum are non-negative. Rearranging (2.27), we can apply the Cauchy–Schwarz inequality to obtain

$$\Lambda_p(\nu) \geqslant \frac{1}{w_p(\vartheta)} \left(\sum_{L \subseteq X} \frac{\vartheta(L) d_\nu(L)^2}{p^{|L|}} \right) \left(\sum_{L \subseteq X} \vartheta(L) p^{|L|} \right) \geqslant \frac{1}{w_p(\vartheta)} \left(\sum_{L \subseteq X} \vartheta(L) d_\nu(L) \right)^2. \tag{2.28}$$

Now, note that, as ϑ satisfies (2.24), we have

$$\sum_{L \subseteq X} \vartheta(L) d_\nu(L) = \sum_{E \in \mathcal{G}[X]} \nu(E) \sum_{L \subseteq E} \vartheta(L) \geqslant \sum_{E \in \mathcal{G}[X]} \nu(E) = e(\nu). \tag{2.29}$$

Combining (2.26) and (2.29) with (2.28), we obtain

$$\Lambda_p(\nu) \geqslant \frac{e(\nu)^2}{\zeta p|X|},$$

which completes the proof because ν was arbitrary. □

We will apply Theorem 2.15 with $\mathcal{G} = \mathcal{H}$ and define $\mathcal{X} = \{\psi(S) : S \in \mathcal{S}\}$ to be our container family. Combining item (i) in the statement of that theorem with Observation 2.13 guarantees that whenever G' and G satisfy $G[U] = \tilde{G}$ and $G' \in \Gamma_G$, there is some container $X \in \mathcal{X}$ for which $\iota(G', G) \subseteq X$. Our goal is now to show that $|X^{(0)}| \leqslant (1-\gamma)|U|$ for some constant $\gamma > 0$ using item (iii) in Theorem 2.15 and a supersaturation result. This will be sufficient to obtain the probability bound in Lemma 2.11, because $\mathrm{N}_G(v)^{\mathsf{c}} \subseteq X^{(0)}$ and this non-neighbourhood is a binomial random set, due to $G \sim \mathbb{G}(m+1, 1/2)$.

To prove the required supersaturation result, we will use our assumption that $\mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}$ is $(p, 2pm)$-Janson for every $W \subseteq U$ with $|W| \geqslant m/16$. However, to connect this assumption with the fact that $\mathcal{H}[X]$ is not $(p, p|X|)$-Janson for a container $X \in \mathcal{X}$, we must relate the Janson properties of the hypergraphs $\mathcal{H}$ and $\mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}$.

Towards that goal, let $\pi : U \times \{0,1\} \to U$ be the projection of each pair (u, i) onto its first coordinate u. Observe that if $E = E_L \in \mathcal{H}$ for some $L \in \mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}$, then $\pi(E) = L$. Moreover, for every $L \in \mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}$, there is $E \in \mathcal{H}$ such that $\pi(E) = L$. We conclude that $\pi(\mathcal{H}) = \mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}$, where we extend the application of π from a single vertex to hypergraphs $\mathcal{G}$ by

$$V\big(\pi(\mathcal{G})\big) = \pi(V(\mathcal{G})) \qquad \text{and} \qquad E\big(\pi(\mathcal{G})\big) = \big\{\pi(E) : E \in \mathcal{G}\big\}.$$

Observation 2.16. *If $\pi : U \times \{0,1\} \to U$ is the projection onto the first coordinate, then*

$$\pi(\mathcal{H}) = \mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}.$$

The next lemma will allow us to prove that if $\mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}$ is (p,σ)-Janson, then so is $\mathcal{H}$. We state it in greater generality than we need, but it is a trivial exercise to observe that π satisfies this more general condition. Lemma 2.17 is in fact a corollary of Lemma 2.52, so we will defer its proof to when that other result is proven (see Section 2.7).

Lemma 2.17. *Let $\sigma \geqslant 0$ and $p > 0$. Further let $\mathcal{G}$ be a hypergraph and $\pi : V(\mathcal{G}) \to U$ be a function satisfying*

$$|\pi(E)| = |E| \qquad \text{for every } E \in \mathcal{G}. \tag{2.30}$$

If $\mathcal{G}$ is not (p,σ)-Janson, then $\pi(\mathcal{G})$ is also not (p,σ)-Janson.

We now record the trivial observation that π satisfies (2.30) when $\mathcal{G} = \mathcal{H}$ for future reference.

Observation 2.18. *$|\pi(E)| = |E|$ for every $E \in \mathcal{H}$.*

It follows from item (iii) in Theorem 2.15 and Lemma 2.17 that $\pi(\mathcal{H}[X])$ is not $(p, p|X|)$-Janson for each container X. However, our assumption in Lemma 2.11 is that $\pi(\mathcal{H})[W]$ is $(p, 2pm)$-Janson for all large sets $W \subseteq U$, and this does not immediately imply anything about the size of X. For example, we might try to apply this assumption with $W = \pi(X)$, and hope that

$$\pi(\mathcal{H})[W] \subseteq \pi(\mathcal{H}[X]), \tag{2.31}$$

but unfortunately (2.31) does not always hold: if $X = U \times \{0\}$, then we have $\pi(\mathcal{H})[\pi(X)] = \pi(\mathcal{H})$ and $\mathcal{H}[X] = \emptyset$. In order to deal with this issue, we will instead apply our assumption to the set $W = X^{(0)} \cap X^{(1)}$. By the following lemma, this choice *does* satisfy (2.31).

Lemma 2.19. *For all $Y \subseteq U \times \{0,1\}$, we have*

$$\pi(\mathcal{H})[W] \subseteq \pi(\mathcal{H}[Y]), \tag{2.32}$$

where $W = Y^{(0)} \cap Y^{(1)}$.

Proof. Take $L \in \pi(\mathcal{H})[W]$ with the goal of proving that $L \in \pi(\mathcal{H}[Y])$. By definition of W,

$$L \subseteq Y^{(0)} \cap Y^{(1)}. \tag{2.33}$$

Moreover, since $L \in \pi(\mathcal{H})$, there is $E \in \mathcal{H}$ such that $\pi(E) = L$. Therefore,

$$E \subseteq \pi^{-1}(L) = \big\{(u,a) : u \in L,\ a \in \{0,1\}\big\},$$

but we also have, by (2.33), that

$$\pi^{-1}(L) \subseteq \big(Y^{(0)} \times \{0\}\big) \cup \big(Y^{(1)} \times \{1\}\big) = Y,$$

and so $E \subseteq Y$. We conclude that $E \in \mathcal{H}[Y]$ and also $L = \pi(E) \in \pi(\mathcal{H}[Y])$, as required. □

Now, taking $W = X^{(0)} \cap X^{(1)}$ for some container $X \in \mathcal{X}$ and applying Lemma 2.19, we have that $\pi(\mathcal{H})[W] \subseteq \pi(\mathcal{H}[X])$. Using item (iii) of Theorem 2.15, Observation 2.16, Lemma 2.17 and Observation 2.5, we will conclude that $\mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}[W]$ is not $(p, 2pm)$-Janson (see Claim 2.20). Our assumption in Lemma 2.11 then implies that $|W| < m/16$, so every $X \in \mathcal{X}$ satisfies

$$\big|X^{(0)} \cap X^{(1)}\big| < \frac{m}{16},$$

by definition of W, and, since $X^{(0)} \cup X^{(1)} \subseteq U$, it follows that

$$\big|X^{(0)}\big| + \big|X^{(1)}\big| = \big|X^{(0)} \cup X^{(1)}\big| + \big|X^{(0)} \cap X^{(1)}\big| < \Big(1 + \frac{1}{16}\Big)m. \tag{2.34}$$

To establish that $X^{(0)}$ is not large when $I = \iota(G', G) \subseteq X$, recall that it follows from $G' \in \Gamma_G$ that

$$|X^{(1)}| \geqslant d_{G'}(v) \geqslant \frac{m}{8} \tag{2.35}$$

where the first inequality holds because $\mathrm{N}_{G'}(v) \subseteq X^{(1)}$ by (2.20). Combining (2.34) and (2.35) yields

$$|X^{(0)}| \leqslant \Big(1 - \frac{1}{16}\Big)m,$$

which, together with

$$\mathrm{N}_G(v)^{\mathsf{c}} \subseteq X^{(0)} \tag{2.36}$$

will allow us to use the randomness in the distribution of $G \sim \mathbb{G}(m+1, 1/2)$ to bound the probability of (2.36) by an exponentially small term. The resulting upper bound is sufficiently strong to overcome the union bound over all $X \in \mathcal{X}$, so we can now formalise the proof of Lemma 2.11.

Proof of Lemma 2.11. Applying Theorem 2.15 to $\mathcal{H}$, defined in (2.16), with $\zeta = 1$ we obtain a family $\mathcal{S}$ and functions ϕ, ψ satisfying items (i)–(iii) in Theorem 2.15, so let

$$\mathcal{X} = \{\psi(S) : S \in \mathcal{S}\}.$$

With the goal of taking a union bound over $X \in \mathcal{X}$, fix G on $m+1$ vertices satisfying $G[U] = \tilde{G}$, and observe that $\iota(G', G) \in \mathcal{I}(\mathcal{H})$ for all $G' \in \Gamma_G$ by Observation 2.13. Now, it follows from item (i) of Theorem 2.15 that there exists $X \in \mathcal{X}$ such that $\iota(G', G) \subseteq X$, and therefore

$$\mathrm{N}_G(v)^{\mathsf{c}} = I^{(0)} \subseteq X^{(0)} \qquad \text{and} \qquad \mathrm{N}_{G'}(v) = I^{(1)} \subseteq X^{(1)}, \tag{2.37}$$

where $I = \iota(G', G)$. Using (2.37), we further refine $\mathcal{X}$ to

$$\mathcal{X}' = \big\{X \in \mathcal{X} : |X^{(1)}| \geqslant m/8\big\}.$$

This refinement preserves the property that for every $G' \in \Gamma_G$, there is $X \in \mathcal{X}'$ such that $\iota(G', G) \subseteq X$, because

$$|X^{(1)}| \geqslant d_{G'}(v) \geqslant \frac{m}{8} \tag{2.38}$$

follows from $G' \in \Gamma_G$, defined in (2.18). That is, if $X \in \mathcal{X}$ contains some $\iota(G', G)$, then it is also in $\mathcal{X}'$ by (2.38), and crucially, for every $G' \in \Gamma_G$, there exists $X \in \mathcal{X}'$ such that

$$\mathrm{N}_G(v)^{\mathsf{c}} \subseteq X^{(0)}. \tag{2.39}$$

Taking a union bound over $\mathcal{X}'$ then yields, for the probability in (2.15),

$$\mathbb{P}\big(\exists G' \subseteq G : G' \in \Gamma_G \mid G[U] = \tilde{G}\big) \leqslant \sum_{X \in \mathcal{X}'} \mathbb{P}\big(\mathrm{N}_G(v)^{\mathsf{c}} \subseteq X^{(0)}\big), \tag{2.40}$$

where we used (2.39) to bound the probability of the event $\{\iota(G', G) \subseteq X\}$ by that of the event $\{\mathrm{N}_G(v)^{\mathsf{c}} \subseteq X^{(0)}\}$. We shift our focus to obtaining an upper bound for the probability of this event that holds for every $X \in \mathcal{X}'$, so we fix $X \in \mathcal{X}'$ and set $W = X^{(0)} \cap X^{(1)}$. The following claim, together with the assumption that $\mathfrak{I}_{F^-, \tilde{G}', \tilde{G}}[W']$ is $(p, 2pm)$-Janson for every $W' \subseteq U$ with $|W'| \geqslant m/16$, will allow us to bound the size of $X^{(0)}$ from above.

Claim 2.20. *$\mathfrak{I}_{F^-, \tilde{G}', \tilde{G}}[W]$ is not $(p, 2pm)$-Janson.*

Proof. By Observation 2.16 and Lemma 2.19, we have

$$\mathfrak{I}_{F^-, \tilde{G}', \tilde{G}}[W] = \pi(\mathcal{H})[W] \subseteq \pi(\mathcal{H}[X]).$$

Therefore, if we establish that the hypergraph $\pi(\mathcal{H}[X])$ is not $(p, 2pm)$-Janson, then we are done, because being $(p, 2pm)$-Janson is increasing with respect to inclusion, by Observation 2.5.

To show this, observe first that $\mathcal{H}[X]$ is not $(p, p|X|)$-Janson by item (iii) in Theorem 2.15 and our choice of $\zeta = 1$. Since $|X| \leqslant 2m = v(\mathcal{H})$ it follows from Observation 2.3 that the hypergraph $\mathcal{H}[X]$ is also not $(p, 2pm)$-Janson. Thus, by Lemma 2.17, we deduce that $\pi(\mathcal{H}[X])$ is not $(p, 2pm)$-Janson, and, by our previous reasoning, that neither is $\mathfrak{I}_{F^-, \tilde{G}', \tilde{G}}[W]$, as claimed. ∎

Since we know, by assumption, that $\mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}[W']$ is $(p, 2pm)$-Janson for every $W' \subseteq U$ with $|W'| \geqslant m/16$, Claim 2.20 implies that $|W| < m/16$ and therefore

$$|X^{(0)}| + |X^{(1)}| = \left|X^{(0)} \cup X^{(1)}\right| + \left|X^{(0)} \cap X^{(1)}\right| < m + \frac{m}{16}.$$

As $X \in \mathcal{X}'$, we have $|X^{(1)}| \geqslant m/8$, and therefore

$$|X^{(0)}| < \Big(1 + \frac{1}{16}\Big)m - |X^{(1)}| \leqslant m + \frac{m}{16} - \frac{m}{8} = \Big(1 - \frac{1}{16}\Big)m,$$

that is,

$$|U \setminus X^{(0)}| \geqslant \frac{m}{16}. \tag{2.41}$$

We conclude from (2.41) that

$$\mathbb{P}\Big(\mathrm{N}_G(v)^{\mathsf{c}} \cap \big(U \setminus X^{(0)}\big) = \emptyset\Big) = 2^{-|U \setminus X^{(0)}|} \leqslant 2^{-m/16}, \tag{2.42}$$

where we used that $G \sim \mathbb{G}(m+1, 1/2)$ and $v \notin U$. Replacing (2.42) in (2.40), we obtain

$$\sum_{X \in \mathcal{X}'} \mathbb{P}\big(\mathrm{N}_G(v)^{\mathsf{c}} \subseteq X^{(0)}\big) \leqslant |\mathcal{X}'|\, 2^{-m/16}. \tag{2.43}$$

To bound the size of $\mathcal{X}'$ by $|\mathcal{X}|$, we enumerate the latter using that each container X is a function of $S \in \mathcal{S}$. As each $S \in \mathcal{S}$ satisfies

$$|S| \leqslant 16ps^2m \leqslant 2^{-16}m$$

by item (ii) in Theorem 2.15, our choices of $\zeta = 1$ and of $p = 2^{-20}k^{-2}$, and by the assumption that $v(F) = s \leqslant k$, we have

$$|\mathcal{X}'| \leqslant |\mathcal{X}| \leqslant \sum_{t=0}^{2^{-16}m} \binom{2m}{t} \leqslant 2^{m/32}, \tag{2.44}$$

where, recall, $\mathcal{H}$ has $2m$ vertices. Combining (2.44) and (2.43),

$$\mathbb{P}\big(\exists G' \subseteq G : G' \in \Gamma_G \mid G[U] = \tilde{G}\big) \leqslant 2^{m/32}\, 2^{-m/16} \leqslant 2^{-m/32}$$

which completes the proof. □

2.4 A general container theorem for non-Janson sets

In this section, we prove a preliminary version of the main technical result of this chapter. Even though it is not the version of the container theorem that we will end up using to prove the main result of the chapter, it exhibits most of the techniques required to prove the version that we indeed use. Moreover, we believe that Theorem 2.21 should be more broadly applicable than Theorem 2.35.

Theorem 2.21. *Let $s, n \in \mathbb{N}$ with $s \leqslant n$, and let $q, p, \sigma, \eta > 0$ satisfy*

$$q \leqslant \frac{1}{16}, \qquad p \leqslant \frac{q}{2^{10}s^2}, \qquad \sigma \geqslant 2^{-6}pn \qquad \text{and} \qquad \eta = 2^{-2s-2}.$$

For every s-uniform hypergraph $\mathcal{H}$ with n vertices, there exists a family $\mathcal{X} \subseteq 2^{V(\mathcal{H})}$ with

$$|\mathcal{X}| \leqslant \left(\frac{2}{q}\right)^{8qn} \tag{2.45}$$

such that the following hold.

(i) *If $L \subseteq V(\mathcal{H})$ and $\mathcal{H}[L]$ is not $(p/q, \eta\sigma)$-Janson, then $L \subseteq X$ for some $X \in \mathcal{X}$.*

(ii) *For each $X \in \mathcal{X}$, the hypergraph $\mathcal{H}[X]$ is not (p, σ)-Janson.*

We now discuss the proof of Theorem 2.21. The first step is defining the auxiliary hypergraph

$$\mathcal{J} = \big\{L \subseteq V : \mathcal{H}[L] \text{ is } (p/q, \eta\sigma)\text{-Janson}\big\},$$

where $V = V(\mathcal{H})$. It follows immediately from the definition that sets $L \subseteq V$ for which $\mathcal{H}[L]$ is not $(p/q, \eta\sigma)$-Janson are not edges of $\mathcal{J}$, but, more importantly, Observation 2.5 implies that each such L is an independent set in $\mathcal{J}$.

We can now see that item (i) in Theorem 2.21 could be equivalently phrased as "for all $I \in \mathcal{I}(\mathcal{J})$, there exists $X \in \mathcal{X}$ such that $I \subseteq X$". This formulation suggests applying a container theorem to this hypergraph, but the edges of $\mathcal{J}$ could have size comparable to $n = |V|$. To avoid that issue, we first reduce to an alternate setting involving independent sets in s-uniform hypergraphs, and the theorem that we use in that reduction requires the definition of the non-strict link of a hypergraph, which is deceptively similar to the ordinary notion of hypergraph link.

Definition 2.22. For a hypergraph $\mathcal{G}$ and a set Y, let

$$\underline{\partial}_Y \mathcal{G} = \big\{E \setminus Y : E \in \mathcal{G}\big\}$$

denote the non-strict link of $\mathcal{G}$ with respect to Y.

Interestingly, this reduction to s-uniform hypergraphs is proven by applying a container theorem (Theorem 2.23, below) that has no dependency on the uniformity of the hypergraph. It is not immediate that Theorem E in [34] implies Theorem 2.23, so we give a short (and, in fact, self-contained) proof of this result. In the statement and in the rest of this chapter, when $0 \leqslant q \leqslant 1$, we write V_q to denote a q-random subset of V.

Theorem 2.23 (Campos and Samotij [34, modified Theorem E])**.** *Let $\mathcal{G}$ be a hypergraph with vertex set V. For all $\alpha, q \in \mathbb{R}$ satisfying $0 < q \leqslant \alpha < 1$, there exists a family $\mathcal{T} \subseteq 2^V$ and a function $\varphi : \mathcal{I}(\mathcal{G}) \to \mathcal{T}$ such that:*

(a) *For each $I \in \mathcal{I}(\mathcal{G})$, we have $\varphi(I) \subseteq I$.*

(b) *Each $T \in \mathcal{T}$ has at most qn/α elements, where $n = |V|$.*

(c) *For every $T \in \mathcal{T}$, there exists a hypergraph $\mathcal{C}_T$ with vertex set V that covers $\mathcal{G}$ and satisfies*

$$\mathbb{P}\big(L \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big) > (1-\alpha)^{|L|} q^{|L|} \tag{2.46}$$

for all $L \notin \mathcal{C}_T$; moreover, for all $I \in \mathcal{I}(\mathcal{G})$ such that $T = \varphi(I)$, we have $I \in \mathcal{I}(\mathcal{C}_T)$.

Proof. Given $I \in \mathcal{I}(\mathcal{G})$, let $T \subseteq I$ be maximal with respect to

$$\mathbb{P}\big(T \subseteq V_q \mid V_q \in \mathcal{I}(\mathcal{G})\big) \leqslant (1-\alpha)^{|T|} q^{|T|}. \tag{2.47}$$

Set $\varphi(I) = T$ and $\mathcal{T} = \{\varphi(I) : I \in \mathcal{I}(\mathcal{G})\}$. We claim that these choices satisfy the requirements of Theorem 2.23. Notice that item (a) trivially holds, since $\varphi(I)$ is defined as a subset of I.

Now, take an arbitrary $T \in \mathcal{T}$. We have, on one hand,

$$(1-q)^{n-|T|} q^{|T|} = \mathbb{P}(V_q = T) \leqslant \mathbb{P}(T \subseteq V_q \wedge V_q \in \mathcal{I}(\mathcal{G})), \tag{2.48}$$

where the inequality is using that $T \in \mathcal{I}(\mathcal{G})$, and, on the other hand,

$$\mathbb{P}(T \subseteq V_q \wedge V_q \in \mathcal{I}(\mathcal{G})) \leqslant \mathbb{P}(T \subseteq V_q \mid V_q \in \mathcal{I}(\mathcal{G})) \leqslant (1-\alpha)^{|T|} q^{|T|} \tag{2.49}$$

due to (2.47). Combining (2.48) and (2.49), we obtain

$$(1-q)^{n-|T|} q^{|T|} \leqslant (1-\alpha)^{|T|} q^{|T|},$$

which, as $x \mapsto (1-x)^{1/x}$ is decreasing on $(0,1)$, implies that

$$|T| \leqslant qn/\alpha. \tag{2.50}$$

But we took T arbitrarily, so (2.50) establishes item (b).

To show that item (c) holds, take any $T \in \mathcal{T}$ and set

$$\mathcal{C}_T = \Big\{ L \subseteq V(\mathcal{G}) : \mathbb{P}\big(L \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big) \leqslant (1-\alpha)^{|L|} q^{|L|} \Big\}. \tag{2.51}$$

Observe that not only $\mathcal{C}_T$ is a cover of $\mathcal{G}$, but also $\mathcal{G} \subseteq \mathcal{C}_T$: for all $E \in \mathcal{G}$, we have

$$\mathbb{P}\big(E \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big) = 0,$$

because the combination of $V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})$ with $E \subseteq V_q$ implies that $(E \setminus T) \in \mathcal{I}(\underline{\partial}_T \mathcal{G})$, which directly contradicts the definition of $\underline{\partial}_T \mathcal{G}$. The definition in (2.51) also immediately implies that, for all $L \notin \mathcal{C}_T$,

$$\mathbb{P}\big(L \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big) > (1-\alpha)^{|L|} q^{|L|}$$

holds, and that is exactly (2.46) in item (c). It remains only to establish the "moreover" part of item (c).

Our goal is to show that for all $I \in \mathcal{I}(\mathcal{G})$, if $\varphi(I) = T$, then $I \in \mathcal{I}(\mathcal{C}_T)$, so we take $L \in \mathcal{C}_T$

and must determine that $L \not\subseteq I$. Note that $L \not\subseteq T$ follows from $L \in \mathcal{C}_T$, as otherwise

$$\mathbb{P}\big(L \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big) = \mathbb{P}(L \subseteq V_q) = q^{|L|} > (1-\alpha)^{|L|} q^{|L|},$$

because $\{V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\}$ and $\{L \subseteq V_q\}$ are independent when $L \subseteq T$.

We claim that

$$\mathbb{P}\big(L \cup T \subseteq V_q \mid V_q \in \mathcal{I}(\mathcal{G})\big) = \mathbb{P}\big(T \subseteq V_q \mid V_q \in \mathcal{I}(\mathcal{G})\big)\, \mathbb{P}\big(L \setminus T \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big). \tag{2.52}$$

First, observe that we can reveal $T_q = T \cap V_q$ and then $V_q' = V_q \setminus T$, obtaining as a result

$$\mathbb{P}\big(L \cup T \subseteq V_q \mid V_q \in \mathcal{I}(\mathcal{G})\big) = \mathbb{P}\big(T \subseteq V_q \mid V_q \in \mathcal{I}(\mathcal{G})\big)\, \mathbb{P}\big(L \setminus T \subseteq V_q' \mid T \subseteq V_q \wedge V_q \in \mathcal{I}(\mathcal{G})\big).$$

But $\{T \subseteq V_q\} = \{T_q = T\}$ and $V_q \in \mathcal{I}(\mathcal{G})$ are together equivalent to $V_q' \in \mathcal{I}(\underline{\partial}_T \mathcal{G})$, so we have

$$\mathbb{P}\big(L \setminus T \subseteq V_q \setminus T \mid T \subseteq V_q \wedge V_q \in \mathcal{I}(\mathcal{G})\big) = \mathbb{P}\big(L \setminus T \subseteq V_q' \mid V_q' \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big),$$

and therefore

$$\mathbb{P}\big(L \cup T \subseteq V_q \mid V_q \in \mathcal{I}(\mathcal{G})\big) = \mathbb{P}(T \subseteq V_q \mid V_q \in \mathcal{I}(\mathcal{G}))\, \mathbb{P}\big(L \setminus T \subseteq V_q' \mid V_q' \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big). \tag{2.53}$$

Finally, $L \setminus T$ and $V_q \setminus V_q' \subseteq T$ are disjoint and T_q is independent of $\{V_q' \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\}$, so

$$\mathbb{P}\big(L \setminus T \subseteq V_q' \mid V_q' \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big) = \mathbb{P}\big(L \setminus T \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big) \tag{2.54}$$

and substituting (2.54) in (2.53) yields (2.52).

Now, partitioning L according to its intersection with T, we have

$$\mathbb{P}\big(L \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big) = \mathbb{P}\big(L \cap T \subseteq V_q\big)\, \mathbb{P}\big(L \setminus T \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big)$$

because $\{L \cap T \subseteq V_q\}$ and $\{L \setminus T \subseteq V_q\}$ are independent, and so are $\{L \cap T \subseteq V_q\}$ and $\{V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\}$. Therefore,

$$\mathbb{P}\big(L \setminus T \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big) = \mathbb{P}\big(L \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big) q^{-|L \cap T|}. \tag{2.55}$$

Combining (2.55) with (2.52), we obtain

$$\mathbb{P}\big(L \cup T \subseteq V_q \mid V_q \in \mathcal{I}(\mathcal{G})\big) = \mathbb{P}\big(L \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big)\, \mathbb{P}(T \subseteq V_q \mid V_q \in \mathcal{I}(\mathcal{G}))\, q^{-|L \cap T|},$$

and now we can use that our choice of T satisfies (2.47), and L satisfies (2.51) to obtain

$$\mathbb{P}\big(L \cup T \subseteq V_q \mid V_q \in \mathcal{I}(\mathcal{G})\big) \leqslant (1-\alpha)^{|T|+|L|} q^{|T|+|L|-|L \cap T|} \leqslant (1-\alpha)^{|L \cup T|} q^{|L \cup T|} \tag{2.56}$$

because $0 < \alpha < 1$.

Taking $T' = L \cup T$, it is obvious that $T \subseteq T'$. We have picked $T \subseteq I$ to be maximal satisfying (2.47), so it follows from (2.56) that if $T' \subseteq I$, then we would have picked it instead

of T. Therefore, $T' \not\subseteq I$ and thus $L \not\subseteq I$. As we took $L \in \mathcal{C}_T$ arbitrarily and established that $L \not\subseteq I$, we conclude that $I \in \mathcal{I}(\mathcal{C}_T)$ and so item (c) holds, completing the proof. □

We will apply Theorem 2.23 to $\mathcal{J}$, obtaining as a result a family of sets $T \in \mathcal{T}$, each with a corresponding hypergraph $\mathcal{C}_T$. Now, for each $I \in \mathcal{I}(\mathcal{J})$, there is some $T \in \mathcal{T}$ such that $I \in \mathcal{I}(\mathcal{C}_T)$ by item (c) in Theorem 2.23. Therefore,

$$\mathcal{I}(\mathcal{J}) \subseteq \bigcup_{T \in \mathcal{T}} \mathcal{I}(\mathcal{C}_T). \tag{2.57}$$

However, we also want the hypergraphs in the right-hand side of (2.57) to be s-uniform. The first step is replacing each $\mathcal{C}_T$ by its up-set.

Observation 2.24. *Let $\mathcal{G}$ be a hypergraph. If $I \in \mathcal{I}(\mathcal{G})$, then $I \in \mathcal{I}(\langle \mathcal{G} \rangle)$.*

Proof. Assume that $I \notin \mathcal{I}(\langle \mathcal{G} \rangle)$. Hence, there is $E \in \langle \mathcal{G} \rangle$ such that $E \subseteq I$, and by definition of the up-set, there is also $E' \in \mathcal{G}$ such that $E' \subseteq E \subseteq I$, so $I \notin \mathcal{I}(\mathcal{G})$. □

The trivial observation that completes the reduction to s-uniform hypergraphs says that Observation 2.24 also holds if we replace $\mathcal{C}_T$ by a subhypergraph. In our case, the implication is that if $\mathcal{C}'_T \subseteq \langle \mathcal{C}_T \rangle$ for all $T \in \mathcal{T}$, then

$$\mathcal{I}(\mathcal{J}) \subseteq \bigcup_{T \in \mathcal{T}} \mathcal{I}(\mathcal{C}'_T).$$

The subhypergraph $\mathcal{C}'_T$ that we will take is simply the set of edges with size s, i.e.

$$\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_{=s}$$

where we define

$$\mathcal{G}_{=s} = \{E \in \mathcal{G} : |E| = s\}$$

for any hypergraph $\mathcal{G}$ and $s \in \mathbb{N}$.

Having reduced the problem to a collection of hypergraphs with edges of size s, we can apply another container theorem to each $\mathcal{C}'_T$. Recall that we have previously used the following statement in Section 2.3.

Theorem 2.15 (Campos and Samotij [34, modified Theorem A])**.** *Let $\mathcal{G}$ be an s-uniform hypergraph with n vertices. For all $0 < \zeta \leqslant 1$ and $0 < p \leqslant \zeta/(8s^2)$, there is a family $\mathcal{S} \subseteq 2^{V(\mathcal{G})}$ and functions*

$$\phi : \mathcal{I}(\mathcal{G}) \to \mathcal{S} \quad \text{and} \quad \psi : \mathcal{S} \to 2^{V(\mathcal{G})} \tag{2.58}$$

such that:

(i) *For each $I \in \mathcal{I}(\mathcal{G})$, we have $\phi(I) \subseteq I \subseteq \psi(\phi(I))$.*

(ii) *Each $S \in \mathcal{S}$ has at most $8s^2pn/\zeta$ elements.*

(iii) *For every $S \in \mathcal{S}$, letting $X = \psi(S)$, $\mathcal{G}[X]$ is not $(p, \zeta p|X|)$-Janson.*

Consider the following recap of our overview so far. First, we take an $I \in \mathcal{I}(\mathcal{J})$, and from Theorem 2.23 and Observation 2.24 we obtain $\varphi(I) = T \subseteq I$ and $\mathcal{C}'_T$ such that $I \in \mathcal{I}(\mathcal{C}'_T)$. Then, applying Theorem 2.15 with $\mathcal{G} = \mathcal{C}'_T$ yields a family $\mathcal{S}_T$ from which we retrieve a container $X \supset I$. This container has the property that $\mathcal{C}'_T[X]$ is not $(p, \zeta p|X|)$-Janson, and we want to take X to be the container for this fixed I in Theorem 2.21. It is easy to deduce the claimed bound (2.45) on the number of such containers from the bounds on $|\mathcal{T}|$ and $|\mathcal{S}_T|$ given by the corresponding container theorems. However, it is not yet clear how to deduce that $\mathcal{H}[X]$ is not (p, σ)-Janson when we only know the Janson properties of $\mathcal{C}'_T[X]$.

To show that $\mathcal{H}[X]$ is not (p, σ)-Janson, we fix an arbitrary $\nu : \mathcal{H}[X] \to \mathbb{R}_{\geqslant 0}$ with the objective of establishing that

$$\Lambda_p(\nu) \geqslant \frac{e(\nu)^2}{\sigma} \tag{2.59}$$

which, recall, is the definition of what it means to not be (p, σ)-Janson. We will consider two cases, depending on the relation between ν and ν', the restriction of ν to $\mathcal{C}'_T[X]$, where $T = \varphi(I)$. The first (easier) case is when $e(\nu') \geqslant e(\nu)/2$. Here, we can easily show that (2.59) follows from $\mathcal{C}'_T[X]$ not being $(p, \zeta p|X|)$-Janson (see Claim 2.28).

In the other, more delicate case, we will have that $\nu'' = \nu - \nu'$ satisfies

$$e(\nu'') \geqslant \frac{e(\nu)}{2}. \tag{2.60}$$

Our goal is now to obtain bounds relating $\Lambda_p(\nu)$ and $e(\nu'')^2$ – it will be easy to see that combining them with (2.60) will reach (2.59). Concretely, we will show that

$$2^{2s}\Lambda_p(\nu) > \mathbb{E}\big[\Lambda_{p/q}(\nu''_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J})\big] \geqslant \frac{\mathbb{E}\big[e(\nu''_q)^2 \mid V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J})\big]}{\eta\sigma} \geqslant \frac{e(\nu'')^2}{\eta\sigma} \tag{2.61}$$

where $\nu''_q : \mathcal{H}[X] \to \mathbb{R}_{\geqslant 0}$ is defined by

$$\nu''_q(E) = \frac{\mathbb{1}[E \subseteq V_q]}{\mathbb{P}\big(E \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J})\big)}\,\nu''(E).$$

The simple proof of Claim 2.29 will establish, by inspecting the definitions, that

$$\mathbb{E}\big[e(\nu''_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J})\big] = e(\nu''),$$

which implies the rightmost inequality of (2.61) by convexity. To prove the second inequality, we will in fact show that (deterministically) whenever $V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J})$, we have

$$\Lambda_{p/q}(\nu''_q) \geqslant \frac{e(\nu''_q)^2}{\eta\sigma}. \tag{2.62}$$

Proving (2.62) will require a trivial observation about independent sets in the non-strict link of hypergraphs with respect to a set T.

Observation 2.25. *Let $\mathcal{G}$ be a hypergraph and $T \subseteq V(\mathcal{G})$. If $I \in \mathcal{I}(\underline{\partial}_T \mathcal{G})$, then $I \in \mathcal{I}(\mathcal{G})$.*

Proof. For each $E \in \mathcal{G}$, observe that $E \setminus T \subseteq E$ and $E \setminus T \in \underline{\partial}_T \mathcal{G}$ by definition, so $\mathcal{G} \subseteq \langle \underline{\partial}_T \mathcal{G} \rangle$. The statement now follows from Observation 2.24. □

With Observation 2.25, we will see that the definition of $\mathcal{J}$ and the monotonicity of being $(p/q, \eta\sigma)$-Janson will imply (2.62), see Claim 2.30. The remaining inequality, Claim 2.31, establishes that

$$2^{2s}\Lambda_p(\nu) > \mathbb{E}\big[\Lambda_{p/q}(\nu''_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})\big]. \tag{2.63}$$

To prove it, we will crucially rely on the fact that, for all $E \in \mathcal{H}[X] \setminus \mathcal{C}'_T$,

$$\mathbb{P}\big(E \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})\big) > \Big(\frac{q}{2}\Big)^{|E|}$$

since $\mathcal{H}[X] \setminus \mathcal{C}'_T$ and $\mathcal{C}_T$ are disjoint, and $\mathcal{C}_T$ satisfies item (c) of Theorem 2.23 with $\alpha = 1/2$.

The preceding overview and proof strategy in fact proves Theorem 2.26, which strengthens the original statement and furthermore adds the characterization of sets L for which $\mathcal{H}$ is not $(p/q, \eta\sigma)$-Janson as independent sets in the auxiliary hypergraph $\mathcal{J}$. To obtain this stronger statement, we will redefine $\mathcal{J}$ in (2.65) to have instead as its edges sets L for which $\mathcal{H}[L]$ is $(p/(q-p), \eta\sigma)$-Janson. We will then deduce Theorem 2.21 from Theorem 2.26 by applying the latter with q being $q + p$.

Theorem 2.26. *Let $s, n \in \mathbb{N}$ with $s \leqslant n$, and let $p, q, \alpha, \sigma, \eta > 0$ satisfy*

$$p \leqslant \frac{1}{2^{11}s^2}, \qquad 2p \leqslant q \leqslant \alpha < 1, \qquad \sigma \geqslant 2^{-6}pn \qquad \text{and} \qquad \eta \leqslant \frac{(1-\alpha)^{2s}}{4}. \tag{2.64}$$

Furthermore, for every s-uniform hypergraph $\mathcal{H}$ with n vertices, let

$$\mathcal{J} = \big\{L \subseteq V(\mathcal{H}) : \mathcal{H}[L] \text{ is } (p/(q-p), \eta\sigma)\text{-Janson}\big\}. \tag{2.65}$$

There exists a family $\mathcal{Y} \subseteq 2^{V(\mathcal{H})} \times 2^{V(\mathcal{H})}$ and functions

$$g : \mathcal{I}(\mathcal{J}) \to \mathcal{Y} \qquad \text{and} \qquad f : \mathcal{Y} \to 2^{V(\mathcal{H})}$$

such that:

(1) *For every $I \in \mathcal{I}\big(\mathcal{J}\big)$, if $g(I) = (S, T)$, then $S \cup T \subseteq I \subseteq f(S, T)$.*

(2) *Each $(S, T) \in \mathcal{Y}$ satisfies*

$$|S| \leqslant 2^{11}ps^2n \qquad \text{and} \qquad |T| \leqslant qn/\alpha. \tag{2.66}$$

(3) *For all $Y \in \mathcal{Y}$, the hypergraph $\mathcal{H}[X]$ is not (p, σ)-Janson, where $X = f(Y)$.*

Before proving Theorem 2.26, let us quickly observe that it implies Theorem 2.21.

Proof that Theorem 2.26 implies Theorem 2.21. Apply Theorem 2.26 to $\mathcal{H}$ with $\alpha = 1/2$ and q replaced by $q + p$. Note that

$$p \leqslant \frac{q}{2^{10}s^2} \leqslant \frac{1}{2^{11}s^2} \qquad \text{and} \qquad 2p \leqslant q + p \leqslant \alpha,$$

since $q \leqslant 1/16$. As a result, we obtain functions g, f and a family $\mathcal{Y}$, so let

$$\mathcal{X} = \{f(S,T) : (S,T) \in \mathcal{Y}\}.$$

Observe that

$$|\mathcal{X}| \leqslant \sum_{m=0}^{2^{11}ps^2n} \binom{n}{m} \sum_{t=0}^{2(q+p)n} \binom{n}{t} \leqslant \left(\frac{2}{q}\right)^{8qn} \tag{2.67}$$

by enumerating every possible S and T, and combining the bound on their sizes, (2.66), with

$$2^{11}ps^2n \leqslant 2(q+p)n \leqslant 4qn \leqslant \frac{n}{4}. \tag{2.68}$$

This choice of $\mathcal{X}$ therefore satisfies the bound on the size of the family in Theorem 2.21.

Take an arbitrary $L \subseteq V(\mathcal{H})$ such that $\mathcal{H}[L]$ is not $(p/q, \eta\sigma)$-Janson, and observe that $L \notin E(\mathcal{J})$, by our choice of q as $q + p$, and hence $L \in \mathcal{I}(\mathcal{J})$, by Observation 2.5. Applying item (1) of Theorem 2.26 implies that there is $(S,T) \in \mathcal{Y}$ such that $L \subseteq f(S,T)$, which establishes item (i) of Theorem 2.21. Item (ii) of Theorem 2.21 follows from our choice of $\mathcal{X}$ and item (3) of Theorem 2.26. □

Before finally proceeding to prove Theorem 2.26, we recall the statement of the following classical inequality due to Harris [76].

Lemma 2.27 ([76])**.** *Let $0 \leqslant q \leqslant 1$, and $V \subseteq \mathcal{U}$ be a finite set. If $\mathcal{A} \subseteq 2^{\mathcal{U}}$ is an increasing event and $\mathcal{D} \subseteq 2^{\mathcal{U}}$ is a decreasing event, then*

$$\mathbb{P}(V_q \in \mathcal{A}, V_q \in \mathcal{D}) \leqslant \mathbb{P}(V_q \in \mathcal{A})\mathbb{P}(V_q \in \mathcal{D}).$$

Proof of Theorem 2.26. Apply Theorem 2.23 with $\mathcal{G} = \mathcal{J}$ and parameters q and α to obtain $\mathcal{T}$ and φ. Now, for each $T \in \mathcal{T}$, there is $\mathcal{C}_T$ satisfying item (c) in Theorem 2.23, so we let $\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_{=s}$ be the edges of $\langle \mathcal{C}_T \rangle$ with size s. As $\mathcal{C}'_T$ is s-uniform and $p \leqslant 1/(2^{11}s^2)$ by (2.64), we can apply Theorem 2.15 with $\mathcal{G} = \mathcal{C}'_T$ and $\zeta = 2^{-8}$, obtaining as a result $\mathcal{S}_T$, ψ_T and ϕ_T.

Fix an $I \in \mathcal{I}(\mathcal{J})$ and note that if $T = \varphi(I)$, then it follows from the "moreover" part in item (c) of Theorem 2.23 that $I \in \mathcal{I}(\mathcal{C}_T)$. Combining this with Observation 2.24 and $\mathcal{C}'_T \subseteq \langle \mathcal{C}_T \rangle$, we conclude that $I \in \mathcal{I}(\mathcal{C}'_T)$. As we have applied Theorem 2.15 with $\mathcal{G} = \mathcal{C}'_T$ for every $T \in \mathcal{T}$, and since $I \in \mathcal{I}(\mathcal{C}'_T)$, we obtain, by item (i), sets $S = \phi_T(I) \in \mathcal{S}_T$ and $X = \psi_T(S)$, where $T = \varphi(I)$, such that

$$S \cup T \subseteq I \subseteq X. \tag{2.69}$$

We then define $g(I) = (S,T)$ and $f(S,T) = X$ for $T = \varphi(I)$ and $S = \phi_T(I)$ and set

$$\mathcal{Y} = \{g(I) : I \in \mathcal{I}(\mathcal{J})\},$$

which, by (2.69) and the fact that I was arbitrary, is a definition that satisfies item (1) of Theorem 2.26. Moreover, each $T \in \mathcal{T}$ and $S \in \mathcal{S}_T$ satisfy

$$|T| \leqslant qn/\alpha \qquad \text{and} \qquad |S| \leqslant 2^{11}s^2pn \tag{2.70}$$

by item (b) in Theorem 2.23 and item (ii) in Theorem 2.15 with our choice of $\zeta = 2^{-8}$, which proves item (2). It remains only to show that item (3) holds.

Take an arbitrary $(S,T) \in \mathcal{Y}$ with $X = f(S,T)$ with the goal of showing that $\mathcal{H}[X]$ is not (p,σ)-Janson. To do so, it suffices to show that, for any $\nu : \mathcal{H}[X] \to \mathbb{R}_{\geqslant 0}$, we have

$$\Lambda_p(\nu) \geqslant \frac{e(\nu)^2}{\sigma}, \tag{2.71}$$

so we fix a measure $\nu : \mathcal{H}[X] \to \mathbb{R}_{\geqslant 0}$ and want to establish (2.71).

Recall that we have defined $\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_{=s}$, where $\mathcal{C}_T$ is the hypergraph given by item (c) in Theorem 2.23 for $T \in \mathcal{T}$. Let ν' be the restriction of ν to $\mathcal{H}[X] \cap \mathcal{C}'_T[X]$, i.e. for each $E \in \mathcal{H}[X]$, let

$$\nu'(E) = \begin{cases} \nu(E) & \text{if } E \in \mathcal{C}'_T[X], \\ 0 & \text{otherwise.} \end{cases}$$

Claim 2.28. *If the measure ν' satisfies*

$$e(\nu') \geqslant \frac{e(\nu)}{2}, \tag{2.72}$$

then (2.71) holds.

Proof. Assume that (2.72) holds. We have

$$\Lambda_p(\nu) \geqslant \Lambda_p(\nu') \geqslant \frac{2^8 e(\nu')^2}{p|X|} \geqslant \frac{4e(\nu')^2}{\sigma} \geqslant \frac{e(\nu)^2}{\sigma},$$

first because $\nu' \leqslant \nu$ and $\Lambda_p(\cdot)$ is monotone increasing, second since $\mathcal{C}'_T[X]$ is not $(p, 2^{-8}p|X|)$-Janson by item (iii) of Theorem 2.15 and our choice of $\zeta = 2^{-8}$, then because $\sigma \geqslant 2^{-6}pn$ by (2.64), and the last step is due to (2.72). ∎

We now define the measure $\nu'' = \nu - \nu'$, which corresponds to the restriction of ν to the hypergraph $\mathcal{H}' := \mathcal{H}[X] \setminus \mathcal{C}'_T = \mathcal{H}[X] \setminus \mathcal{C}'_T[X]$. By Claim 2.28, we may assume that

$$e(\nu'') = e(\nu) - e(\nu') > \frac{e(\nu)}{2}, \tag{2.73}$$

otherwise we are done.

With the goal of defining a random measure ν''_q on the hypergraph induced by the random set $X_q = V_q \cap X$, where V_q is a q-random subset of V, we introduce some notation. First, let

$$P_q(E) = \mathbb{P}\big(E \subseteq V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})\big) \tag{2.74}$$

and observe that

$$P_q(E) > (1-\alpha)^{|E|}q^{|E|} \tag{2.75}$$

for all $E \in \mathcal{H}'$, by (2.46), since the hypergraphs $\mathcal{H}'$ and $\mathcal{C}_T$ are disjoint. Indeed, $\mathcal{H}$ is s-uniform, which means that so are $\mathcal{H}'$ and $\mathcal{H}' \cap \mathcal{C}_T$. The only edges of $\mathcal{C}_T$ that could be edges of $\mathcal{H}'$ thus have size exactly equal to s. But every s-sized edge of $\mathcal{C}_T$ is also an edge of $\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_{=s}$, and is therefore not in $\mathcal{H}' = \mathcal{H}[X] \setminus \mathcal{C}'_T$.

Now, let $\nu''_q : \mathcal{H}' \to \mathbb{R}_{\geqslant 0}$ be defined by

$$\nu''_q(E) = \frac{\mathbb{1}\big[E \subseteq V_q\big]}{\mathbb{P}\big(E \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})\big)}\,\nu''(E).$$

We will first show that

$$\mathbb{E}\big[e(\nu''_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})\big] = \sum_{E \in \mathcal{H}'} \nu''(E) = e(\nu''),$$

which will allow us to relate $e(\nu''_q)^2$ to $e(\nu)^2$.

Claim 2.29.

$$\mathbb{E}\big[e(\nu''_q)^2 \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})\big] \geqslant e(\nu'')^2.$$

Proof. The definitions of $e(\nu''_q)$ and ν''_q,

$$e(\nu''_q) = \sum_{E \in \mathcal{H}'} \nu''_q(E) = \sum_{E \in \mathcal{H}'} \frac{\nu''(E)\mathbb{1}\big[E \subseteq V_q\big]}{P_q(E)},$$

imply that

$$\mathbb{E}\big[e(\nu''_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})\big] = \sum_{E \in \mathcal{H}'} \nu''(E) = e(\nu''), \tag{2.76}$$

since, for all $E \in \mathcal{H}'$, we have that

$$\mathbb{E}\big[\mathbb{1}[E \subseteq V_q] \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})\big] = P_q(E)$$

from (2.74), the definition of $P_q(E)$. The claim follows from (2.76) by Jensen's inequality. ∎

The next step is relating $\Lambda_{p/(q-p)}(\nu''_q)$ and $e(\nu''_q)$ when $V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})$.

Claim 2.30. *If $V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})$, then*

$$\Lambda_{p/(q-p)}(\nu''_q) \geqslant \frac{e(\nu''_q)^2}{\eta\sigma}.$$

Proof. It follows from $V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})$ and Observation 2.25 that $V_q \in \mathcal{I}(\mathcal{J})$, and hence $\mathcal{H}[V_q]$ is not $(p/(q-p), \eta\sigma)$-Janson, by the definition of $\mathcal{J}$, (2.65). In particular, it follows that

$$\Lambda_{p/(q-p)}(\nu''_q) \geqslant \frac{e(\nu''_q)^2}{\eta\sigma}$$

as ν''_q is also a measure supported on $\mathcal{H}[V_q]$ by $\mathcal{H}' \subseteq \mathcal{H}$. ∎

Our final inequality bounds $\Lambda_{p/(q-p)}(\nu''_q)$ in expectation by $\Lambda_p(\nu)$, up to an exponential factor.

Claim 2.31.

$$\mathbb{E}\big[\Lambda_{p/(q-p)}(\nu''_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J})\big] < (1-\alpha)^{-2s}\Lambda_p(\nu).$$

Proof. Recall that $d_{\nu''_q}(L)$ is defined as

$$d_{\nu''_q}(L) = \sum_{L\subseteq E\in\mathcal{H}'} \nu''_q(E) = \sum_{L\subseteq E\in\mathcal{H}'} \frac{\nu''(E)\mathbb{1}\big[E\subseteq V_q\big]}{P_q(E)},$$

where the last equality is using the definition of ν''_q, and hence we can write, for every $L\subseteq V$,

$$d_{\nu''_q}(L)^2 = \sum_{L\subseteq E_1\in\mathcal{H}'} \frac{\nu''(E_1)}{P_q(E_1)} \sum_{L\subseteq E_2\in\mathcal{H}'} \frac{\nu''(E_2)}{P_q(E_2)}\,\mathbb{1}\big[E_1\cup E_2\subseteq V_q\big].$$

Observe that the event $V_q\in\mathcal{I}(\underline{\partial}_T\mathcal{J})$ is decreasing and also that, for every $E_1,E_2\in\mathcal{H}'$, the event $E_1\cup E_2\subseteq V_q$ is increasing. We can therefore use Harris' inequality, Lemma 2.27, to bound, for $E_1,E_2\in\mathcal{H}'$,

$$\mathbb{P}\big(E_1\cup E_2\subseteq V_q \mid V_q\in\mathcal{I}(\underline{\partial}_T\mathcal{J})\big) \leqslant \mathbb{P}(E_1\cup E_2\subseteq V_q) = q^{|E_1\cup E_2|} = q^{2s-|E_1\cap E_2|} \tag{2.77}$$

because $\mathcal{H}'$ is s-uniform. Taking the conditional expectation and applying (2.77), we obtain

$$\mathbb{E}\big[d_{\nu''_q}(L)^2 \mid V_q\in\mathcal{I}(\underline{\partial}_T\mathcal{J})\big] < (1-\alpha)^{-2s} \sum_{L\subseteq E_1\in\mathcal{H}'} \nu(E_1) \sum_{L\subseteq E_2\in\mathcal{H}'} \nu(E_2)\, q^{-|E_1\cap E_2|} \tag{2.78}$$

where we used $P_q(E) > (1-\alpha)^s q^s$, by (2.75) and since $\mathcal{H}'$ is s-uniform, and the fact that $\nu''\leqslant\nu$.

By the definition (2.2) of $\Lambda_{p/(q-p)}(\nu''_q)$,

$$\mathbb{E}\big[\Lambda_{p/(q-p)}(\nu''_q) \mid V_q\in\mathcal{I}(\underline{\partial}_T\mathcal{J})\big] = \sum_{\substack{L\subseteq V\\ |L|\geqslant 2}} \mathbb{E}\big[d_{\nu''_q}(L)^2 \mid V_q\in\mathcal{I}(\underline{\partial}_T\mathcal{J})\big] \left(\frac{q-p}{p}\right)^{|L|}, \tag{2.79}$$

and then applying (2.78) to each $d_{\nu''_q}(L)$ term in (2.79) yields

$$\mathbb{E}\big[\Lambda_{p/(q-p)}(\nu''_q) \mid V_q\in\mathcal{I}(\underline{\partial}_T\mathcal{J})\big] < (1-\alpha)^{-2s} \sum_{\substack{L\subseteq V\\ |L|\geqslant 2}} \sum_{L\subseteq E_1\in\mathcal{H}'} \sum_{L\subseteq E_2\in\mathcal{H}'} \frac{\nu(E_1)\nu(E_2)}{q^{|E_1\cap E_2|}} \left(\frac{q-p}{p}\right)^{|L|}$$

or, equivalently,

$$\mathbb{E}\big[\Lambda_{p/(q-p)}(\nu''_q) \mid V_q\in\mathcal{I}(\underline{\partial}_T\mathcal{J})\big] < \sum_{\substack{E_1,E_2\in\mathcal{H}'\\ |E_1\cap E_2|\geqslant 2}} \frac{\nu(E_1)\nu(E_2)}{(1-\alpha)^{2s}q^{|E_1\cap E_2|}} \sum_{\ell=2}^{s} \left(\frac{q-p}{p}\right)^{\ell} \binom{|E_1\cap E_2|}{\ell} \tag{2.80}$$

by first choosing $E_1,E_2\in\mathcal{H}'$ and then $L\subseteq E_1\cap E_2$, grouping terms according to $\ell=|L|$.

Bounding the innermost sum in (2.80) for fixed $E_1, E_2 \in \mathcal{H}'$ with $|E_1 \cap E_2| \geqslant 2$ then yields

$$\sum_{\ell=2}^{s} \left(\frac{q-p}{p}\right)^{\ell} \binom{|E_1 \cap E_2|}{\ell} \leqslant \left(\frac{q}{p}\right)^{|E_1 \cap E_2|},$$

which replaced in (2.80) and simplified, results in

$$\mathbb{E}\big[\Lambda_{p/(q-p)}(\nu_q'') \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})\big] < (1-\alpha)^{-2s} \sum_{\substack{E_1, E_2 \in \mathcal{H}' \\ |E_1 \cap E_2| \geqslant 2}} \frac{\nu(E_1)\nu(E_2)}{p^{|E_1 \cap E_2|}}. \tag{2.81}$$

To complete the proof, note that

$$\sum_{\substack{E_1, E_2 \in \mathcal{H}' \\ |E_1 \cap E_2| \geqslant 2}} \frac{\nu(E_1)\nu(E_2)}{p^{|E_1 \cap E_2|}} \leqslant \sum_{\substack{L \subseteq V \\ |L| \geqslant 2}} \sum_{L \subseteq E_1 \in \mathcal{H}} \sum_{L \subseteq E_2 \in \mathcal{H}} \frac{\nu(E_1)\nu(E_2)}{p^{|L|}} = \sum_{\substack{L \subseteq V \\ |L| \geqslant 2}} d_\nu(L)^2 p^{-|L|},$$

where the last term is equal to $\Lambda_p(\nu)$ by definition, so we obtain, replacing it back in (2.81), the inequality that we wanted. ■

Observe that Claim 2.30 implies that

$$\mathbb{E}\big[\Lambda_{p/(q-p)}(\nu_q'') \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})\big] \geqslant \frac{\mathbb{E}\big[e(\nu_q'')^2 \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})\big]}{\eta\sigma}. \tag{2.82}$$

Combining Claim 2.31 and Claim 2.29 with (2.82) yields

$$(1-\alpha)^{-2s} \Lambda_p(\nu) \geqslant \frac{e(\nu'')^2}{\eta\sigma}$$

and thus, since we are in the case where (2.73) holds, it follows that

$$\Lambda_p(\nu) \geqslant \frac{e(\nu)^2}{\sigma} \tag{2.83}$$

by our choice of η satisfying $4\eta \leqslant (1-\alpha)^{2s}$. As (2.83) was exactly our goal, (2.71), and ν was arbitrary, we conclude that $\mathcal{H}[X]$ is not (p,σ)-Janson. Moreover, our choice of $(S,T) \in \mathcal{Y}$ was also arbitrary, so we have established that item (3) holds, and the proof is complete. □

2.5 Extending collections of copies

In this section, we use a novel container theorem to prove the core statement that we need in the proof of Lemma 2.8. We defer the proof of this container theorem to Section 2.7, since that is the most technical part of the entire argument.

The setting is very similar to Lemma 2.11, but now we will be able to extend many copies of F^- to F by adding a single vertex v to $U = V(\tilde{G}) = V(\tilde{G}')$. Moreover, we will be able to show that the set of copies created by adding v to U will be well-distributed in relation to the copies of F fully contained in U. To obtain this stronger conclusion, we assume that, besides

$\mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}[W]$ being (p,σ)-Janson for every $W \subseteq U = V(\tilde{G})$ with $|W| \geqslant m/(8r)$, we also have that $\mathfrak{I}_{F,\tilde{G}',\tilde{G}}$ is (p,σ')-Janson. Under these circumstances, the lemma states that, when G is distributed as $\mathbb{G}(m+1,1/2)$ conditioned on $\{G[U] = \tilde{G}\}$, the following holds with extremely high probability: for every choice of $G' \subseteq G$ such that $\mathrm{N}_{G'}(v)$ is not too small, the hypergraph $\mathfrak{I}_{F,G',G}$ is $(p,\sigma'+1)$-Janson.

Lemma 2.32. *There exists a constant $C'' > 0$ such that the following holds. Let $m,k,r,s \in \mathbb{N}$ with $r \geqslant 2$, $s < k \leqslant m$, and let*

$$p = \frac{1}{2^{25}k^2r^4}, \qquad m \geqslant r^{C''k}, \qquad \sigma = 2^{-5}r^{-1}pm \qquad \text{and} \qquad 0 \leqslant \sigma' \leqslant \frac{\sigma}{16}. \tag{2.84}$$

Further let F, $\tilde{G}'$ and $\tilde{G}$ be graphs such that $v(F) = s+1 < m$, $\tilde{G}' \subseteq \tilde{G}$ and $v(\tilde{G}) = m$.

If $\mathfrak{I}_{F,\tilde{G}',\tilde{G}}$ is (p,σ')-Janson and $\mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}[W]$ is (p,σ)-Janson for every $W \subseteq U = V(\tilde{G})$ with $|W| \geqslant m/(8r)$, then

$$\mathbb{P}\left(\exists\, G' \subseteq G : \begin{array}{c} G'[U] = \tilde{G}',\ d_{G'}(v) \geqslant m/(4r) \text{ and} \\ \mathfrak{I}_{F,G',G} \text{ is not } (p,\sigma'+1)\text{-Janson} \end{array} \,\middle|\, G[U] = \tilde{G}\right) \leqslant 2^{-m/(32r)}, \tag{2.85}$$

where $G \sim \mathbb{G}(m+1,1/2)$ and $V(G) = U \cup \{v\}$.

The first change that we need to make to the proof in Section 2.3 is to replace Γ_G, defined in (2.18), by Ψ_G, which is just the collection of $G' \subseteq G$ satisfying the event in (2.85):

$$\Psi_G = \left\{ G' \subseteq G : \begin{array}{c} G'[U] = \tilde{G}',\ d_{G'}(v) \geqslant m/(4r) \text{ and} \\ \mathfrak{I}_{F,G',G} \text{ is not } (p,\sigma'+1)\text{-Janson} \end{array} \right\}. \tag{2.86}$$

That is, we now require that $\mathfrak{I}_{F,G',G}$ is not $(p,\sigma'+1)$-Janson, instead of requiring it to be empty as in Lemma 2.11. The argument closely follows the one in Section 2.3, so we briefly summarise the ideas in the proof of Lemma 2.11, referring to some of the definitions in that section as we progress.

Recall that we defined

$$\mathcal{H} = \Big\{ E_L \subseteq U \times \{0,1\} : L \in \mathfrak{I}_{F^-,\tilde{G}',\tilde{G}} \Big\},$$

where

$$E_L = \Big\{ \Big(u, \mathbb{1}\big[\phi_L(u) \in \mathrm{N}_F(w)\big] \Big) : u \in L \Big\},$$

with $\phi_L : L \to V(F^-)$ being a fixed bijection and w being the unique vertex in $V(F) \setminus V(F^-)$. We then defined

$$\iota(G',G) = \big\{(u,0) : u \in \mathrm{N}_G(v)^{\mathsf{c}}\big\} \cup \big\{(u,1) : u \in \mathrm{N}_{G'}(v)\big\}$$

in (2.19), proved Observation 2.13, which states that if both $G[U] = \tilde{G}$ and $G' \in \Gamma_G$, then $\iota(G',G) \in \mathcal{I}(\mathcal{H})$, and applied a container theorem to bound the probability of that event. However, this is not immediately possible here, since Observation 2.13 is not true if we replace $G' \in \Gamma_G$ by $G' \in \Psi_G$, the analogous collection for this section: requiring $\mathfrak{I}_{F,G',G}$ to not be

$(p, \sigma' + 1)$-Janson instead of $\mathfrak{I}_{F,G',G} = \emptyset$ means that we are not interested in independent sets in $\mathcal{H}$, but in vertex subsets I such that $\mathcal{H}[I]$ is not $(p, \sigma' + 1)$-Janson.

We remedy that situation by relying on a container theorem for such sets, like the one that we proved in Section 2.4. However, Theorem 2.21 is not adequate for several reasons, which we discuss while introducing some notation and new definitions. We then prove the analogue of Observation 2.13 for this section, Observation 2.34, and state the container theorem that we end up using, Theorem 2.35.

It will be helpful to partition the copies of $F \subseteq G'[U \cup \{v\}]$ in two natural classes. The first one consists of the copies of F that use v, which correspond to copies of $F^- \subseteq \tilde{G}'[U]$ that are extended with the addition of $v \notin U$. Every other copy of $F \subseteq G'$, i.e. those that do not use v, belong in the second class, and are contained in $\tilde{G}' = G'[U]$. The next definition will relate the hypergraph of copies of F^- that can be extended with v and the hypergraph of the resulting copies of F.

Definition 2.33. For a hypergraph $\mathcal{G}$ and a vertex v not in $V(\mathcal{G})$, let

$$\overline{\partial}_v \mathcal{G} = \big\{E \cup \{v\} : E \in \mathcal{G}\big\}$$

denote the edge-wise inclusion of v in $\mathcal{G}$.

Like in Section 2.3, let $\pi : V(\mathcal{H}) \to U$ be the projection onto the first coordinate and define $\pi_v = \overline{\partial}_v \circ \pi$. Recalling Observation 2.16, that is,

$$\pi(\mathcal{H}) = \mathfrak{I}_{F^-,\tilde{G}',\tilde{G}},$$

we now relate $\mathfrak{I}_{F,G',G}$ to both $\mathfrak{I}_{F,\tilde{G}',\tilde{G}}$ and $\mathcal{H}$ when $G[U] = \tilde{G}$ and $G'[U] = \tilde{G}'$. We will use this fact to conclude that, if $\mathfrak{I}_{F,G',G}$ is not $(p, \sigma'+1)$-Janson, then neither is $\pi_v\big(\mathcal{H}[I]\big) \cup \mathfrak{I}_{F,\tilde{G}',\tilde{G}}$ when $I = \iota(G', G)$.

Observation 2.34. *For all graphs G' and G such that $G[U] = \tilde{G}$, $G' \subseteq G$ and $G'[U] = \tilde{G}'$, if $I = \iota(G', G)$, then*

$$\pi_v\big(\mathcal{H}[I]\big) \cup \mathfrak{I}_{F,\tilde{G}',\tilde{G}} \subseteq \mathfrak{I}_{F,G',G}. \tag{2.87}$$

The inclusion in Observation 2.34 is in fact an equality, but we will not use that fact, and therefore avoid giving its (trivial) proof for the sake of brevity. As we will see, Observation 2.34 follows easily from expanding the definitions, especially after we recall (2.20), that is, if $I = \iota(G', G)$, then

$$I^{(0)} = \mathrm{N}_G(v)^{\mathsf{c}} \qquad \text{and} \qquad I^{(1)} = \mathrm{N}_{G'}(v). \tag{2.88}$$

Proof of Observation 2.34. It follows immediately from

$$G'[U] = \tilde{G}' \qquad \text{and} \qquad G[U] = \tilde{G} \tag{2.89}$$

that $\mathfrak{I}_{F,\tilde{G}',\tilde{G}} \subseteq \mathfrak{I}_{F,G',G}$, so it only remains to show that

$$\pi_v\big(\mathcal{H}[I]\big) \subseteq \mathfrak{I}_{F,G',G}.$$

Let $E \in \mathcal{H}[I]$ be of the form $E = E_L$ for $L \in \mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}$ and recall that $\pi(E) = L$, so our goal is to show that

$$L \cup \{v\} \in \mathfrak{I}_{F,G',G}, \tag{2.90}$$

where $\pi_v(E) = L \cup \{v\}$ by the definition of π_v. As $L \in \mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}$, (2.90) follows from Observation 2.12 using (2.88), (2.89) and the fact that $E \subseteq I$. □

In analogy to the proof in Section 2.3, by Observation 2.34 and the fact that the Janson property is increasing, Observation 2.5, it suffices to have a family of containers $\mathcal{X}$ with the following property. For all $\iota(G', G) = I \subseteq V(\mathcal{H})$ such that $\pi_v(\mathcal{H}[I]) \cup \mathfrak{I}_{F,G',G}$ is not $(p, \sigma' + 1)$-Janson, there is $X \in \mathcal{X}$ with $I \subseteq X$. This is the statement of Theorem 2.35, the container theorem that we need to prove Lemma 2.32. We state that theorem below, but defer its proof, an implementation of the methods discussed in Section 2.4 tailored to this specific setting, to Section 2.7.

Continuing the comparison with Lemma 2.11, ideally each $X \in \mathcal{X}$ would be such that $\pi(\mathcal{H}[X])$ is not (p, σ)-Janson. We are unable to prove such a statement, because the inclusion of v by π_v adds a constraint in the one-degrees of the vertices in $\mathcal{H}$. To deal with this extra constraint, we (roughly) delete a small proportion of vertices to reduce the maximum degree.

Implementing this modification to our method yields something slightly weaker that nonetheless suffices: we show that if $X \in \mathcal{X}$ is sufficiently large, then there is Y covering almost all of X such that $\pi(\mathcal{H}[Y])$ is not (p, σ)-Janson. Our final note before the statement is that, despite applying Theorem 2.35 with the function π being a projection, as we previously defined it in this section, we state the theorem in a slightly more general setting.

Theorem 2.35. *Let $n, r, s \in \mathbb{N}$ with $n \geqslant s$ and $r \geqslant 2$, and let $q, p, \sigma, \sigma', \eta \in \mathbb{R}$ satisfy*

$$0 < q < \frac{1}{8}, \quad 0 < p \leqslant \frac{q}{2^{10} r^2 s^2}, \quad \sigma = 2^{-6} r^{-1} pn, \quad 0 \leqslant \sigma' \leqslant \frac{\sigma}{16} \quad \text{and} \quad \eta = p^4 \Big(\frac{q}{2}\Big)^{4s}. \tag{2.91}$$

Further let $\mathcal{F}$ be a $(s+1)$-uniform hypergraph with vertex set U that is (p, σ')-Janson, let $\mathcal{H}$ be an s-uniform hypergraph with vertex set V, where $|V| = n$, and let $\pi : V \to U$ satisfy

$$|\pi(L)| \geqslant \frac{|L|}{2} \quad \text{for every } L \subseteq V \qquad \text{and} \qquad |\pi(E)| = |E| \quad \text{for every } E \in \mathcal{H}. \tag{2.92}$$

Finally, let v be a vertex not in U. There exists a family $\mathcal{X} \subseteq 2^V$ with

$$|\mathcal{X}| \leqslant \left(\frac{2}{q}\right)^{2qn} \tag{2.93}$$

such that the following hold.

(1) *If $I \subseteq V$ and $\pi_v(\mathcal{H}[I]) \cup \mathcal{F}$ is not $(p, \sigma' + \eta\sigma)$-Janson, then $I \subseteq X$ for some $X \in \mathcal{X}$.*

(2) *For each $X \in \mathcal{X}$ with $|X| \geqslant n/(8r)$, there exists $Y \subseteq X$ with*

$$|Y| \geqslant |X| - 2^{-8} r^{-1} n \tag{2.94}$$

such that $\pi(\mathcal{H}[Y])$ is not (p, σ)-Janson.

We are now ready to prove Lemma 2.32.

Proof of Lemma 2.32. The first step is applying Theorem 2.35 with $q = 2^{-15}r^{-2}$ and $\mathcal{F} = \mathfrak{I}_{F,\tilde{G}',\tilde{G}}$, so we must check that these choices satisfy the assumptions of the theorem.

We assumed that $\mathcal{F} = \mathfrak{I}_{F,\tilde{G}',\tilde{G}}$ is (p,σ')-Janson and $(s+1)$-uniform, and $\mathcal{H}$ being s-uniform follows from its definition and the fact that $v(F^-) = v(F) - 1$. To apply Theorem 2.35 with this choice of $\mathcal{H}$, we implicitly set

$$V = U \times \{0,1\}, \qquad n = 2m \qquad \text{and} \qquad \sigma = 2^{-5}r^{-1}pm = 2^{-6}r^{-1}pn$$

and therefore the value of σ coincides in both statements, resulting in the condition $0 \leqslant \sigma' \leqslant \sigma/16$ also being satisfied by the identical assumption in Lemma 2.32. Furthermore, our choices for the parameters $0 < q = 2^{-15}r^{-2} < 1/8$ and $p > 0$ satisfy

$$p = \frac{1}{2^{25}k^2r^4} < \frac{q}{2^{10}s^2r^2}$$

because $s < k$, and we have checked that all the conditions in (2.91) hold.

We now check that π satisfies (2.92). The first assumption follows trivially from the choice of $V = U \times \{0,1\}$ and $\pi : V \to U$ being a projection into the first coordinate, while the other requirement is Observation 2.18. This concludes the checking of the assumptions and requirements in Theorem 2.35.

Applying Theorem 2.35, we obtain a family $\mathcal{X}$ satisfying (2.93) and items (1) and (2) in its statement. The next claim, a simple combination of Observation 2.34, the definition of Ψ_G and the choice of q, shows that item (1) holds for $I = \iota(G',G)$ when $G' \in \Psi_G$, where, recall,

$$\iota(G',G) = \big\{(u,0) : u \in \mathrm{N}_G(v)^{\mathsf{c}}\big\} \cup \big\{(u,1) : u \in \mathrm{N}_{G'}(v)\big\}.$$

Claim 2.36. *Let G' and G be graphs and let $I = \iota(G',G)$. If $G[U] = \tilde{G}$ and $G' \in \Psi_G$, then the hypergraph $\pi_v\big(\mathcal{H}[I]\big) \cup \mathcal{F}$ is not $(p,\sigma'+\eta\sigma)$-Janson.*

Proof. Recall that the definition of Ψ_G, (2.86), implies that $G'[U] = \tilde{G}'$, so we can apply Observation 2.34 to conclude that

$$\pi_v\big(\mathcal{H}[I]\big) \cup \mathcal{F} \subseteq \mathfrak{I}_{F,G',G}, \tag{2.95}$$

where we replaced $\mathcal{F} = \mathfrak{I}_{F,\tilde{G}',\tilde{G}}$ in (2.87).

It also follows from $G' \in \Psi_G$ that $\mathfrak{I}_{F,G',G}$ is not $(p,\sigma'+1)$-Janson, so we can combine (2.95) with the fact that being Janson is increasing, Observation 2.5, to deduce that $\pi_v\big(\mathcal{H}[I]\big) \cup \mathcal{F}$ is also not $(p,\sigma'+1)$-Janson. But now, as we chose

$$\eta = p^4\Big(\frac{q}{2}\Big)^{4s}, \qquad \sigma = 2^{-5}r^{-1}pm, \qquad q = 2^{-15}r^{-2} \qquad \text{and} \qquad p = 2^{-25}k^{-2}r^{-4},$$

one can verify that

$$\eta\sigma = 2^{-4s-5}r^{-1}p^5q^{4s}m \geqslant 1$$

by $s \leqslant k$ and $m \geqslant r^{C''k}$ if we take $C'' \geqslant 300$. We therefore conclude that $\pi_v\big(\mathcal{H}[I]\big) \cup \mathcal{F}$ is not $(p, \sigma' + \eta\sigma)$-Janson, because being (p,σ)-Janson is decreasing in σ by Observation 2.3. ■

By item (1) in Theorem 2.35 and Claim 2.36, for all graphs G' and G such that $G[U] = \tilde{G}$ and $G' \in \Psi_G$, there exists $X \in \mathcal{X}$ such that $\iota(G', G) \subseteq X$. Let

$$\mathcal{X}' = \big\{X \in \mathcal{X} : |X^{(1)}| \geqslant m/(4r)\big\},$$

and we will show that there is also $X \in \mathcal{X}'$ such that $\iota(G', G) \subseteq X$.

Claim 2.37. *If $G[U] = \tilde{G}$ and $G' \in \Psi_G$, then there exists $X \in \mathcal{X}'$ such that $\iota(G', G) \subseteq X$.*

Proof. Fix G with $G[U] = \tilde{G}$ and $G' \in \Psi_G$, and let $X \in \mathcal{X}$ be such that $I = \iota(G', G) \subseteq X$. By the definition of ι, we have

$$\mathrm{N}_{G'}(v) = I^{(1)} \subseteq X^{(1)} \qquad \text{and} \qquad \mathrm{N}_G(v)^{\mathsf{c}} = I^{(0)} \subseteq X^{(0)},$$

and therefore

$$|X^{(1)}| \geqslant d_{G'}(v) \geqslant \frac{m}{4r} \tag{2.96}$$

where the last inequality is due to $G' \in \Psi_G$. We conclude that $X \in \mathcal{X}'$. ■

Taking a union bound over choices of $\mathcal{X}'$, we can bound the probability in (2.85) from above by

$$\mathbb{P}\big(\exists G' \subseteq G : G' \in \Psi_G \mid G[U] = \tilde{G}\big) \leqslant \sum_{X \in \mathcal{X}'} \mathbb{P}\big(\mathrm{N}_G(v)^{\mathsf{c}} \subseteq X^{(0)}\big) \tag{2.97}$$

using Claim 2.37 and replacing the event $\{\iota(G', G) \subseteq X\}$ by $\big\{\mathrm{N}_G(v)^{\mathsf{c}} \subseteq X^{(0)}\big\}$, which it implies by (2.88). We now want an upper bound for the probability of this event for each $X \in \mathcal{X}'$.

Claim 2.38. *For every $X \in \mathcal{X}'$, we have*

$$|U \setminus X^{(0)}| \geqslant \frac{m}{16r}. \tag{2.98}$$

Proof. By item (2) in Theorem 2.35, there is $Y \subseteq X$ with

$$|Y| \geqslant |X| - \frac{n}{2^8 r} = |X^{(0)}| + |X^{(1)}| - \frac{n}{2^8 r} \geqslant |X^{(0)}| + \Big(\frac{1}{4r} - \frac{1}{2^7 r}\Big)m \tag{2.99}$$

such that the hypergraph $\pi(\mathcal{H}[Y])$ is not (p,σ)-Janson, where we used (2.96) and $n = 2m$.

Taking $W = Y^{(0)} \cap Y^{(1)}$, observe that $\mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}[W]$ is not (p,σ)-Janson. To check that, recall that the hypergraph $\pi(\mathcal{H}[Y])$ is not (p,σ)-Janson by item (2) in Theorem 2.35. It then follows from

$$\mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}[W] = \pi(\mathcal{H})[W] \subseteq \pi\big(\mathcal{H}[Y]\big)$$

by Observation 2.16 and Lemma 2.19 and the fact that being (p,σ)-Janson is increasing, Observation 2.5, that $\mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}[W]$ cannot be (p,σ)-Janson.

Since we have assumed in the statement of Lemma 2.32 that $\mathfrak{I}_{F^-,\tilde{G}',\tilde{G}}[W]$ is (p,σ)-Janson whenever $W \subseteq U$ satisfies $|W| \geqslant m/(8r)$, the fact that choosing $W = Y^{(0)} \cap Y^{(1)}$ results in a

subhypergraph that is not (p,σ)-Janson implies that

$$|Y^{(0)} \cap Y^{(1)}| < \frac{m}{8r}. \tag{2.100}$$

Manipulating (2.100), we obtain

$$|Y| = |Y^{(0)} \cup Y^{(1)}| + |Y^{(0)} \cap Y^{(1)}| < \Big(1 + \frac{1}{8r}\Big)|U|,$$

which combined with (2.99) yields

$$|X^{(0)}| < \Big(1 - \frac{1}{16r}\Big)m,$$

and hence (2.98), as desired. ■

Applying Claim 2.38 to each term in (2.97), we obtain

$$\mathbb{P}\big(\mathrm{N}_G(v)^{\mathsf{c}} \subseteq X^{(0)}\big) = \mathbb{P}\Big(\mathrm{N}_G(v)^{\mathsf{c}} \cap \big(U \setminus X^{(0)}\big) = \emptyset\Big) = 2^{-|U\setminus X^{(0)}|} \leqslant 2^{-m/(16r)} \tag{2.101}$$

for each $X \in \mathcal{X}'$, using that $G \sim \mathbb{G}(m+1,1/2)$. Replacing (2.101) back in (2.97) yields

$$\sum_{X\in\mathcal{X}'} \mathbb{P}\big(\mathrm{N}_G(v)^{\mathsf{c}} \subseteq X^{(0)}\big) \leqslant |\mathcal{X}'|\, 2^{-m/(16r)}.$$

Now, we can bound the size of $\mathcal{X}'$ using (2.93) and $2qn \leqslant m/(2^{13}r^2)$, where the latter holds by $n = 2m$ and our choice of $q = 2^{-15}r^{-2}$, to obtain

$$\mathbb{P}\big(\exists G' \subseteq G : G' \in \Psi_G \mid G[U] = \tilde{G}\big) \leqslant \left(\frac{2}{q}\right)^{2qn} 2^{-m/(16r)} \leqslant 2^{m/(2^9 r) - m/(16r)} \leqslant 2^{-m/(32r)}$$

since $r \geqslant 2$. □

2.6 Proof of Lemma 2.8

The purpose of this section is to give a proof of Lemma 2.8, restated below. We begin with an intuitive and informal overview of the proof, with the purpose of motivating the intermediate results in the section, and then introduce the details and technicalities in the sections that follow.

Lemma 2.8. *There exists an absolute constant $C' > 0$ such that the following holds. For all $k, r \in \mathbb{N}$ with $r \geqslant 2$ and for all graphs $H_1, \dots, H_r$ with at most k vertices, if $N \in \mathbb{N}$ satisfies*

$$N \geqslant r^{C'(k+t)} \tag{2.102}$$

for $t = \sum_{i=1}^r v(H_i)$, then

$$\mathbb{P}\big(G \in \mathcal{B}(\mathbf{H};p) \cap \mathcal{E}(\mathbf{s};p)\big) \leqslant 2^{-\delta^2 N^2},$$

where $G \sim \mathbb{G}(N,1/2)$, $p = 1/(2^{25}k^2r^4)$, $\delta = r^{-50}$, $\mathbf{H} = (H_i)_{i\in[r]}$ and $\mathbf{s} = \big(v(H_i)\big)_{i\in[r]}$.

Our informal overview of the proof of Lemma 2.8 starts with a statement of our setting and strategy, where throughout this section, we fix $\delta = r^{-50}$ and $p = 1/(2^{25}k^2r^4)$. We assume that $G \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})$, i.e. G admits a "bad" colouring $\chi : E(G) \to [r]$ in which the copies of $H_i \subseteq G_i$ that are induced in G are not (p, pN)-Janson, even though G satisfies the inductive assumption, represented here by the event $\mathcal{E}(\mathbf{s})$. Our goal is to show that such G are extremely rare when $G \sim \mathbb{G}(N, 1/2)$, which we accomplish by applying Lemma 2.32 with $\tilde{G} = G[U]$ and $\tilde{G}' = G_\ell^{(\chi)}[U]$ for a certain vertex subset U and a specific colour $\ell \in [r]$. Proving the existence of this set U is not difficult, and is the main purpose of the intermediate results in this section.

To reach a point where we can apply Lemma 2.32, that is, to show that this set U exists, we combine Lemma 2.43 and Lemma 2.44. The proof of the former lemma uses the induction hypothesis to conclude that $\mathfrak{I}_{H_i^-, G_i^{(\chi)}, G}[W]$ is $(p, p|W|)$-Janson for all "bad" colourings χ and all large $W \subseteq U$.

In the proof Lemma 2.44, we (roughly) construct a set $U \subseteq V(G)$ vertex-by-vertex, starting from an arbitrary vertex subset of size δN. Adding a vertex v to U increments the Janson parameter of $H_i \subseteq G_i[U]$ for some colour i, in the sense that if $\mathfrak{I}_{H_i,G_i,G}[U]$ was (p, σ_i)-Janson for some $\sigma_i \geqslant 0$, then $\mathfrak{I}_{H_i,G_i,G}[U \cup \{v\}]$ is $(p, \sigma_i + 1)$-Janson. The set U is complete when there are no more vertices whose addition to U would increase the Janson parameter of H_i for some $i \in [r]$, and we show that the final size of U is at most $2\delta N$.

These are the two preliminaries that we require before the proof of Lemma 2.8, which we briefly and informally discuss now. With U given by Lemma 2.43 and Lemma 2.44, we will apply Lemma 2.32 to the graphs $\tilde{G} = G[U]$ and $\tilde{G}' = G_\ell^{(\chi)}[U]$ for a certain colour ℓ and many vertices $v \notin U$ when the colouring χ is bad, relying on the independence of these events for each v to obtain the required bound on their joint probability. These vertices $v \notin U$ are chosen first to ensure that their degree to U is not small, so one colour $i_v \in [r]$ also has sufficiently many neighbours in U. We then select ℓ as the majority colour among the i_v, and further restrict to those v for which $i_v = \ell$.

To formally implement this outline, we first address a technicality: we are not able to directly prove that the final collection of copies of H_ℓ is (p, pN)-Janson, only $(p, 2^{-9}r^{-1}\delta pN)$-Janson[1]. As we will see, this is not a problem, because we show in Lemma 2.40, using double counting, that a hypergraph $\mathcal{G}$ that is not (p, pN)-Janson contains a subset S of $\delta^{2/3}N$ vertices such that $\mathcal{G}[S]$ is not $(p, 2^{-9}r^{-1}\delta pN)$-Janson, and we will be able to work entirely inside this set S. The next section proves Lemma 2.40, and the rest of the section implements the above outline to show that U exists, and finally to prove Lemma 2.8 in Section 2.6.3.

2.6.1 Changing to another bad event

First, recall the definition of the bad event $\mathcal{B}(\mathbf{H})$.

Definition 2.7. Given $p > 0$ and a collection of graphs $H_1, \dots, H_r$, let $\mathbf{H} = (H_i)_{i \in [r]}$ and let $\mathcal{B}(\mathbf{H}) = \mathcal{B}(\mathbf{H}; p)$ be the family of graphs G with the following property. There exists a colouring $\chi : E(G) \to [r]$ such that, for every $i \in [r]$, the hypergraph $\mathfrak{I}_{H_i,G_i,G}$ is not $(p, p\,v(G))$-Janson.

[1] This is because we can only apply Lemma 2.32 for $\sigma' \leqslant \sigma/16$, where the collection of copies of H_ℓ^- contained in W is (p, σ)-Janson for every $W \subseteq U$ with $|W| \geqslant |U|/(8r)$, and the event $\mathcal{E}(\mathbf{s})$ only tells us that this hypergraph is $(p, p|W|)$-Janson.

As previously mentioned, the first stage in the proof is to show that we can replace $\mathcal{B}(\mathbf{H})$ by an alternative event $\mathcal{B}'(\mathbf{H})$. The main advantage of this change is that it reduces the Janson parameter by a factor of order $r^{-1}\delta$, at the cost of assuming that it only holds for a single subset S of size $|S| \geqslant \delta^{2/3}N$.

Definition 2.39. Given a collection of graphs $H_1, \dots, H_r$, let $\mathbf{H} = (H_i)_{i\in[r]}$ and let $\mathcal{B}'(\mathbf{H})$ be the family of graphs G with the following property. There exists $S \subseteq V(G)$ with $|S| \geqslant \delta^{2/3}v(G)$ and a colouring $\chi : E(G[S]) \to [r]$ such that $\mathfrak{I}_{H_i,G_i,G}[S]$ is not $(p, 2^{-9}r^{-1}\delta p\, v(G))$-Janson for all $i \in [r]$.

It is crucial that the original bad event $\mathcal{B}(\mathbf{H})$ is contained in the variant $\mathcal{B}'(\mathbf{H})$, which we now prove with a simple double counting argument.

Lemma 2.40. *For every $k \in \mathbb{N}$ and graphs $H_1, \dots, H_r$,*

$$\mathcal{B}(\mathbf{H}) \subseteq \mathcal{B}'(\mathbf{H}),$$

where $\mathbf{H} = (H_i)_{i\in[r]}$.

As we need to work with measures in the Janson property, the following observation, albeit a trivial consequence of the definitions, will be useful when proving Lemma 2.40.

Observation 2.41. *For all $s \in \mathbb{N}$ and hypergraphs $\mathcal{G}$, if $\vartheta_1, \dots, \vartheta_s : \mathcal{G} \to \mathbb{R}_{\geqslant 0}$ satisfy*

$$\vartheta = \sum_{i=1}^{s} \vartheta_i,$$

then,

$$e(\vartheta) = \sum_{i=1}^{s} e(\vartheta_i) \qquad \text{and} \qquad d_\vartheta(L) = \sum_{i=1}^{s} d_{\vartheta_i}(L)$$

for all $L \subseteq V(\mathcal{G})$.

We can now prove Lemma 2.40.

Proof of Lemma 2.40. Assume that $G \notin \mathcal{B}'(\mathbf{H})$ and let $V = V(G)$ with $N = |V|$. Let $\chi : E(G) \to [r]$ be an arbitrary r-colouring of the edges of G, and observe that, by Definition 2.39, for every $S \subseteq V$ with $|S| \geqslant \delta^{2/3}N$ there exists $j \in [r]$ such that $\mathfrak{I}_{H_j,G_j,G}[S]$ is $(p, 2^{-9}r^{-1}\delta pN)$-Janson. Thus, if for each $j \in [r]$, we define

$$\mathbf{V}_j = \left\{ S \subseteq V : |S| = \delta^{2/3}N \text{ and } \mathfrak{I}_{H_j,G_j,G}[S] \text{ is } (p, 2^{-9}r^{-1}\delta pN)\text{-Janson} \right\}$$

then

$$\bigcup_{j=1}^{r} \mathbf{V}_j = \binom{V}{\delta^{2/3}N}$$

and hence there exists $i \in [r]$ such that

$$|\mathbf{V}_i| \geqslant \frac{1}{r}\binom{N}{\delta^{2/3}N}, \tag{2.103}$$

so fix such an i. For each $S \in \mathbf{V}_i$, as $\mathfrak{J}_{H_i,G_i,G}[S]$ is $(p, 2^{-9}r^{-1}\delta pN)$-Janson by assumption, there exists $\nu_S : \mathfrak{J}_{H_i,G_i,G}[S] \to \mathbb{R}_{\geqslant 0}$ satisfying

$$e(\nu_S) = 1 \qquad \text{and} \qquad \Lambda_p(\nu_S) < \frac{2^9 r}{\delta pN} \tag{2.104}$$

by Observation 2.4. Now, define $\nu : \mathfrak{J}_{H_i,G_i,G} \to \mathbb{R}_{\geqslant 0}$ by

$$\nu = \sum_{S \in \mathbf{V}_i} \nu_S$$

and note that if

$$\Lambda_p(\nu) < \frac{e(\nu)^2}{pN}, \tag{2.105}$$

then $\mathfrak{J}_{H_i,G_i,G}$ is (p, pN)-Janson and therefore $G \notin \mathcal{B}(\mathbf{H})$ by definition, since χ is an arbitrary r-colouring of the edges of G.

Recall that $V = V(\mathfrak{J}_{H_i,G_i,G})$ and also that, by definition,

$$\Lambda_p(\nu) = \sum_{\substack{L \subseteq V \\ |L| \geqslant 2}} d_\nu(L)^2 p^{-|L|} = \sum_{\substack{L \subseteq V \\ |L| \geqslant 2}} \Big(\sum_{S \in \mathbf{V}_i} d_{\nu_S}(L) \Big)^2 p^{-|L|}$$

where the last equality is due to Observation 2.41. Further observe that since ν_S is supported only on $\mathfrak{J}_{H_i,G_i,G}[S]$, then we can only have $d_{\nu_S}(L) > 0$ if $L \subseteq S$. Hence, denoting the subfamily

$$\mathbf{T}_L = \{S \in \mathbf{V}_i : L \subseteq S\}$$

for every $L \subseteq V$, we have

$$\Lambda_p(\nu) = \sum_{\substack{L \subseteq V \\ |L| \geqslant 2}} \Big(\sum_{S \in \mathbf{T}_L} d_{\nu_S}(L) \Big)^2 p^{-|L|} \leqslant \sum_{\substack{L \subseteq V \\ |L| \geqslant 2}} |\mathbf{T}_L| \sum_{S \in \mathbf{T}_L} d_{\nu_S}(L)^2 p^{-|L|}, \tag{2.106}$$

where the last step holds by the Cauchy–Schwarz inequality. Now, note that, as every L in the sums of (2.106) satisfies $|L| \geqslant 2$, we can bound $|\mathbf{T}_L|$ for these L by

$$|\mathbf{T}_L| \leqslant \binom{N-2}{\delta^{2/3}N - 2} \leqslant \delta^{4/3} \binom{N}{\delta^{2/3}N} \leqslant r\delta^{4/3}|\mathbf{V}_i|,$$

where we used (2.103) in the last inequality. Substituting this into (2.106) and using (2.104) together with the definition of $\Lambda_p(\nu)$ yields

$$\Lambda_p(\nu) \leqslant r\delta^{4/3}|\mathbf{V}_i| \sum_{S \in \mathbf{V}_i} \Lambda_p(\nu_S) < r\delta^{4/3}|\mathbf{V}_i|^2 \frac{2^9 r}{\delta pN} = 2^9 r^2 \delta^{1/3} \frac{|\mathbf{V}_i|^2}{pN}. \tag{2.107}$$

Note that, since $e(\nu_S) = 1$ for all $S \in \mathbf{V}_i$ by (2.104), it follows from Observation 2.41 that

$$e(\nu) = |\mathbf{V}_i|,$$

which, replaced in (2.107), results in our goal, (2.105),

$$\Lambda_p(\nu) < 2^9 r^2 \delta^{1/3} \frac{e(\nu)^2}{pN} < \frac{e(\nu)^2}{pN}$$

where the last inequality is due to our choice of $\delta = r^{-50}$. Therefore, the hypergraph $\mathfrak{I}_{H_i,G_i,G}$ is (p, pN)-Janson, which implies that $G \notin \mathcal{B}(\mathbf{H})$ since the colouring $\chi : E(G) \to [r]$ was arbitrary. □

2.6.2 Finding a set U to apply Lemma 2.32

Using Lemma 2.40, we can bound

$$\mathbb{P}\big(G \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})\big) \leqslant \mathbb{P}\big(G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})\big)$$

where $G \sim \mathbb{G}(N, 1/2)$, which leads us to the second stage in the proof of Lemma 2.8. In it, we will apply Lemma 2.32 to bound $\mathbb{P}\big(G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})\big)$, but the setup of this application requires some work. First we show, with a deterministic argument, that for any graph $G \in \mathcal{B}'(\mathbf{H})$ with a "bad" colouring χ, we can find a set U that satisfies the requirements of Lemma 2.32.

Before describing the concrete properties of U, we establish a correspondence between $G \in \mathcal{B}'(\mathbf{H})$, the sets S that appear in Definition 2.39, and these bad colourings χ. To do that, it is helpful to define the common setting for the rest of this section. Fix then $k \in \mathbb{N}$ and $s_1, \dots, s_r \in \mathbb{N}$ such that $s_i \leqslant k$ for each $i \in [r]$. Further fix graphs $H_1, \dots, H_r$ such that $v(H_i) \leqslant s_i$ for all $i \in [r]$, let $\mathbf{s} = (s_i)_{i\in[r]}$ and $\mathbf{H} = (H_i)_{i\in[r]}$, and fix $N \in \mathbb{N}$ satisfying (2.102).

Definition 2.42. For a graph G, define the collection $\mathbf{S}(G)$ by

$$\mathbf{S}(G) = \left\{ (S, \chi) : \begin{array}{c} S \subseteq V(G) \text{ with } |S| \geqslant \delta^{2/3} N \text{ and } \chi : E(G[S]) \to [r] \\ \text{such that } \forall i \in [r],\ \mathfrak{I}_{H_i,G_i,G}[S] \text{ is not } (p, 2^{-9} r^{-1} \delta p N)\text{-Janson} \end{array} \right\}.$$

Recall from Definition 2.39 that if $G \in \mathcal{B}'(\mathbf{H})$, then there are $S \subseteq V(G)$ and $\chi : E(G[S]) \to [r]$ such that $(S, \chi) \in \mathbf{S}(G)$. We will next show that if $(S, \chi) \in \mathbf{S}(G)$, then for every large subset $W \subseteq S$, the hypergraph $\mathfrak{I}_{H_i^-,G_i,G}[W]$ is $(p, p|W|)$-Janson. To prove that, we will require the event $\mathcal{E}(\mathbf{s})$, whose definition we recall for the reader's convenience.

Definition 2.6. Given $p > 0$, $r \in \mathbb{N}$ and $\mathbf{s} = (s_i)_{i\in[r]} \in \mathbb{N}^r$, let $\mathcal{E}(\mathbf{s}) = \mathcal{E}(\mathbf{s}; p)$ be the family of graphs G with the following property. Fix $\delta = r^{-50}$. For all graphs $F_1, \dots, F_r$ satisfying

$$\sum_{i=1}^{r} v(F_i) = \sum_{i=1}^{r} s_i - 1 \qquad \text{and} \qquad v(F_i) \leqslant s_i \ \text{ for each } i \in [r],$$

for every $W \subseteq V(G)$ with

$$|W| \geqslant \frac{\delta}{8r}\, v(G),$$

and every colouring $\chi : E(G[W]) \to [r]$, there is $i \in [r]$ such that $\mathfrak{I}_{F_i,G_i,G}[W]$ is $(p, p|W|)$-Janson.

The following lemma is the only place in the proof of Lemma 2.8 where we will use the event

$\mathcal{E}(\mathbf{s})$. Since later we will choose the set U to be a subset of S, the lemma immediately implies that $\mathfrak{I}_{H_i^-,G_i,G}[W]$ is $(p,p|W|)$-Janson for every large subset $W \subseteq U$.

Lemma 2.43. *For all $G \in \mathcal{E}(\mathbf{s})$ and $(S,\chi) \in \mathbf{S}(G)$, the hypergraph $\mathfrak{I}_{H_i^-,G_i,G}[W]$ is $(p,p|W|)$-Janson for every $i \in [r]$ and every $W \subseteq S$ with $|W| \geqslant \delta N/(8r)$.*

Proof. Fix an arbitrary $i \in [r]$, let $F_j = H_j$ for all $j \in [r] \setminus \{i\}$, and take $F_i = H_i^-$, a choice that satisfies

$$\sum_{j=1}^{r} v(F_j) = \sum_{j=1}^{r} s_j - 1. \tag{2.108}$$

By $G \in \mathcal{E}(\mathbf{s})$ and (2.108), for every $W \subseteq S \subseteq V(G)$ with $|W| \geqslant \delta N/(8r)$, there exists a colour $\ell \in [r]$ such that $\mathfrak{I}_{F_\ell,G_\ell,G}[W]$ is $(p,p|W|)$-Janson. However, as

$$p|W| \geqslant 2^{-9} r^{-1} \delta p N,$$

we conclude that $\mathfrak{I}_{F_\ell,G_\ell,G}[W]$ is $(p,2^{-9}r^{-1}\delta pN)$-Janson. Since the Janson property is increasing by Observation 2.5, it follows that $\mathfrak{I}_{F_\ell,G_\ell,G}[S]$ is also $(p,2^{-9}r^{-1}\delta pN)$-Janson.

Now, recall that $\mathfrak{I}_{H_j,G_j,G}[S]$ is not $(p,2^{-9}r^{-1}\delta pN)$-Janson for all $j \in [r]$ since $(S,\chi) \in \mathbf{S}(G)$. Combining our choice of $F_j = H_j$ for all $j \neq i$ with the fact that $\mathfrak{I}_{F_\ell,G_\ell,G}[S]$ is $(p,2^{-9}r^{-1}\delta pN)$-Janson, the only remaining possibility is that $\ell = i$. It follows that, for every $W \subseteq S$ satisfying $|W| \geqslant \delta N/(8r)$, the hypergraph $\mathfrak{I}_{H_i^-,G_i,G}[W]$ is $(p,p|W|)$-Janson. Since i was arbitrary, this completes the proof of the lemma. □

The remaining properties that U will have are all guaranteed simultaneously by the way we construct it. For each $i \in [r]$, the hypergraph $\mathfrak{I}_{H_i,G_i,G}[U]$ will be (p,σ_i)-Janson for some $\sigma_i \geqslant 0$, and $\mathfrak{I}_{H_i,G_i,G}[U \cup \{v\}]$ will not be (p,σ_i+1)-Janson for any $v \notin U$. The proof of Lemma 2.44 shows that we can find such a U with a routine induction.

Lemma 2.44. *Let G be a graph. If $(S,\chi) \in \mathbf{S}(G)$, then there exist $U \subseteq S$ and $\sigma_1,\dots,\sigma_r \in \mathbb{Z}_{\geqslant 0}$ such that*

$$|U| = \delta N + \sum_{i=1}^{r} \sigma_i \qquad \text{and} \qquad \sigma_1,\dots,\sigma_r \leqslant \frac{p|U|}{2^9 r} \tag{2.109}$$

which further satisfy, for all $i \in [r]$,

(a) *$\mathfrak{I}_{H_i,G_i,G}[U]$ is (p,σ_i)-Janson, and*

(b) *$\mathfrak{I}_{H_i,G_i,G}\big[U \cup \{v\}\big]$ is not (p,σ_i+1)-Janson for all $v \in S \setminus U$.*

Proof. Observe first that every set $U \subseteq S$ of size δN satisfies item (a) with $\sigma_i = 0$ for all $i \in [r]$, since $\mathfrak{I}_{H_i,G_i,G}[U]$ is $(p,0)$-Janson by definition. Now take U of size $\delta N + \sum_{i=1}^r \sigma_i$ maximising $\sum_{i=1}^r \sigma_i$ among the choices that satisfy item (a). Note that $\mathfrak{I}_{H_i,G_i,G}[U \cup \{v\}]$ is (p,σ_i)-Janson for every $v \in S$, since $\mathfrak{I}_{H_i,G_i,G}[U]$ is (p,σ_i)-Janson, and by Observation 2.5, so item (b) holds by the maximality of $\sum_{i=1}^r \sigma_i$. It therefore remains to prove the inequality in (2.109).

Suppose for a contradiction that $\sigma_i > p|U|/(2^9 r)$ for some $i \in [r]$ and observe that

$$\sigma_i > \frac{p|U|}{2^9 r} \geqslant \frac{\delta p N}{2^9 r},$$

so $\mathfrak{I}_{H_i,G_i,G}[U]$ is $(p, 2^{-9}r^{-1}\delta pN)$-Janson by Observation 2.3. Since $U \subseteq S$, Observation 2.5 implies that $\mathfrak{I}_{H_i,G_i,G}[S]$ is also $(p, 2^{-9}r^{-1}\delta pN)$-Janson, contradicting $(S,\chi) \in \mathbf{S}(G)$. □

2.6.3 The proof

Having established that there is a U satisfying most of the requirements of Lemma 2.32, we will combine the following trivial consequence of the Chernoff bound with a pigeonhole argument to find many vertices v whose degree to U in some colour $\ell \in [r]$ is not small.

Observation 2.45. *Let $G \sim \mathbb{G}(N,1/2)$, let $U \subseteq S \subseteq V(G)$ and set*

$$S' = \big\{v \in S \setminus U : d_G(v,U) > |U|/4\big\}.$$

If $2^{10} \leqslant |U| \leqslant |S|/4$, then

$$\mathbb{P}\big(|S'| \leqslant |S|/4\big) \leqslant \exp\big(-2^{-6}|U||S|\big).$$

Proof. For a vertex $v \in S \setminus U$, the Chernoff bound implies that

$$\mathbb{P}(v \notin S') = \mathbb{P}\big(d_G(v,U) \leqslant |U|/4\big) \leqslant \exp\big(-|U|/16\big),$$

and observe that these events are independent for distinct vertices of $S \setminus U$. If $|S'| \leqslant |S|/4$, then at least $|S|/2$ vertices of $S \setminus U$ fail to be in S'. Taking a union bound then yields

$$\mathbb{P}\big(|S'| \leqslant |S|/4\big) \leqslant 2^{|S|}\exp\left(-\frac{|S|}{2}\frac{|U|}{16}\right) \leqslant \exp\big(-2^{-6}|U||S|\big),$$

where last inequality uses that $|U| \geqslant 2^{10}$. □

With all of the above, the proof of Lemma 2.8 proceeds by fixing the sets S, U and the colouring of $G[U]$, all of which we will take union bounds over. We will use Observation 2.45 to find many vertices whose degree to U in colour ℓ is not small, and we will then be able to check that the graphs $\tilde{G}' = G_\ell^{(\chi)}[U]$ and $\tilde{G} = G[U]$ satisfy all the conditions required by Lemma 2.32. To complete the proof, it will then suffice to observe that for each v the events bounded by Lemma 2.32 are independent, and hence the probabilities multiply.

Proof of Lemma 2.8. Recall first that we fixed $p = \big(2^{25}k^2r^4\big)^{-1}$ and $\delta = r^{-50}$ in Lemma 2.8. We claim that if $v(H_i) = 1$ for some $i \in [r]$, then trivially we have $\mathcal{B}(\mathbf{H}) = \emptyset$. Indeed, every non-empty 1-uniform hypergraph is (p,σ)-Janson for all $p > 0$ and σ, since $\Lambda_p(\nu) = 0$ regardless of the measure ν. We may therefore assume that $v(H_i) \geqslant 2$ for all $i \in [r]$.

By Lemma 2.40, we have $\mathcal{B}(\mathbf{H}) \subseteq \mathcal{B}'(\mathbf{H})$, and therefore

$$\mathbb{P}\big(G \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})\big) \leqslant \mathbb{P}\big(G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})\big),$$

where $G \sim \mathbb{G}(N,1/2)$, here and in every probability statement in this proof. It will also be convenient to assume that every such graph shares the same vertex set $V(G) = [N]$.

Let $\mathbf{U}$ denote the collection of tuples $(S, U, (\tilde{G}_i)_{i\in[r]})$ with the following properties. The sets $U \subseteq S \subseteq [N]$ satisfy

$$|S| \geqslant \delta^{2/3}N \quad \text{and} \quad |U| = \delta N + \sum_{i=1}^{r} \sigma_i, \quad \text{with} \quad 0 \leqslant \sigma_1, \dots, \sigma_r \leqslant \frac{p|U|}{2^9 r}. \tag{2.110}$$

Moreover, letting

$$\tilde{G} = \bigcup_{i\in[r]} \tilde{G}_i, \tag{2.111}$$

each tuple $(S, U, (\tilde{G}_i)_{i\in[r]}) \in \mathbf{U}$ satisfies, for all $i \in [r]$, that $V(\tilde{G}_i) = U$,

(1) $\mathfrak{I}_{H_i^-,\tilde{G}_i,\tilde{G}}[W]$ is $(p, p|W|)$-Janson for every $W \subseteq U$ with $|W| \geqslant |U|/(8r)$, and

(2) $\mathfrak{I}_{H_i,\tilde{G}_i,\tilde{G}}[U]$ is (p, σ_i)-Janson.

This collection is important because of the following claim. Before stating it, define $f(G) = (S, \chi)$ to map each $G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})$ to a fixed choice of $(S, \chi) \in \mathbf{S}(G)$.

Claim 2.46. *For each $G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})$ and $(S, \chi) = f(G)$, there is $\tau = \big(S, U, (\tilde{G}_i)_{i\in[r]}\big) \in \mathbf{U}$ for which $\tilde{G}_i = G_i^{(\chi)}[U]$ and the hypergraph $\mathfrak{I}_{H_i,G_i,G}\big[U \cup \{v\}\big]$ is not $(p, \sigma_i + 1)$-Janson for all $i \in [r]$ and all $v \in S \setminus U$.*

Note that the latter property of U and every $v \in S \setminus U$ in Claim 2.46 cannot be defined only in terms of $(\tilde{G}_i)_{i\in[r]}$, because it also depends on the colouring $\chi : E(G[S]) \to [r]$ in $(S, \chi) \in f(G)$.

Proof of Claim 2.46. Fix $G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})$ and $(S, \chi) = f(G)$. Let $U \subseteq S$ be given by Lemma 2.44, which implies that it satisfies (2.110). Further set $\tilde{G}_i = G_i^{(\chi)}[U]$ for every $i \in [r]$, and recall that $\tilde{G} = \bigcup_{i\in[r]} \tilde{G}_i$ by (2.111). It follows from item (a) in Lemma 2.44 that this choice satisfies item (2) in the definition of $\mathbf{U}$ and also that $\mathfrak{I}_{H_i,G_i,G}\big[U \cup \{v\}\big]$ is not $(p, \sigma_i + 1)$-Janson for all $i \in [r]$ and all $v \in S \setminus U$, where the values of each σ_i are given by Lemma 2.44.

To prove that this choice also satisfies item (1) in the definition of $\mathbf{U}$, observe first that Lemma 2.43 implies that $\mathfrak{I}_{H_i^-,G_i,G}[W]$ is $(p, p|W|)$-Janson for every $i \in [r]$ and every $W \subseteq S$ with $|W| \geqslant \delta N/(8r)$. As $U \subseteq S$, this conclusion also holds whenever $W \subseteq U$ and $|W| \geqslant |U|/(8r)$, since $|U| \geqslant \delta N$ by (2.110). The final observation is that

$$\mathfrak{I}_{H_i^-,\tilde{G}_i,\tilde{G}}[W] = \mathfrak{I}_{H_i^-,G_i,G}[W]$$

by our choice of $\tilde{G}_i = G_i^{(\chi)}[U]$, so the previous reasoning indeed establishes item (1). ∎

Now, for $\tau = \big(S, U, (\tilde{G}_i)_{i\in[r]}\big)$, let $\mathcal{A}(\tau)$ be the collection of pairs of graphs G and colourings $\chi : E(G) \to [r]$ such that

(a) $G_i[U] = G_i^{(\chi)}[U] = \tilde{G}_i$ for all $i \in [r]$, and

(b) the hypergraph $\mathfrak{I}_{H_i,G_i,G}\big[U \cup \{v\}\big]$ is not $(p, \sigma_i + 1)$-Janson for all $i \in [r]$ and all $v \in S \setminus U$.

By Claim 2.46, we know that if $G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})$, then there exists $\tau \in \mathbf{U}$ and a colouring $\chi : E(G) \to [r]$ such that $(G,\chi) \in \mathcal{A}(\tau)$. Taking a union bound over choices of $\tau \in \mathbf{U}$, but, crucially, *not* over the choices of $\chi : E(G) \to [r]$, then yields

$$\mathbb{P}\big(G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})\big) \leqslant \sum_{\tau \in \mathbf{U}} \mathbb{P}\big(\exists \chi \in [r]^{E(G)} : (G,\chi) \in \mathcal{A}(\tau)\big). \tag{2.112}$$

Most of the remainder of the proof will be dedicated to proving the following claim.

Claim 2.47. *For every $\tau \in \mathbf{U}$,*

$$\mathbb{P}\big(\exists \chi \in [r]^{E(G)} : (G,\chi) \in \mathcal{A}(\tau)\big) \leqslant 2^{-8r\delta^2 N^2}. \tag{2.113}$$

To prove Claim 2.47, we will modify $\mathcal{A}(\tau)$ until the events whose probabilities we need to bound become $\big\{\exists G' \subseteq G : (G',G) \in \mathcal{M}_{v,\ell}\big\}$, where

$$\mathcal{M}_{v,\ell} = \left\{(G',G) : \begin{array}{c} G'[U] = \tilde{G}_\ell,\ d_{G'}(v,U) \geqslant |U|/(4r) \ \text{ and} \\ \mathfrak{I}_{H_\ell,G',G}\big[U \cup \{v\}\big] \ \text{ is not } \ (p,\sigma_\ell+1)\text{-Janson} \end{array}\right\}. \tag{2.114}$$

We will then observe that not only the definition of $\mathcal{M}_{v,\ell}$ in (2.114) corresponds to the event whose probability Lemma 2.32 bounds if we take $F = H_\ell$, $\sigma' = \sigma_\ell$ and $\tilde{G}' = \tilde{G}_\ell$, but also that the current setting satisfies the assumptions to apply that lemma.

To make the proof easier to follow, we will first establish several intermediate claims towards Claim 2.47. Because of that, we will fix $\tau = \big(S,U,(\tilde{G}_i)_{i\in[r]}\big) \in \mathbf{U}$ and abbreviate $\mathcal{A}(\tau) = \mathcal{A}$ until the proof of Claim 2.47.

The first step towards proving (2.113) is to replace the set S with a large subset $S' \subseteq S$ by discarding vertices with small degree into U. By Observation 2.45, S contains such a subset with very high probability. More precisely, let

$$S' = S'(G) = \big\{v \in S \setminus U : d_G(v,U) > |U|/4\big\} \tag{2.115}$$

for every graph G, and define

$$\mathcal{A}' = \big\{(G,\chi) \in \mathcal{A} : |S'(G)| \geqslant |S|/4\big\}.$$

Claim 2.48.

$$\mathbb{P}\big(\exists \chi \in [r]^{E(G)} : (G,\chi) \in \mathcal{A}\big) \leqslant \mathbb{P}\big(\exists \chi \in [r]^{E(G)} : (G,\chi) \in \mathcal{A}'\big) + \exp\big(-2^{-6}\delta^{5/3}N^2\big). \tag{2.116}$$

Proof. Recalling (2.110) in the definition of $\mathbf{U}$ and that $p \leqslant 1$, we have

$$|U| \leqslant \delta N + \frac{p|U|}{2^9} \leqslant \delta N + \frac{|U|}{2},$$

which, since we also have $|U| \geqslant \delta N$, implies that

$$\delta N \leqslant |U| \leqslant 2\delta N. \tag{2.117}$$

As $2\delta N \leqslant |S|/4$, we can apply Observation 2.45 to S and U, obtaining as a result

$$\mathbb{P}\big(|S'| \leqslant |S|/4\big) \leqslant \exp\big(-2^{-6}|U||S|\big) \leqslant \exp\big(-2^{-6}\delta^{5/3}N^2\big), \tag{2.118}$$

where the last inequality follows from $|S| \geqslant \delta^{2/3}N$ by (2.110) and $|U| \geqslant \delta N$ by (2.117). The claim now follows from splitting $\mathcal{A}$ into $\mathcal{A}'$ and $\mathcal{A}\setminus\mathcal{A}'$ and bounding the latter using (2.118). ■

We now focus our attention on bounding the first term in the right-hand side of (2.116). Recalling that our goal is to apply Lemma 2.32 and that this lemma requires only a single subgraph, instead of a colouring, we will restrict our attention to the subgraph G_ℓ whose colour $\ell \in [r]$ is the majority colour of the neighbourhood of most vertices in S'.

Claim 2.49.

$$\mathbb{P}\big(\exists \chi \in [r]^{E(G)} : (G,\chi) \in \mathcal{A}'\big) \leqslant \sum_{\ell\in[r]} \sum_{\substack{A\subseteq S\setminus U\\ |A|\geqslant \frac{|S|}{4r}}} \mathbb{P}\Big(\exists G' \subseteq G : (G',G) \in \bigcap_{v\in A} \mathcal{M}_{v,\ell} \text{ and } G[U] = \tilde{G}\Big).$$

Proof. Let G be a graph, and suppose that there exists a colouring $\chi : E(G) \to [r]$ such that $(G,\chi) \in \mathcal{A}'$. We claim that $G[U] = \tilde{G}$, and that there exists a subgraph $G' \subseteq G$, a colour $\ell \in [r]$ and a subset $A \subseteq S \setminus U$ with $|A| \geqslant |S|/(4r)$, such that $(G',G) \in \mathcal{M}_{v,\ell}$ for every $v \in A$. Claim 2.49 will then follow by taking a union bound over the choices of ℓ and A.

To show this, observe first that $G[U] = \tilde{G}$ holds because $(G,\chi) \in \mathcal{A}' \subseteq \mathcal{A}$, where $\tilde{G} = \bigcup_{i\in[r]} \tilde{G}_i$ was defined in (2.111). Next, note that for each $\chi : E(G) \to [r]$ and $v \in S'(G)$, the colouring χ partitions the edges connecting v and U into r sets. As each $v \in S'(G)$ satisfies

$$d_G(v,U) > \frac{|U|}{4}$$

by the definition of $S'(G)$, (2.115), there exists a colour $j(v) = j \in [r]$ such that

$$d_{G_j}(v,U) > \frac{|U|}{4r}. \tag{2.119}$$

By the pigeonhole principle and (2.119), then, there exists a colour $\ell \in [r]$ such that, letting

$$A = A(G,\chi) = \{v \in S'(G) : j(v) = \ell\},$$

we have, as a consequence of $(G,\chi) \in \mathcal{A}'$, that

$$|A| \geqslant \frac{|S'(G)|}{r} \geqslant \frac{|S|}{4r}.$$

Letting $G' = G_\ell$ be the graph of edges spanned by colour ℓ completes the proof of the claim because $A \subseteq S\setminus U$ and $(G,\chi) \in \mathcal{A}$ satisfy item (b) in the definition of the event $\mathcal{A}$, which shows that the last property of $\mathcal{M}_{v,\ell}$ holds for this choice of G'. ■

The final claim that we need before the proof of Claim 2.47 is the observation that the events $\mathcal{M}_{v,\ell}$ are independent for each $v \notin U$ when conditioned on $\{G[U] = \tilde{G}\}$. This is a

simple consequence of the conditioning causing $\mathcal{M}_{v,\ell}$ to depend only on the subgraph $G[v,U]$ corresponding to the edges between v and U. Although Claim 2.49 has $\{G[U] = \tilde{G}\}$ as an intersecting event, we can instead condition on it because its probability is non-zero.

Claim 2.50. *For any set $A \subseteq S \setminus U$, we have*

$$\mathbb{P}\Big(\exists G' \subseteq G : (G', G) \in \bigcap_{v\in A} \mathcal{M}_{v,\ell} \,\Big|\, G[U] = \tilde{G}\Big) = \prod_{v\in A} \mathbb{P}\Big(\exists G' \subseteq G : (G', G) \in \mathcal{M}_{v,\ell} \,\Big|\, G[U] = \tilde{G}\Big).$$

Proof. First recall that $\mathcal{M}_{v,\ell}$ is defined in (2.114) as

$$\mathcal{M}_{v,\ell} = \left\{(G', G) : \begin{array}{c} G'[U] = \tilde{G}_\ell,\ d_{G'}(v, U) \geqslant |U|/(4r) \ \text{ and} \\ \mathfrak{I}_{H_\ell,G',G}\big[U \cup \{v\}\big] \ \text{ is not } \ (p, \sigma_\ell + 1)\text{-Janson} \end{array}\right\}.$$

Now, observe that after conditioning on $G[U]$, the existence of a subgraph G' satisfying the three properties in the definition of $\mathcal{M}_{v,\ell}$ depends only on the edges of $G[v,U]$. Since these edges are chosen independently for each vertex $v \in A$, the claim follows. ∎

We can now combine the previous claims to prove Claim 2.47.

Proof of Claim 2.47. Fix $\tau = \big(S, U, (\tilde{G}_i)_{i\in[r]}\big) \in \mathbf{U}$ and let $\mathcal{A} = \mathcal{A}(\tau)$. By Claim 2.48, we have

$$\mathbb{P}\big(\exists \chi \in [r]^{E(G)} : (G, \chi) \in \mathcal{A}\big) \leqslant \mathbb{P}\big(\exists \chi \in [r]^{E(G)} : (G, \chi) \in \mathcal{A}'\big) + \exp\big(-2^{-6}\delta^{5/3}N^2\big). \quad (2.120)$$

Claim 2.49 then implies that the first term in this right-hand side is at most

$$\mathbb{P}\big(\exists \chi \in [r]^{E(G)} : (G, \chi) \in \mathcal{A}'\big) \leqslant \sum_{\ell\in[r]} \sum_{A} \mathbb{P}\Big(\exists G' \subseteq G : (G', G) \in \bigcap_{v\in A} \mathcal{M}_{v,\ell} \,\Big|\, G[U] = \tilde{G}\Big), \quad (2.121)$$

where the sum over A ranges over all $A \subseteq S \setminus U$ such that $|A| \geqslant |S|/(4r)$, and we moved the non-zero probability event $\{G[U] = \tilde{G}\}$ from the intersection to the conditioning. As a consequence of Claim 2.50, the probability term inside the two sums of (2.121) satisfies

$$\mathbb{P}\Big(\exists G' \subseteq G : (G', G) \in \bigcap_{v\in A} \mathcal{M}_{v,\ell} \,\Big|\, G[U] = \tilde{G}\Big) = \prod_{v\in A} \mathbb{P}\Big(\exists G' : (G', G) \in \mathcal{M}_{v,\ell} \,\Big|\, G[U] = \tilde{G}\Big). \quad (2.122)$$

We now want to apply Lemma 2.32 to obtain an upper bound for the probabilities on the right-hand side of ((2.122)), recalling that (2.114), the definition of $\mathcal{M}_{v,\ell}$, corresponds to the event whose probability we bound in (2.85) if we take $F = H_\ell$, $\sigma' = \sigma_\ell$ and $\tilde{G}' = \tilde{G}_\ell$. To check that the choice of parameters for this application is admissible, first observe that $s = v(H_\ell) - 1$ by assumption, that

$$m = |U| \geqslant \delta N \geqslant \delta r^{C'(k+t)} \geqslant r^{C''k},$$

where the final inequality follows from $k, t \geqslant 1$ and $\delta = r^{-50}$ if we take $C' \geqslant C'' + 50$. It follows from $(S, U, (\tilde{G}_i)_{i\in[r]}) \in \mathbf{U}$ and the definition of $\mathbf{U}$ that not only

$$\sigma_\ell \leqslant \frac{p|U|}{2^9 r} = \frac{\sigma}{16}$$

by (2.110) and our choice of $\sigma = 2^{-5}r^{-1}p|U|$ in (2.84), but also that $\tilde{G}_\ell \subseteq \tilde{G}$ by (2.111). Finally, again by the properties of the tuples in $\mathbf{U}$, we have that

(1) $\mathfrak{I}_{H_\ell,\tilde{G}_\ell,\tilde{G}}[U]$ is (p,σ_ℓ)-Janson, and

(2) $\mathfrak{I}_{H_\ell^-,\tilde{G}_\ell,\tilde{G}}[W]$ is (p,σ)-Janson for every $W \subseteq U$ with $|W| \geqslant |U|/(8r)$, because

$$p|W| \geqslant \frac{p|U|}{8r} \geqslant 2^{-5}r^{-1}p|U| = \sigma$$

and the Janson property is increasing (Observation 2.3),

so we can apply Lemma 2.32.

Applying Lemma 2.32, we then obtain, for fixed $\ell \in [r]$ and $v \in S \setminus U$, that

$$\mathbb{P}\Big(\exists G' \subseteq G : (G',G) \in \mathcal{M}_{v,\ell} \,\Big|\, G[U] = \tilde{G}\Big) \leqslant 2^{-|U|/(2^5 r)},$$

which, replaced in ((2.122)), yields

$$\mathbb{P}\Big(\exists G' \subseteq G : (G',G) \in \bigcap_{v\in A} \mathcal{M}_{v,\ell} \,\Big|\, G[U] = \tilde{G}\Big) \leqslant 2^{-|U||A|/(2^5 r)} \leqslant 2^{-\delta^{5/3}N^2/(2^7 r^2)} \tag{2.123}$$

by $|A| \geqslant |S|/(4r) \geqslant \delta^{2/3}N/(4r)$, where $A \subseteq S \setminus U$, and $|U| \geqslant \delta N$. Substituting (2.123) in (2.121) and bounding the number of choices for ℓ and A respectively by r and 2^N then yields

$$\mathbb{P}\big(\exists \chi \in [r]^{E(G)} : (G,\chi) \in \mathcal{A}'\big) \leqslant r2^N 2^{-\delta^{5/3}N^2/(2^7 r^2)}. \tag{2.124}$$

To deduce the claim, we combine (2.120) and (2.124) to obtain

$$\mathbb{P}\big(\exists \chi \in [r]^{E(G)} : (G,\chi) \in \mathcal{A}\big) \leqslant r2^N 2^{-\delta^{5/3}N^2/(2^7 r^2)} + \exp\big(-2^{-6}\delta^{5/3}N^2\big) \leqslant 2^{-8r\delta^2 N^2}$$

by $r \geqslant 2$, our assumptions that $N \geqslant r^{C'(k+t)}$ and $\delta = r^{-50}$. ■

We now apply Claim 2.47 in every term of (2.112) to obtain

$$\mathbb{P}\big(G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})\big) \leqslant \sum_{\tau\in\mathbf{U}} 2^{-8r\delta^2 N^2} \tag{2.125}$$

To bound the right-hand side of (2.125), we count the number of tuples $\tau = \big(S, U, (\tilde{G}_i)_{i\in[r]}\big)$ in $\mathbf{U}$. There are at most 2^{2N} choices for both $S \subseteq V$ and $U \subseteq V$, and at most $2^{|U|^2}$ choices for each $\tilde{G}_i$. It follows from $|U| \leqslant 2\delta N$ in (2.117) that there are at most

$$2^{r|U|^2} \leqslant 2^{4r\delta^2 N^2}$$

tuples $(\tilde{G}_i)_{i\in[r]}$, which replaced back in (2.125) yields

$$\mathbb{P}\big(G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})\big) \leqslant 2^{2N+4r\delta^2N^2-8r\delta^2N^2} \leqslant 2^{-\delta^2N^2}$$

because we assumed that $N \geqslant r^{C'(k+t)}$ and $\delta = r^{-50}$. □

2.7 Containers for non-Janson sets

In this section, we prove our main technical result, and with it complete the proof of Theorem 1.2. The statement of Theorem 2.35 has three components that differ from Theorem 2.21: another hypergraph $\mathcal{F}$, which is (p,σ')-Janson by assumption, a function π and a vertex v not in the set $U = V(\mathcal{F})$. Recall that when applying this theorem, we will take $\mathcal{F}$ to correspond to copies of F completely contained in U, and π to be the projection from $U \times \{0,1\}$ onto U. To explain the role of v in the statement of Theorem 2.35, recall Definition 2.33,

$$\overline{\partial}_v \mathcal{G} = \{E \cup \{v\} : E \in \mathcal{G}\},$$

the edge-wise inclusion of v in a hypergraph $\mathcal{G}$.

Rather than obtaining containers for sets $L \subseteq V$ such that $\mathcal{H}[L]$ is not $(p/q, \eta\sigma)$-Janson, Theorem 2.35 provides containers for sets $L \subseteq V$ such that $\pi_v(\mathcal{H}[L]) \cup \mathcal{F}$ is not $(p, \sigma' + \eta\sigma)$-Janson, where $\pi_v = \overline{\partial}_v \circ \pi$. Another difference between this theorem and Theorem 2.21 is in the properties of the containers $X \in \mathcal{X}$. Previously, we concluded that $\mathcal{H}[X]$ was not (p,σ)-Janson, but we were not able to establish the same thing here due to vertices with high-degree. Instead, what we show is that whenever $X \subseteq V$ has linear size, we have a set $Y \subseteq X$ containing almost all elements of X such that $\pi(\mathcal{H}[Y])$ is not (p,σ)-Janson.

Theorem 2.35. *Let $n, r, s \in \mathbb{N}$ with $n \geqslant s$ and $r \geqslant 2$, and let $q, p, \sigma, \sigma', \eta \in \mathbb{R}$ satisfy*

$$0 < q < \frac{1}{8}, \quad 0 < p \leqslant \frac{q}{2^{10} r^2 s^2}, \quad \sigma = 2^{-6} r^{-1} pn, \quad 0 \leqslant \sigma' \leqslant \frac{\sigma}{16} \quad \text{and} \quad \eta = p^4 \Big(\frac{q}{2}\Big)^{4s}. \tag{2.126}$$

Further let $\mathcal{F}$ be a $(s+1)$-uniform hypergraph with vertex set U that is (p,σ')-Janson, let $\mathcal{H}$ be an s-uniform hypergraph with vertex set V, where $|V| = n$, and let $\pi : V \to U$ satisfy

$$|\pi(L)| \geqslant \frac{|L|}{2} \quad \text{for every } L \subseteq V \qquad \text{and} \qquad |\pi(E)| = |E| \quad \text{for every } E \in \mathcal{H}. \tag{2.127}$$

Finally, let v be a vertex not in U. There exists a family $\mathcal{X} \subseteq 2^V$ with

$$|\mathcal{X}| \leqslant \left(\frac{2}{q}\right)^{2qn} \tag{2.128}$$

such that the following hold.

(1) *If $I \subseteq V$ and $\pi_v(\mathcal{H}[I]) \cup \mathcal{F}$ is not $(p, \sigma' + \eta\sigma)$-Janson, then $I \subseteq X$ for some $X \in \mathcal{X}$.*

(2) *For each $X \in \mathcal{X}$ with $|X| \geqslant n/(8r)$, there exists $Y \subseteq X$ with*

$$|Y| \geqslant |X| - 2^{-8} r^{-1} n$$

such that $\pi(\mathcal{H}[Y])$ is not (p,σ)-Janson.

We now highlight differences between the proof of Theorem 2.35 and the one in Section 2.4. The first one is already in the auxiliary hypergraph $\mathcal{J}'$, which is defined here by

$$\mathcal{J}' = \Big\{L \subseteq V : \pi_v\big(\mathcal{H}[L]\big) \cup \mathcal{F} \text{ is } (p, \sigma' + \eta\sigma)\text{-Janson}\Big\}.$$

Note that, as the Janson property is increasing by Observation 2.5, each $L \subseteq V$ for which the hypergraph $\pi_v(\mathcal{H}[L]) \cup \mathcal{F}$ is not $(p, \sigma' + \eta\sigma)$-Janson is also an independent set in $\mathcal{J}'$.

The start of the proof of Theorem 2.35 is analogous to the proof of Theorem 2.26: we apply Theorem 2.23 with $\mathcal{G} = \mathcal{J}'$, define $\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_{=s}$ for each $T \in \mathcal{T}$, and apply Theorem 2.15 with $\mathcal{G} = \mathcal{C}'_T$. This application of Theorem 2.15 also provides our candidate containers X, and the bulk of the argument is proving that they fulfil the conclusions of the theorem, in particular that $\pi(\mathcal{H}[Y])$ is not (p, σ)-Janson for some large $Y \subseteq X$.

Towards determining that $\mathcal{H}[X]$ was not (p, σ)-Janson for fixed $X \in \mathcal{X}$, in the previous proof we took an arbitrary measure $\nu : \mathcal{H}[X] \to \mathbb{R}_{\geqslant 0}$ and showed that

$$\Lambda_p(\nu) \geqslant \frac{e(\nu)^2}{\sigma}, \tag{2.129}$$

by considering two cases, depending on whether $e(\nu')$ was sufficiently large, where ν' was the restriction of ν to $\mathcal{C}'_T[X]$. Our goal in item (2) of Theorem 2.35 is to establish something like (2.129) for measures μ supported on $\pi(\mathcal{H}[Y])$ for some large $Y \subseteq X$. In order to do so, we will need to introduce some additional machinery, which will allow us to relate Janson properties of $\mathcal{H}$ and those of $\pi(\mathcal{H})$.

To reason about the Janson properties under the effect of π, we define the pullback of a measure ϑ with respect to π. The resulting measure, denoted by $\vartheta \mathbin{\hat{\circ}} \pi$, distributes the mass of $E \in \pi(\mathcal{G})$ equally among edges that are entirely contained in its pre-image.

Definition 2.51. Let $\mathcal{G}$ be a hypergraph, U be a set, and let $\pi : V(\mathcal{G}) \to U$. If $\vartheta : \pi(\mathcal{G}) \to \mathbb{R}_{\geqslant 0}$ is a measure, then $\vartheta \mathbin{\hat{\circ}} \pi : \mathcal{G} \to \mathbb{R}_{\geqslant 0}$ is defined by

$$\vartheta \mathbin{\hat{\circ}} \pi(E) = \frac{\vartheta(\pi(E))}{\big|\{E_0 \in \mathcal{G} : \pi(E_0) = \pi(E)\}\big|} \tag{2.130}$$

for all $E \in \mathcal{G}$.

The crucial property of pullback measures is that, using them, we can show that for any hypergraph $\mathcal{G}$, if $\pi(\mathcal{G})$ is (p, σ)-Janson, then so is $\mathcal{G}$.

Lemma 2.52. *Let $\sigma > 0$ and $p > 0$. Further let $\mathcal{G}$ be a hypergraph, U be a set, $\pi : V(\mathcal{G}) \to U$ be a function satisfying*

$$|\pi(E)| = |E| \qquad \text{for every } E \in \mathcal{G},$$

and $\vartheta : \pi(\mathcal{G}) \to \mathbb{R}_{\geqslant 0}$ be a measure. If

$$\Lambda_p(\vartheta) < \frac{e(\vartheta)^2}{\sigma},$$

then

$$\Lambda_p(\vartheta \mathbin{\hat{\circ}} \pi) < \frac{e(\vartheta \mathbin{\hat{\circ}} \pi)^2}{\sigma}.$$

In particular, if $\pi(\mathcal{G})$ is (p, σ)-Janson, then so is $\mathcal{G}$.

We will postpone proving Lemma 2.52 to Section 2.7.1, after we have completed our overview of the proof of Theorem 2.35.

With the definition of the pullback measure and its crucial property, we can state the main inequalities in the proof that $\pi(\mathcal{H}[Y])$ is not (p,σ)-Janson. We will start with a simplified view of the proof, and add details as we proceed. For technical reasons, it will be useful to assume instead the converse of our goal, i.e. we will fix one $X \in \mathcal{X}$ such that, for all large Y, the hypergraph $\pi(\mathcal{H}[Y])$ is (p,σ)-Janson, and reach a contradiction. In particular, $\pi(\mathcal{H}[X])$ is (p,σ)-Janson, so there is a measure $\mu : \pi(\mathcal{H}[X]) \to \mathbb{R}_{\geqslant 0}$ such that

$$\Lambda_p(\mu) < \frac{e(\mu)^2}{\sigma} \tag{2.131}$$

and the pullback $\nu = \mu \mathbin{\hat{\circ}} \pi$ satisfies

$$\Lambda_p(\nu) < \frac{e(\nu)^2}{\sigma}$$

by Lemma 2.52.

Recall that a critical step in the proof of the simpler version of our container theorem was the inequality

$$\mathbb{E}\big[\Lambda_{p/q}(\nu_q'') \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})\big] \geqslant \frac{\mathbb{E}\big[e(\nu_q'')^2 \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J})\big]}{\eta\sigma}, \tag{2.132}$$

established in Claim 2.30 via the characterization of independent sets in $\mathcal{J}$. The analogous statement here depends on the definition of $\mathcal{J}'$, which implies that when $I \subseteq V(\mathcal{J}')$ is independent, then $\pi_v\big(\mathcal{H}[I]\big) \cup \mathcal{F}$ is not $(p, \sigma' + \eta\sigma)$-Janson. To avoid too many technical details at once, let us assume that $\sigma' > 0$, and that we have $\pi(\cdot)$ instead of $\pi_v(\cdot)$. These simplifications mean that $\pi\big(\mathcal{H}[I]\big) \cup \mathcal{F}$ is not $(p, \sigma' + \eta\sigma)$-Janson when $I \in \mathcal{I}(\mathcal{J}')$.

In this simpler setup, we now use the assumption that $\mathcal{F}$ is (p,σ')-Janson and $\sigma' > 0$ to choose $\rho : \mathcal{F} \to \mathbb{R}_{\geqslant 0}$ with

$$\Lambda_p(\rho) < \frac{e(\rho)^2}{\sigma'}. \tag{2.133}$$

A case analysis, the same as in Claim 2.28 in the proof of Theorem 2.26, will allow us to focus on $\mathcal{H}' = \mathcal{H}[X] \setminus \mathcal{C}_T'$, so we define $\mu_q : \pi(\mathcal{H}[X]) \to \mathbb{R}_{\geqslant 0}$ by

$$\mu_q(E) = \frac{\gamma \cdot \mathbb{1}\big[E \in \pi\big(\mathcal{H}'[V_q]\big)\big]}{P_q'(E)}\,\mu(E),$$

where we will choose $\gamma = \sqrt{8\eta}$ and

$$P_q'(E) = \mathbb{P}\big(E \in \pi\big(\mathcal{H}'[V_q]\big) \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big).$$

The definition of $\mathcal{H}'$ will allow us to assume that

$$P_q'(E) > \Big(\frac{q}{2}\Big)^{|E|},$$

cf. (2.75), so μ_q is well-defined.

The inequality corresponding to (2.132) in this simplified overview of the proof will then be

$$\mathbb{E}\big[\Lambda_p(\rho + \mu_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big] \geqslant \frac{\mathbb{E}\big[e(\rho + \mu_q)^2 \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big]}{\sigma' + \eta\sigma}, \tag{2.134}$$

which, note, has Λ_p instead of $\Lambda_{p/q}$ due to some still undiscussed numerics related to the fact that, in this case, the value of η is much smaller here than in Theorem 2.21. In the other direction, we would ideally like to show that

$$\mathbb{E}\big[\Lambda_p(\rho+\mu_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J}')\big] \leqslant \Lambda_p(\rho) + 2\gamma\sqrt{\Lambda_p(\rho)\Lambda_p(\mu)} + \gamma^2\Big(\frac{q}{2}\Big)^{-2s}\Lambda_p(\mu), \tag{2.135}$$

and

$$\mathbb{E}\big[e(\rho+\mu_q)^2 \mid V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J}')\big] \geqslant \big(e(\rho)+\gamma e(\mu)\big)^2, \tag{2.136}$$

both of which, when combined with (2.131), (2.133) and our choice of γ, yield

$$\mathbb{E}\big[\Lambda_p(\rho+\mu_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J}')\big] < \frac{\mathbb{E}\big[e(\rho+\mu_q)^2 \mid V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J}')\big]}{\sigma'+\eta\sigma}$$

a direct contradiction of (2.134). Although we can establish inequalities resembling (2.135) and (2.136) with methods similar to those in Section 2.4, we have not justified some of our assumptions.

The first detail that we overlooked in the preceding overview is the assumption that $\sigma' > 0$. In the case $\sigma' = 0$, we can simply take $\rho = 0$. To handle both cases together, we use Observation 2.4 when $\sigma' > 0$ and assume instead that

$$e(\rho) = \sqrt{\sigma'} \qquad \text{and} \qquad \Lambda_p(\rho) < 1,$$

which is sufficient for our purposes.

The other omission in our discussion so far is that we simplified the definition of $\mathcal{J}'$. The correct definition means that, to satisfy something like (2.134), we require a measure supported on $\pi_v\big(\mathcal{H}[I]\big) \cup \mathcal{F}$, instead of $\pi\big(\mathcal{H}[I]\big) \cup \mathcal{F}$, where $I \in \mathcal{I}(\mathcal{J}')$. To address this change, we extend $\mu : \pi\big(\mathcal{H}[X]\big) \to \mathbb{R}_{\geqslant 0}$ to a measure $\bar{\mu} : \pi_v\big(\mathcal{H}[X]\big) \to \mathbb{R}_{\geqslant 0}$ by setting

$$\bar{\mu}\big(E \cup \{v\}\big) = \mu(E)$$

for all $E \in \pi(\mathcal{H})$, recalling that $v \notin U \supset \pi(V)$.

Adjusting for this seemingly innocuous change requires some technicalities. In order to prove the appropriate version of (2.134), with $\bar{\mu}_q$ replacing μ_q, we now need an upper bound for $\sum_{u\in\pi(X)} d_\mu(u)^2$. This is how we use Lemma 2.53 below, and is the main reason why we prove item (2) in Theorem 2.35 for a large $Y \subseteq X$ instead of all of $X \in \mathcal{X}$; it is also why it is simpler to prove that same item by contradiction. The proof of Lemma 2.53 is straightforward: we simply restrict the measure to avoid vertices with high degree.

Lemma 2.53. *Let $s \in \mathbb{N}$, $\sigma, p, \beta > 0$, and $\mathcal{G}$ be an s-uniform hypergraph. If for every $W \subseteq V(\mathcal{G})$ with $|W| \geqslant (1-\beta)v(\mathcal{G})$, we have that $\mathcal{G}[W]$ is (p,σ)-Janson, then there exists $\mu : \mathcal{G} \to \mathbb{R}_{\geqslant 0}$ with*

$$e(\mu) = \sqrt{\sigma}, \qquad \Lambda_p(\mu) < \frac{e(\mu)^2}{\sigma} \qquad \text{and} \qquad \sum_{v\in V(\mathcal{G})} d_\mu(v)^2 \leqslant \frac{2s^2 e(\mu)^2}{\beta v(\mathcal{G})}. \tag{2.137}$$

Proof. Observe that taking $W = V(\mathcal{G})$, we conclude that $\mathcal{G} = \mathcal{G}[W]$ is (p,σ)-Janson by assump-

tion. Therefore, we can apply Observation 2.4 to obtain $\mu : \mathcal{G} \to \mathbb{R}_{\geqslant 0}$ such that

$$e(\mu) = \sqrt{\sigma} \qquad \text{and} \qquad \Lambda_p(\mu) < \frac{e(\mu)^2}{\sigma}. \tag{2.138}$$

Take such a μ minimising $\sum_{v \in V(\mathcal{G})} d_\mu(v)^2$, and assume by contradiction that

$$\sum_{v \in V(\mathcal{G})} d_\mu(v)^2 > \frac{2s^2 e(\mu)^2}{\beta v(\mathcal{G})}. \tag{2.139}$$

Now, take

$$W = \{v \in V(\mathcal{G}) : d_\mu(v) \leqslant s\, e(\mu)/(\beta v(\mathcal{G}))\} \tag{2.140}$$

observe that it satisfies $|W| \geqslant (1-\beta)v(\mathcal{G})$ since $\mathcal{G}$ is s-uniform, and therefore $G[W]$ is (p,σ)-Janson by assumption. We conclude that there is a measure $\mu' : \mathcal{G}[W] \to \mathbb{R}_{\geqslant 0}$ which satisfies

$$\Lambda_p(\mu') < \frac{e(\mu')^2}{\sigma} \qquad \text{and} \qquad e(\mu') = \sqrt{\sigma}, \tag{2.141}$$

again by Observation 2.4 applied with $y = \sqrt{\sigma} > 0$. Moreover, we claim that setting

$$\mu'' = (1-\xi)\mu + \xi\mu',$$

for a suitable $0 < \xi \leqslant 1$, also yields a measure satisfying

$$\Lambda_p(\mu'') < \frac{e(\mu'')^2}{\sigma}, \qquad e(\mu'') = \sqrt{\sigma} \tag{2.142}$$

and

$$\sum_{v \in V(\mathcal{G})} d_{\mu''}(v)^2 < \sum_{v \in V(\mathcal{G})} d_\mu(v)^2. \tag{2.143}$$

If μ'' satisfies both (2.142) and (2.143), it contradicts the minimality of our original choice of μ.

We start the proof of our claims by showing that (2.142) holds. The equality $e(\mu'') = \sqrt{\sigma}$ follows by linearity of $e(\cdot)$ and the fact that both μ and μ' have edge measure equal to $\sqrt{\sigma}$. Expand $\Lambda_p(\mu'')$ as

$$\Lambda_p(\mu'') = (1-\xi)^2\Lambda_p(\mu) + 2\xi(1-\xi)\sum_{\substack{L \subseteq V(\mathcal{G}) \\ |L| \geqslant 2}} d_\mu(L) d_{\mu'}(L) p^{-|L|} + \xi^2 \Lambda_p(\mu'). \tag{2.144}$$

Applying the Cauchy–Schwarz inequality to the second term in (2.144), we obtain

$$\sum_{\substack{L \subseteq V(\mathcal{G}) \\ |L| \geqslant 2}} \frac{d_\mu(L) d_{\mu'}(L)}{p^{|L|}} \leqslant \left(\sum_{\substack{L \subseteq V(\mathcal{G}) \\ |L| \geqslant 2}} \frac{d_\mu(L)^2}{p^{|L|}} \right)^{1/2} \left(\sum_{\substack{L \subseteq V(\mathcal{G}) \\ |L| \geqslant 2}} \frac{d_{\mu'}(L)^2}{p^{|L|}} \right)^{1/2} = \sqrt{\Lambda_p(\mu)\Lambda_p(\mu')}. \tag{2.145}$$

Replacing (2.145) back in (2.144) and simplifying yields

$$\Lambda_p(\mu'') \leqslant (1-\xi)^2\Lambda_p(\mu) + 2\xi(1-\xi)\sqrt{\Lambda_p(\mu)\Lambda_p(\mu')} + \xi^2\Lambda_p(\mu') = \left((1-\xi)\Lambda_p(\mu)^{1/2} + \xi\Lambda_p(\mu')^{1/2}\right)^2,$$

which, by (2.138) and (2.141), establishes (2.142):

$$\Lambda_p(\mu'') < \frac{1}{\sigma}\big((1-\xi)e(\mu)+\xi e(\mu')\big)^2 = \frac{e(\mu'')^2}{\sigma}.$$

Towards establishing (2.143), we expand

$$\sum_{v\in V(\mathcal{G})} d_{\mu''}(v)^2 = (1-\xi)^2 \sum_{v\in V(\mathcal{G})} d_\mu(v)^2 + 2\xi(1-\xi)\sum_{v\in V(\mathcal{G})} d_\mu(v)d_{\mu'}(v) + \xi^2 \sum_{v\in V(\mathcal{G})} d_{\mu'}(v)^2, \quad (2.146)$$

an expression whose terms we will bound separately. Recall that $d_{\mu'}(v) = 0$ for $v \notin W$, so

$$\sum_{v\in V(\mathcal{G})} d_\mu(v)d_{\mu'}(v) = \sum_{v\in W} d_\mu(v)d_{\mu'}(v)$$

but now, (2.140) implies that

$$\sum_{v\in W} d_\mu(v)d_{\mu'}(v) \leqslant \frac{s\,e(\mu)}{\beta v(\mathcal{G})}\sum_{v\in W} d_{\mu'}(v) = \frac{s^2 e(\mu)^2}{\beta v(\mathcal{G})} \quad (2.147)$$

since $\mathcal{G}$ is s-uniform and $e(\mu') = e(\mu)$. Using (2.139) in (2.147), we obtain, for the second term,

$$\sum_{v\in W} d_\mu(v)d_{\mu'}(v) < \frac{1}{2}\sum_{v\in V(\mathcal{G})} d_\mu(v)^2. \quad (2.148)$$

The s-uniformity of $\mathcal{G}$ and $e(\mu) = e(\mu')$ also imply an easy bound on the third term in (2.146):

$$\sum_{v\in V(\mathcal{G})} d_{\mu'}(v)^2 \leqslant \Big(\sum_{v\in V(\mathcal{G})} d_{\mu'}(v)\Big)^2 = \big(e(\mu)s\big)^2 < \beta v(\mathcal{G})\sum_{v\in V(\mathcal{G})} d_\mu(v)^2, \quad (2.149)$$

where the first inequality holds because d'_μ is always non-negative, and the second is (2.139).

We can now replace (2.148) and (2.149) in (2.146) and use that ξ is positive to obtain, after simplification, that

$$\sum_{v\in V(\mathcal{G})} d_{\mu''}(v)^2 < (1-\xi+\beta v(\mathcal{G})\xi^2)\sum_{v\in V(\mathcal{G})} d_\mu(v)^2.$$

Choosing ξ to satisfy

$$0 < \xi < \min\Big\{1, \frac{1}{\beta v(\mathcal{G})}\Big\}$$

results in

$$\sum_{v\in V(\mathcal{G})} d_{\mu''}(v)^2 < \sum_{v\in V(\mathcal{G})} d_\mu(v)^2$$

which contradicts the fact that μ minimises $\sum_{v\in V(\mathcal{G})} d_\mu(v)^2$ and completes the proof. □

When we account for the term handled by Lemma 2.53 in the final version of (2.134), it dominates the term for $\Lambda_p(\nu)$, so a simpler bound for the latter suffices, and we do not need to consider $\Lambda_{p/q}$.

2.7.1 Properties of the pullback measure

The missing proof of Lemma 2.52 is a trivial combination of Lemma 2.55, which says that $e(\vartheta \mathbin{\hat{\circ}} \pi) = e(\vartheta)$, and Lemma 2.57, which establishes $\Lambda_p(\vartheta \mathbin{\hat{\circ}} \pi) \leqslant \Lambda_p(\vartheta)$. Towards these two lemmas, we first prove an elementary observation about pullback measures.

Observation 2.54. *Let $\mathcal{G}$ be a hypergraph and let $\pi : V(\mathcal{G}) \to U$ for some set U. Further let $\vartheta : \pi(\mathcal{G}) \to \mathbb{R}_{\geqslant 0}$ be a measure. If $E' \in \pi(\mathcal{G})$, then*

$$\sum_{\substack{E \in \mathcal{G} \\ \pi(E) = E'}} \vartheta \mathbin{\hat{\circ}} \pi(E) = \vartheta(E').$$

Proof. Fix $E' \in \pi(\mathcal{G})$. We have by (2.130) that

$$\sum_{\substack{E \in \mathcal{G} \\ \pi(E) = E'}} \vartheta \mathbin{\hat{\circ}} \pi(E) = \sum_{\substack{E \in \mathcal{G} \\ \pi(E) = E'}} \frac{\vartheta(E')}{\big|\{E \in \mathcal{G} : \pi(E) = E'\}\big|} = \vartheta(E')$$

which is what we wanted to show. □

We can now prove Lemma 2.55 directly from the definition of the pullback measure.

Lemma 2.55. *For all hypergraphs $\mathcal{G}$, functions $\pi : V(\mathcal{G}) \to U$ and measures $\vartheta : \pi(\mathcal{G}) \to \mathbb{R}_{\geqslant 0}$, we have*

$$e(\vartheta \mathbin{\hat{\circ}} \pi) = e(\vartheta).$$

Proof. Expanding the definition of $e(\vartheta \mathbin{\hat{\circ}} \pi)$ and using Observation 2.54, we obtain

$$e(\vartheta) = \sum_{E' \in \pi(\mathcal{G})} \vartheta(E') = \sum_{E' \in \pi(\mathcal{G})} \sum_{\substack{E \in \mathcal{G} \\ \pi(E) = E'}} \vartheta \mathbin{\hat{\circ}} \pi(E).$$

But each $E \in \mathcal{G}$ appears on the right-hand side exactly once, only when $E' \in \pi(\mathcal{G})$ satisfies $\pi(E) = E'$. Therefore,

$$\sum_{E' \in \pi(\mathcal{G})} \sum_{\substack{E \in \mathcal{G} \\ \pi(E) = E'}} \vartheta \mathbin{\hat{\circ}} \pi(E) = \sum_{E \in \mathcal{G}} \vartheta \mathbin{\hat{\circ}} \pi(E) = e(\vartheta \mathbin{\hat{\circ}} \pi)$$

as we wanted to show. □

The next lemma requires an assumption about π, motivating its appearance in the statements of Lemma 2.52 and Theorem 2.35. The proof is easy, and follows from expanding the definitions and a simple counting argument. It also requires defining the uniformity-preserving pre-image of $L' \subseteq \pi(V(\mathcal{G}))$, the set

$$\overleftarrow{\pi}(L') = \{L \subseteq V(\mathcal{G}) : \pi(L) = L' \text{ and } |L| = |L'|\}.$$

Lemma 2.56. *Let $\mathcal{G}$ be a hypergraph, let $\pi : V(\mathcal{G}) \to U$ satisfy*

$$|\pi(E)| = |E| \qquad \text{for every } E \in \mathcal{G},$$

and let $\vartheta : \pi(\mathcal{G}) \to \mathbb{R}_{\geqslant 0}$ *be a measure. For all* $L' \subseteq \pi(V(\mathcal{G}))$*, if* $\lambda = \vartheta \mathbin{\hat{\circ}} \pi$*, then*

$$d_\vartheta(L') = \sum_{L \in \overleftarrow{\pi}(L')} d_\lambda(L). \tag{2.150}$$

Proof. Fix $L' \subseteq \pi(V(\mathcal{G}))$. The definition of $d_\vartheta(L')$, combined with Observation 2.54, yields

$$d_\vartheta(L') = \sum_{L' \subseteq E' \in \pi(\mathcal{G})} \vartheta(E') = \sum_{L' \subseteq E' \in \pi(\mathcal{G})} \sum_{\substack{E \in \mathcal{G} \\ \pi(E) = E'}} \lambda(E) = \sum_{\substack{E \in \mathcal{G} \\ L' \subseteq \pi(E)}} \lambda(E), \tag{2.151}$$

where the last step holds because each $E \in \mathcal{G}$ with $L' \subseteq \pi(E)$ appears exactly once in the second-to-last sum. On the other hand, the right-hand side of (2.150) is, by definition, equal to

$$\sum_{L \in \overleftarrow{\pi}(L')} d_\lambda(L) = \sum_{L \in \overleftarrow{\pi}(L')} \sum_{L \subseteq E \in \mathcal{G}} \lambda(E). \tag{2.152}$$

We also know that $|L| = |L'|$ for all $L \in \overleftarrow{\pi}(L')$. It then follows from $|\pi(E)| = |E|$ for all $E \in \mathcal{G}$ that $\pi|_E$ is a bijection, so there is a unique $L \subseteq E$ satisfying $\pi(L) = L'$ and $|\pi(L)| = |L|$. The conclusion is that

$$\sum_{L \in \overleftarrow{\pi}(L')} \sum_{L \subseteq E \in \mathcal{G}} \lambda(E) = \sum_{\substack{E \in \mathcal{G} \\ L' \subseteq \pi(E)}} \lambda(E),$$

which, together with (2.151), (2.152) and the fact that L' was arbitrary, completes the proof. □

The proof of Lemma 2.52 will be complete once we establish Lemma 2.57, which also admits a simple proof from Lemma 2.56.

Lemma 2.57. *Let* $\sigma > 0$ *and* $p > 0$*. Further let* $\mathcal{G}$ *be a hypergraph,* $\pi : V(\mathcal{G}) \to U$ *be a function satisfying*

$$|\pi(E)| = |E| \qquad \text{for every } E \in \mathcal{G}.$$

For all $\vartheta : \pi(\mathcal{G}) \to \mathbb{R}_{\geqslant 0}$*, we have*

$$\Lambda_p(\vartheta \mathbin{\hat{\circ}} \pi) \leqslant \Lambda_p(\vartheta).$$

Proof. Let $V = V(\mathcal{G})$, $\lambda = \vartheta \mathbin{\hat{\circ}} \pi$ and recall (2.2), the definition of Λ_p,

$$\Lambda_p(\vartheta) = \sum_{\substack{L' \subseteq \pi(V) \\ |L'| \geqslant 2}} d_\vartheta(L')^2 p^{-|L'|}.$$

As $|\pi(E)| = |E|$ for every $E \in \mathcal{G}$, we can use Lemma 2.56 to conclude that

$$\Lambda_p(\vartheta) = \sum_{\substack{L' \subseteq \pi(V) \\ |L'| \geqslant 2}} \left(\sum_{L \in \overleftarrow{\pi}(L')} d_\lambda(L) \right)^2 p^{-|L'|} \geqslant \sum_{\substack{L' \subseteq \pi(V) \\ |L'| \geqslant 2}} \sum_{L \in \overleftarrow{\pi}(L')} d_\lambda(L)^2 p^{-|L|} \tag{2.153}$$

where the last inequality uses that $d_\lambda(L)$ is always non-negative and also that $|L| = |L'|$ for $L \in \overleftarrow{\pi}(L')$.

Now, whenever $d_\lambda(L) > 0$ for $L \subseteq V$, we know that there is $L' \subseteq \pi(V)$ such that $L \in \overleftarrow{\pi}(L')$. We conclude that the rightmost part of (2.153) is at least

$$\Lambda_p(\vartheta) \geqslant \sum_{\substack{L' \subseteq \pi(V) \\ |L'| \geqslant 2}} \sum_{L \in \overleftarrow{\pi}(L')} d_\lambda(L)^2 p^{-|L|} \geqslant \sum_{\substack{L \subseteq V \\ |L| \geqslant 2}} d_\lambda(L)^2 p^{-|L|} = \Lambda_p(\lambda)$$

where the last step is the definition, and the proof is complete. □

2.7.2 Proof of Theorem 2.35

We are now ready to prove Theorem 2.35; for ease of reference, let us restate Theorem 2.23.

Theorem 2.23 (Campos and Samotij [34, modified Theorem E])**.** *Let $\mathcal{G}$ be a hypergraph with vertex set V. For all $\alpha, q \in \mathbb{R}$ satisfying $0 < q \leqslant \alpha < 1$, there exists a family $\mathcal{T} \subseteq 2^V$ and a function $\varphi : \mathcal{I}(\mathcal{G}) \to \mathcal{T}$ such that:*

(a) *For each $I \in \mathcal{I}(\mathcal{G})$, we have $\varphi(I) \subseteq I$.*

(b) *Each $T \in \mathcal{T}$ has at most qn/α elements, where $n = |V|$.*

(c) *For every $T \in \mathcal{T}$, there exists a hypergraph $\mathcal{C}_T$ with vertex set V that covers $\mathcal{G}$ and satisfies*

$$\mathbb{P}\big(L \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{G})\big) > (1-\alpha)^{|L|} q^{|L|} \tag{2.154}$$

for all $L \notin \mathcal{C}_T$; moreover, for all $I \in \mathcal{I}(\mathcal{G})$ such that $T = \varphi(I)$, we have $I \in \mathcal{I}(\mathcal{C}_T)$.

Proof of Theorem 2.35. Apply Theorem 2.23 with $\mathcal{G} = \mathcal{J}'$, where

$$\mathcal{J}' = \Big\{ L \subseteq V : \pi_v\big(\mathcal{H}[L]\big) \cup \mathcal{F} \text{ is } (p, \sigma' + \eta\sigma)\text{-Janson} \Big\},$$

and parameters q and $\alpha = 1/2$ to obtain $\mathcal{T}$ and φ. Now, for each $T \in \mathcal{T}$, there is $\mathcal{C}_T$ satisfying item (c) in Theorem 2.23, so we let $\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_{=s}$ be the edges of $\langle \mathcal{C}_T \rangle$ of size s. Since $\mathcal{C}'_T$ is s-uniform and $p \leqslant 1/(2^{11} r s^2)$ by assumption, we can apply Theorem 2.15 with $\mathcal{G} = \mathcal{C}'_T$ and $\zeta = 2^{-8} r^{-1}$. As in the proof of Theorem 2.26 we obtain φ, ψ_T, ϕ_T that satisfy

(i) For each $I \in \mathcal{I}(\mathcal{J}')$, we have $\phi_T(I) \subseteq I \subseteq \psi_T(\phi_T(I))$, for $T = \varphi(I)$.

(ii) Each $S \in \mathcal{S}_T$ has at most $2^{11} r p s^2 n$ elements.

(iii) For every $S \in \mathcal{S}_T$, letting $X = \psi_T(S)$, $\mathcal{C}'_T[X]$ is not $(p, 2^{-8} r^{-1} p |X|)$-Janson.

Setting $f(I) = X = \psi_T(S)$ for $T = \varphi(I)$ and $S = \phi_T(I)$, we define

$$\mathcal{X} = \big\{ f(I) : I \in \mathcal{I}(\mathcal{J}') \big\},$$

which, by item (i) and the fact that I was arbitrary, is a definition that satisfies item (1) in Theorem 2.35.

As in the proofs of Theorem 2.21 and Theorem 2.26 (cf. (2.67), (2.68) and (2.70)), to show that the size of $\mathcal{X}$ is suitable, we count $X \in \mathcal{X}$ by choosing $T \in \mathcal{T}$ and then $S \in \mathcal{S}_T$. We therefore obtain

$$|\mathcal{X}| \leqslant \sum_{T\in\mathcal{T}} |\mathcal{S}_T| \leqslant \sum_{m=0}^{2^{11}ps^2n} \binom{n}{m} \sum_{t=0}^{2qn} \binom{n}{t} \leqslant \left(\frac{2}{q}\right)^{2qn} \tag{2.155}$$

combining item (ii) above with item (b) in Theorem 2.23 and

$$2^{11}rps^2n \leqslant 2qn \leqslant \frac{n}{4}.$$

The bound in (2.155) matches (2.128) in the statement, so it only remains to show that item (2) holds.

Now, assume by contradiction that there exists $I \in \mathcal{I}(\mathcal{J}')$ such that

$$\begin{gathered} f(I) = X \in \mathcal{X} \text{ satisfies } |X| \geqslant n/(8r), \text{ and} \\ \pi\big(\mathcal{H}[Y]\big) \text{ is } (p,\sigma)\text{-Janson for every } Y \subseteq X \text{ with } |Y| \geqslant |X| - 2^{-8}r^{-1}n. \end{gathered} \tag{$*$}$$

This is the converse of item (2) in Theorem 2.35, so contradicting it will complete the proof.

We want to apply Lemma 2.53 with $\mathcal{G} = \pi\big(\mathcal{H}[X]\big)$ and $\beta = 2^{-9}r^{-1}$. To do that, we first verify that this hypergraph satisfies the assumptions in the lemma.

Claim 2.58. *For all*

$$W \subseteq \pi(X) \qquad \textit{with} \qquad |W| \geqslant (1 - 2^{-9}r^{-1})|\pi(X)|, \tag{2.156}$$

the hypergraph $\pi\big(\mathcal{H}[X]\big)[W]$ *is* (p,σ)*-Janson.*

Proof. Fix W satisfying (2.156) and observe that letting $Y = \pi^{-1}(W) \cap X$, we have

$$\pi\big(\mathcal{H}[Y]\big) \subseteq \pi\big(\mathcal{H}[Y]\big)[W] \subseteq \pi\big(\mathcal{H}[X]\big)[W]$$

where the first containment holds because $\pi(E) \subseteq W$, for all $E \subseteq Y$, by our choice of $Y \subseteq \pi^{-1}(W)$, and the second holds since $Y \subseteq X$.

As $\pi(X \setminus Y) = \pi(X) \setminus W$ and we assumed that $|L| \leqslant 2|\pi(L)|$ for every $L \subseteq V$, we have that

$$|Y| = |X| - |X \setminus Y| \geqslant |X| - 2|\pi(X \setminus Y)| = |X| - 2|\pi(X) \setminus W| \geqslant |X| - 2^{-8}r^{-1}n. \tag{2.157}$$

It follows from (2.157) and ($*$) that $\pi\big(\mathcal{H}[Y]\big)$ is (p,σ)-Janson, but this is an increasing property by Observation 2.5, so $\pi\big(\mathcal{H}[X]\big)[W]$ is also (p,σ)-Janson. ∎

Applying Lemma 2.53 with $\mathcal{G} = \pi\big(\mathcal{H}[X]\big)$ and $\beta = 2^{-9}r^{-1}$, we obtain $\mu : \pi\big(\mathcal{H}[X]\big) \to \mathbb{R}_{\geqslant 0}$ satisfying

$$e(\mu) = \sqrt{\sigma}, \qquad \Lambda_p(\mu) < 1 \tag{2.158}$$

and

$$\sum_{u\in\pi(X)} d_\mu(u)^2 \leqslant \frac{2s^2e(\mu)^2}{\beta|\pi(X)|} \leqslant \frac{2^{10}rs^2e(\mu)^2}{|\pi(X)|} \leqslant \frac{2^{14}r^2s^2e(\mu)^2}{n}, \tag{2.159}$$

where the last inequality follows from the assumptions that $|L|/2 \leqslant |\pi(L)|$ for every $L \subseteq V$ and $|X| \geqslant n/(8r)$.

Let $\nu : \mathcal{H}[X] \to \mathbb{R}_{\geqslant 0}$ be the pullback measure of μ with respect to π, i.e. $\nu = \mu \mathbin{\hat{\circ}} \pi$. As π satisfies

$$|\pi(E)| = |E| \qquad \text{for all } E \in \mathcal{H}[X] \subseteq \mathcal{H}$$

by (2.127), we can apply Lemma 2.52 with $\mathcal{G} = \mathcal{H}[X]$ and $\vartheta = \mu$ to conclude that

$$\Lambda_p(\nu) < \frac{e(\nu)^2}{\sigma}. \tag{2.160}$$

Now take $T = \varphi(I)$, and recall that $\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_{=s}$, where $\mathcal{C}_T$ is the cover given by item (c) in Theorem 2.23. Let ν' be the restriction of the measure ν to $\mathcal{H}[X] \cap \mathcal{C}'_T[X]$, that is, for each $E \in \mathcal{H}[X]$, let

$$\nu'(E) = \begin{cases} \nu(E) & \text{if } E \in \mathcal{C}'_T[X], \\ 0 & \text{otherwise.} \end{cases}$$

Claim 2.59. *The measure ν' satisfies*

$$e(\nu') < \frac{e(\nu)}{2}.$$

Proof. Indeed, we have

$$\frac{e(\nu)^2}{\sigma} > \Lambda_p(\nu) \geqslant \Lambda_p(\nu') \geqslant \frac{2^8 r e(\nu')^2}{p|X|} \geqslant \frac{4e(\nu')^2}{\sigma}$$

first by (2.160), second because $\nu \geqslant \nu'$ and $\Lambda_p(\cdot)$ is monotone increasing, then since the hypergraph $\mathcal{C}'_T[X]$ is not $(p, 2^{-8}r^{-1}p|X|)$-Janson by item (iii) of Theorem 2.15 and our choice of $\zeta = 2^{-8}r^{-1}$, and finally because $\sigma = 2^{-6}r^{-1}pn$ by assumption. ∎

We now define the measure $\nu'' = \nu - \nu'$, which corresponds to the restriction of ν to the hypergraph $\mathcal{H}' := \mathcal{H}[X] \setminus \mathcal{C}'_T = \mathcal{H}[X] \setminus \mathcal{C}'_T[X]$. By Claim 2.59, we have

$$e(\nu'') = e(\nu) - e(\nu') > \frac{e(\nu)}{2}. \tag{2.161}$$

Also define μ'' to be the restriction of μ to $\pi(\mathcal{H}')$. Applying Observation 2.54 with $\vartheta = \mu$ and $\nu = \mu \mathbin{\hat{\circ}} \pi$, then using (2.161) and Lemma 2.55, yields

$$e(\mu'') = \sum_{E \in \pi(\mathcal{H}')} \mu(E) = \sum_{E \in \pi(\mathcal{H}')} \sum_{\substack{E_0 \in \mathcal{H} \\ \pi(E_0) = E}} \nu(E_0) \geqslant \sum_{E \in \mathcal{H}'} \nu''(E) = e(\nu'') > \frac{e(\nu)}{2} = \frac{e(\mu)}{2}. \tag{2.162}$$

With the goal of defining a random measure μ''_q, let, for all $E \in \pi(\mathcal{H}')$,

$$P'_q(E) = \mathbb{P}\big(E \in \pi(\mathcal{H}'[V_q]) \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big). \tag{2.163}$$

Claim 2.60. *For all $E \in \pi(\mathcal{H}')$,*

$$P'_q(E) \geqslant \mathbb{P}\big(E_0 \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big) > \left(\frac{q}{2}\right)^s \tag{2.164}$$

where $E_0 \in \mathcal{H}'$ is fixed and satisfies $\pi(E_0) = E$.

Proof. The first inequality in (2.164) follows from the fact that if $E_0 \subseteq V_q$ for $E_0 \in \mathcal{H}'$, then we have $\pi(E_0) \in \pi(\mathcal{H}')[V_q]$. Also notice that, by the same argument used to establish (2.75), the hypergraphs $\mathcal{H}'$ and $\mathcal{C}_T$ are disjoint. Therefore, we conclude that

$$\mathbb{P}\big(E_0 \subseteq V_q \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big) > \left(\frac{q}{2}\right)^s$$

by (2.154) in item (c) of Theorem 2.23. ∎

Now, let $\gamma = \sqrt{8\eta} > 0$ (a choice made with foresight) and $\mu''_q : \pi(\mathcal{H}') \to \mathbb{R}_{\geqslant 0}$ be defined by

$$\mu''_q(E) = \mu''(E)\,\frac{\gamma \cdot \mathbb{1}\big[E \in \pi(\mathcal{H}'[V_q])\big]}{P'_q(E)}.$$

Also define the extension $\bar{\mu}''_q$ supported on $\pi_v(\mathcal{H}'[X_q])$, where $X_q = V_q \cap X$, by

$$\bar{\mu}''_q(E \cup \{v\}) = \mu''_q(E) \qquad \text{for all } E \in \pi(\mathcal{H}'[X_q]).$$

Our goal now is to show that analysing $\bar{\mu}''_q$ for our choice of γ contradicts $I \in \mathcal{I}(\mathcal{J}')$.

To reason about properties of $\mathcal{J}'$, we define a measure ρ supported on $\mathcal{F}$. If $\sigma' > 0$, then we can apply Observation 2.4 with $\mathcal{G} = \mathcal{F}$ and $y = \sqrt{\sigma'}$, since $\mathcal{F}$ is (p, σ')-Janson, to obtain $\rho : \mathcal{F} \to \mathbb{R}_{\geqslant 0}$ satisfying

$$\Lambda_p(\rho) < \frac{e(\rho)^2}{\sigma'} \qquad \text{and} \qquad e(\rho) = \sqrt{\sigma'}. \tag{2.165}$$

If, on the other hand, $\sigma' = 0$, then we take the measure

$$\rho = 0. \tag{2.166}$$

Regardless of the value of σ', or the choice of ρ as either (2.165) or (2.166), we have

$$\Lambda_p(\rho) < 1 \qquad \text{and} \qquad e(\rho) = \sqrt{\sigma'}, \tag{2.167}$$

which, together on being supported on $\mathcal{F}$, are the only properties that we will use of ρ.

Our goal is to obtain inequalities bounding $\Lambda_p(\rho + \bar{\mu}''_q)$ from $e(\rho + \bar{\mu}''_q)^2$, so we start easily, relating the expected value of $e(\bar{\mu}''_q)$ to $e(\mu'')$.

Claim 2.61.

$$\mathbb{E}\big[e(\bar{\mu}''_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big] = \gamma e(\mu'').$$

Proof. The definitions of $e(\bar{\mu}''_q)$, $\bar{\mu}''_q$ and μ''_q,

$$e(\bar{\mu}''_q) = \sum_{E \cup \{v\} \in \pi_v(\mathcal{H}')} \bar{\mu}''_q\big(E \cup \{v\}\big) = \sum_{E \in \pi(\mathcal{H}')} \mu''(E)\,\frac{\gamma \cdot \mathbb{1}\big[E \in \pi(\mathcal{H}'[V_q])\big]}{P'_q(E)},$$

imply that

$$\mathbb{E}\big[e(\bar{\mu}_q'') \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big] = \sum_{E \in \pi(\mathcal{H}')} \gamma \mu''(E) = \gamma e(\mu''),$$

since, for all $E \in \pi(\mathcal{H}')$, we have that

$$\mathbb{E}\big[\mathbb{1}\big[E \in \pi(\mathcal{H}'[V_q])\big] \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big] = P_q'(E) \tag{2.168}$$

from the definition of $P_q'(E)$, (2.163). ■

Combining Claim 2.61 with (2.162), we have

$$\mathbb{E}\big[e(\bar{\mu}_q'') \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big] > \gamma \frac{e(\mu)}{2}. \tag{2.169}$$

Recall that the definition of $d_\rho(L)$ for any set $L \subseteq U$ is

$$d_\rho(L) = \sum_{L \subseteq E \in \mathcal{F}} \rho(E).$$

As ρ is supported over $\mathcal{F}$, we have $\rho(E) = 0$ when $E \not\subseteq U = V(\mathcal{F})$. We can use this observation to obtain the expansion

$$\Lambda_p(\rho + \bar{\mu}_q'') = \sum_{\substack{L \subseteq U \\ |L| \geqslant 2}} d_\rho(L)^2 p^{-|L|} + 2 \sum_{\substack{L \subseteq U \\ |L| \geqslant 2}} d_\rho(L) d_{\bar{\mu}_q''}(L) p^{-|L|} + \sum_{\substack{L \subseteq U \cup \{v\} \\ |L| \geqslant 2}} d_{\bar{\mu}_q''}(L)^2 p^{-|L|} \tag{2.170}$$

where the first two terms range over $L \subseteq U$ instead of $L \subseteq U \cup \{v\}$ because if $v \in L$, then $d_\rho(L) = 0$ as $v \notin V(\mathcal{F})$ by assumption. The first sum in (2.170) is now exactly $\Lambda_p(\rho)$, which we can immediately bound with (2.167), but we need some simple claims before we analyse the other two sums. We start with an observation that relates the degrees $d_{\bar{\mu}_q''}$ to the degrees $d_{\mu_q''}$.

Claim 2.62. *For all $L \subseteq U \cup \{v\}$, we have*

$$d_{\bar{\mu}_q''}(L) = d_{\mu_q''}(L \setminus \{v\}).$$

Proof. Let $E_v = E \cup \{v\}$ for each $E \in \pi(\mathcal{H}')$. The definitions of $d_{\bar{\mu}_q''}$ and $\bar{\mu}_q''$ imply that, for every $L \subseteq U \cup \{v\}$,

$$d_{\bar{\mu}_q''}(L) = \sum_{L \subseteq E_v \in \pi_v(\mathcal{H}')} \bar{\mu}_q''(E_v) = \sum_{L \subseteq E_v \in \pi_v(\mathcal{H}')} \mu_q''(E_v \setminus \{v\}) = \sum_{L \setminus \{v\} \subseteq E \in \pi(\mathcal{H}')} \mu_q''(E) = d_{\mu_q''}(L \setminus \{v\}),$$

where the last equality is the definition of $d_{\mu_q''}(L \setminus \{v\})$. ■

We can now use Claim 2.62 to relate $\Lambda_p(\bar{\mu}_q'')$ and $\Lambda_p(\mu_q'')$.

Claim 2.63.

$$\Lambda_p(\bar{\mu}_q'') = \left(1 + \frac{1}{p}\right) \Lambda_p(\mu_q'') + \frac{1}{p^2} \sum_{u \in U} d_{\mu_q''}(u)^2. \tag{2.171}$$

Proof. First, recall the definition of $\Lambda_p(\bar{\mu}_q'')$,

$$\Lambda_p(\bar{\mu}_q'') = \sum_{\substack{L \subseteq U \cup \{v\} \\ |L| \geqslant 2}} d_{\bar{\mu}_q''}(L)^2 p^{-|L|}.$$

Splitting that sum according to whether $v \in L$ or not, we obtain

$$\Lambda_p(\bar{\mu}_q'') = \sum_{\substack{L \subseteq U \\ |L| \geqslant 2}} d_{\mu_q''}(L)^2 p^{-|L|} + \sum_{u \in U} d_{\mu_q''}(u)^2 p^{-2} + \sum_{\substack{L \subseteq U \\ |L| \geqslant 2}} d_{\bar{\mu}_q''}\big(L \cup \{v\}\big)^2 p^{-|L|-1}, \tag{2.172}$$

where the second term corresponds to $L = \{u, v\}$ for $u \in U$, that is, the case $v \in L$ and $|L| = 2$. Applying Claim 2.62 in the third sum of (2.172) and using the definition of $\Lambda_p(\mu_q'')$ yields (2.171). ∎

Now, we prove a deterministic upper bound for $\Lambda_p(\mu_q'')$ in terms of $\Lambda_p(\mu'')$. We do not optimise this bound, like the analogous one in Section 2.4 (cf. Claim 2.31), since that will not be necessary in this proof.

Claim 2.64.

$$\Lambda_p(\mu_q'') \leqslant \gamma^2 \Big(\frac{q}{2}\Big)^{-2s} \Lambda_p(\mu'').$$

Proof. Recall that

$$P_q'(E) > \Big(\frac{q}{2}\Big)^s$$

for all $E \in \pi(\mathcal{H}')$, by Claim 2.60, so by definition of $d_{\mu_q''}$ and μ_q'', we deterministically have that

$$d_{\mu_q''}(L)^2 = \Bigg(\sum_{L \subseteq E \in \pi(\mathcal{H}')} \mu''(E) \frac{\gamma \cdot \mathbb{1}\big[E \in \pi(\mathcal{H}'[V_q])\big]}{P_q'(E)} \Bigg)^2 \leqslant \gamma^2 \Big(\frac{q}{2}\Big)^{-2s} d_{\mu''}(L)^2, \tag{2.173}$$

holds for all $L \subseteq U$, by ignoring the indicators. The inequality in the claim now follows

$$\Lambda_p(\mu_q'') = \sum_{\substack{L \subseteq U \\ |L| \geqslant 2}} d_{\mu_q''}(L)^2 p^{-|L|} \leqslant \gamma^2 \Big(\frac{q}{2}\Big)^{-2s} \sum_{\substack{L \subseteq U \\ |L| \geqslant 2}} d_{\mu''}(L)^2 p^{-|L|} = \gamma^2 \Big(\frac{q}{2}\Big)^{-2s} \Lambda_p(\mu'')$$

where we used the definition of $\Lambda_p(\cdot)$. ∎

We will now combine the bound given by Lemma 2.53 for the sum of the square of the μ-degrees, (2.159), with Claims 2.63 and 2.64 and another simple calculation to complete the proof of a deterministic inequality relating $\Lambda_p(\bar{\mu}_q'')$ and $e(\mu)^2$.

Claim 2.65.

$$\Lambda_p(\bar{\mu}_q'') < \frac{\gamma^2}{2\sqrt{\eta}}.$$

Proof. Repeating what we did in (2.173), we have

$$\sum_{u \in U} d_{\mu_q''}(u)^2 \leqslant \gamma^2 \Big(\frac{q}{2}\Big)^{-2s} \sum_{u \in U} d_{\mu''}(u)^2 \leqslant \gamma^2 \Big(\frac{q}{2}\Big)^{-2s} \sum_{u \in U} d_{\mu}(u)^2 \tag{2.174}$$

where the last step is using that $\mu'' \leqslant \mu$. Now, recall that the support of μ is $\pi(\mathcal{H}[X])$, so if an edge $E \subseteq U$ is not fully contained in $\pi(X)$, then $\mu(E) = 0$. As so, $d_\mu(u) = 0$ if $u \notin \pi(X)$ and thus

$$\sum_{u\in U} d_\mu(u)^2 = \sum_{u\in\pi(X)} d_\mu(u)^2.$$

We conclude that (2.174) is at most

$$\sum_{u\in U} d_{\mu''_q}(u)^2 \leqslant \gamma^2\Big(\frac{q}{2}\Big)^{-2s} \sum_{u\in\pi(X)} d_\mu(u)^2 \leqslant \gamma^2\Big(\frac{q}{2}\Big)^{-2s} \frac{2^{14}r^2s^2e(\mu)^2}{n} \tag{2.175}$$

where the last step is

$$\sum_{u\in\pi(X)} d_\mu(u)^2 \leqslant \frac{2^{14}r^2s^2e(\mu)^2}{n},$$

the inequality in (2.159).

Combining (2.175) with Claims 2.63 and 2.64 thus yields

$$\Lambda_p(\bar{\mu}''_q) \leqslant \gamma^2\Big(\frac{2}{q}\Big)^{2s}\left(\Big(1+\frac{1}{p}\Big)\Lambda_p(\mu'') + \frac{2^{14}r^2s^2e(\mu)^2}{p^2n}\right). \tag{2.176}$$

Also recall that we chose

$$p \leqslant \frac{q}{2^{10}r^2s^2}, \qquad q < \frac{1}{8}, \qquad \sigma = 2^{-6}r^{-1}pn, \qquad \text{and} \qquad \eta = p^4\Big(\frac{q}{2}\Big)^{4s},$$

in (2.126), and that

$$\Lambda_p(\mu'') \leqslant \Lambda_p(\mu) < \frac{e(\mu)^2}{\sigma} \qquad \text{and} \qquad e(\mu) = \sqrt{\sigma}$$

by $\mu'' \leqslant \mu$ and (2.158), so (2.176) is at most

$$\Lambda_p(\bar{\mu}''_q) < \gamma^2\Big(\frac{2}{q}\Big)^{2s}\left(1+\frac{1}{p}+\frac{2^8rs^2}{p}\right) \leqslant \gamma^2\Big(\frac{2}{q}\Big)^{2s}\frac{2^{10}rs^2}{p} \leqslant \frac{\gamma^2}{2\sqrt{\eta}}$$

as desired. ■

Replacing the bound of Claim 2.65 in (2.170) and then taking the conditional expectation with respect to $\{V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J}')\}$ yields

$$\mathbb{E}\big[\Lambda_p(\rho+\bar{\mu}''_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J}')\big] < 1 + 2\sum_{\substack{L\subseteq U\\ |L|\geqslant 2}} d_\rho(L)\frac{\mathbb{E}\big[d_{\bar{\mu}''_q}(L) \mid V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J}')\big]}{p^{|L|}} + \frac{\gamma^2}{2\sqrt{\eta}}. \tag{2.177}$$

In particular, this inequality motivates our final claim.

Claim 2.66. *If $L \subseteq U$, then*

$$\mathbb{E}\big[d_{\bar{\mu}''_q}(L) \mid V_q \in \mathcal{I}(\underline{\partial}_T\mathcal{J}')\big] = \gamma d_{\mu''}(L).$$

Proof. Note that $v \notin L$ if $L \subseteq U$. It then follows from Claim 2.62 and the definition of $d_{\mu''_q}(L)$ that

$$d_{\bar{\mu}''_q}(L) = d_{\mu''_q}(L) = \sum_{L \subseteq E \in \pi(\mathcal{H}')} \mu''(E) \frac{\gamma \cdot \mathbb{1}\big[E \in \pi(\mathcal{H}'[V_q])\big]}{P'_q(E)},$$

so taking the conditional expectation yields, for all $L \subseteq U$,

$$\mathbb{E}\big[d_{\bar{\mu}''_q}(L) \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big] = \sum_{L \subseteq E \in \pi(\mathcal{H}')} \gamma \mu''(E) = \gamma d_{\mu''}(L)$$

since

$$\mathbb{E}\big[\mathbb{1}\big[E \in \pi(\mathcal{H}'[V_q])\big] \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big] = P'_q(E),$$

(2.168), holds for all $E \in \pi(\mathcal{H}')$. ∎

Substituting the bound in Claim 2.66 in the middle sum of (2.177) and applying the Cauchy–Schwarz inequality, we obtain

$$\gamma \sum_{\substack{L \subseteq U \\ |L| \geqslant 2}} \frac{d_\rho(L) d_{\mu''}(L)}{p^{|L|}} \leqslant \gamma \left(\sum_{\substack{L \subseteq U \\ |L| \geqslant 2}} \frac{d_\rho(L)^2}{p^{|L|}} \right)^{1/2} \left(\sum_{\substack{L \subseteq U \\ |L| \geqslant 2}} \frac{d_{\mu''}(L)^2}{p^{|L|}} \right)^{1/2} = \gamma \sqrt{\Lambda_p(\rho)} \sqrt{\Lambda_p(\mu'')} \leqslant \gamma,$$

which, replaced back in (2.177), yields

$$\mathbb{E}\big[\Lambda_p(\rho + \bar{\mu}''_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big] < 1 + 2\gamma + \frac{\gamma^2}{2\sqrt{\eta}}. \tag{2.178}$$

Now, to bound the expected edge mass of the measure $\rho + \bar{\mu}''_q$, observe that

$$\mathbb{E}\big[e(\rho + \bar{\mu}''_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big] = e(\rho) + \mathbb{E}\big[e(\bar{\mu}''_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big] \geqslant \sqrt{\sigma'} + \frac{\gamma\sqrt{\sigma}}{2},$$

where the last inequality is due to (2.167) and (2.169). Jensen's inequality thus implies that

$$\mathbb{E}\big[e(\rho + \bar{\mu}''_q)^2 \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big] \geqslant \left(\sqrt{\sigma'} + \frac{\gamma\sqrt{\sigma}}{2}\right)^2 > (1 + 4\gamma)\left(\sigma' + \frac{\gamma^2 \sigma}{8}\right) \tag{2.179}$$

because $\sigma \geqslant 16\sigma'$ by assumption, and we can choose $\gamma < 1/4$. We fix $\gamma = \sqrt{8\eta}$, which is less than $1/4$ because

$$\eta = p^4 \left(\frac{q}{2}\right)^{4s} < \frac{1}{2^7}.$$

This choice, replaced in (2.178), implies that

$$\mathbb{E}\big[\Lambda_p(\rho + \bar{\mu}''_q) \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big] < 1 + 4\gamma = \frac{(1 + 4\gamma)(\sigma' + \gamma^2\sigma/8)}{\sigma' + \eta\sigma} \leqslant \frac{\mathbb{E}\big[e(\rho + \bar{\mu}''_q)^2 \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big]}{\sigma' + \eta\sigma}$$

and the last step is (2.179).

To reach a contradiction, observe that if $V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')$ then $V_q \in \mathcal{I}(\mathcal{J}')$ follows from Observation 2.25, and hence $\pi_v(\mathcal{H}'[V_q]) \cup \mathcal{F}$ is not $(p, \sigma' + \eta\sigma)$-Janson, by the definition of $\mathcal{J}'$.

In particular, it follows that

$$\Lambda_p(\rho + \bar{\mu}_q'') \geqslant \frac{e(\rho + \bar{\mu}_q'')^2}{\sigma' + \eta\sigma}$$

since this measure is supported on $\pi_v(\mathcal{H}'[V_q]) \cup \mathcal{F}$, so we cannot have

$$\mathbb{E}\big[\Lambda_p(\rho + \bar{\mu}_q'') \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big] < \frac{\mathbb{E}\big[e(\rho + \bar{\mu}_q'')^2 \mid V_q \in \mathcal{I}(\underline{\partial}_T \mathcal{J}')\big]}{\sigma' + \eta\sigma}$$

like we previously established, and we have a contradiction. We conclude that item (2) in Theorem 2.35 holds, and the proof is complete. □

Chapter 3

On the independence number of sparser random Cayley graphs

This chapter is adapted from [29], which is joint work with Campos and Marciano that appeared in *J. Lond. Math. Soc.* Its main goal is to prove the following result.

Theorem 1.3. *Let n be a prime number and let $p = p(n)$ be such that $(\log n)^{-1/80} \leqslant p \leqslant 1/2$. The random Cayley sum graph G_p of $\mathbb{Z}_n$ satisfies*

$$\alpha(G_p) = \big(2 + o(1)\big) \log_{\frac{1}{1-p}} n \tag{3.1}$$

with high probability as $n \to \infty$.

In the next section, we give a detailed sketch of our proof strategy for Theorem 1.3, and a proof of our fingerprint theorem. Then, in Section 3.2, we give a simple proof of a weaker version of our "Freĭman's lemma via few translates" theorem that nevertheless contains some of the main ideas required for the full proof of Theorem 1.4. Section 3.3 is dedicated to improving the constant to $(1 - 5\gamma)/2$, assuming two additional technical lemmas, which we prove in Sections 3.4 and 3.5. In Section 3.6, we derive our supersaturation result from Theorem 1.4, and in Sections 3.8 and 3.9 we complete the proof of the main theorem in this chapter.

3.1 Overview of the proof

Throughout, $n \in \mathbb{N}$ will always be a sufficiently large prime number; we will also adopt the standard convention of omitting floors and ceilings whenever they are not essential. Let $k \in \mathbb{N}$ be the bound that we want to show for the independence number, and let S be a p-random subset of $\mathbb{Z}_n$. Denoting $\mathcal{A} = \binom{\mathbb{Z}_n}{k}$, we will show that

$$\mathbb{P}\left(\exists\, A \in \mathcal{A} : A \hat{+} A \subseteq S^c\right) \to 0$$

as $n \to \infty$, which is equivalent to proving that $\alpha(G_p) < k$ with high probability.

We will follow Green's general approach of partitioning $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3$ where each

sub-collection is defined based on the doubling $\sigma[A]$,

$$\begin{aligned}\mathcal{A}_1 &= \{A \in \mathcal{A} : \sigma[A] \leqslant k^c\}\,,\\ \mathcal{A}_2 &= \{A \in \mathcal{A} : k^c < \sigma[A] \leqslant \delta k/10\}\,, \text{ and}\\ \mathcal{A}_3 &= \{A \in \mathcal{A} : \delta k/10 < \sigma[A]\}\,,\end{aligned}$$

for $c = 1/40$ and some $\delta = o(1)$, and handle each one differently. Explicitly, we use the union bound to deduce

$$\mathbb{P}\left(\exists\, A \in \mathcal{A} : A \hat{+} A \subseteq S^{\mathsf{c}}\right) \leqslant \sum_{i=1}^{3} \mathbb{P}\left(\exists\, A \in \mathcal{A}_i : A \hat{+} A \subseteq S^{\mathsf{c}}\right).$$

It turns out that we can bound the terms related to $\mathcal{A}_2$ and $\mathcal{A}_3$ using the same techniques of Green [72] and Green and Morris [73]; we therefore defer working out the details regarding these terms to Section 3.8. As we mentioned in Section 1.2.1, one of the key new ideas to prove Theorem 1.3 is in how we handle the term related to the collection $\mathcal{A}_1$: we show that there exists a family $\mathcal{F}$ of fingerprints such that it suffices to consider only the events $\{F \hat{+} F \subseteq S^{\mathsf{c}}\}_{F\in\mathcal{F}}$ instead of the collection $\{A \hat{+} A \subseteq S^{\mathsf{c}}\}_{A\in\mathcal{A}_1}$. Trivially, $\mathcal{F} = \mathcal{A}_1$ is such a family, so, to make this strategy work and improve upon taking the union bound over all A, we must choose $\mathcal{F}$ in a more clever way.

The first property of these fingerprints that we require is that $|\mathcal{F}|$ is sufficiently small, where small vaguely means that a union bound over all $F \in \mathcal{F}$ works. One direct way to achieve that is picking each fingerprint F to be small, but there is a subtle trade-off between the size of F and the upper bound on the probability term $(1-p)^{|F\hat{+}F|}$. To circumvent this issue, we can use the fact that we have a bound (even if a polynomially large one) on the doubling of each $A \in \mathcal{A}_1$. In this setting, Freĭman's theorem [60] says that A is contained in a generalised arithmetic progression P such that both the size s and dimension d of P are small. Recall that a d-dimensional generalised arithmetic progression is a set of the form

$$\left\{a_0 + \sum_{i=1}^{d} w_i a_i : w_i \in \mathbb{Z},\, 0 \leqslant w_i < \ell_i\right\}$$

for some differences $a_1, \dots, a_d \in \mathbb{Z}_n$ and side-lengths $\ell_1, \dots, \ell_d \in \mathbb{N}$, and P is proper when every possible choice of $\{w_1, \dots, w_d\}$ corresponds to a distinct element of P. Therefore, instead of choosing F directly from $\mathbb{Z}_n$, we first choose P and then select F inside P.

Now we can state the two requirements for $F \in \mathcal{F}$ in more detail. Let P be the generalised arithmetic progression given by Freĭman's theorem for A.

First, F should be small enough compared to $|F \hat{+} F|$ for a union bound over choices of $F \subseteq P$ to work:

$$\mathbb{P}\big(\exists\, F \subseteq P : F \hat{+} F \subseteq S^{\mathsf{c}}\big) \leqslant \binom{s}{|F|}(1-p)^{|F\hat{+}F|} \to 0 \tag{3.2}$$

as $n \to \infty$. To get to this point, however, we must also choose this generalised arithmetic progression in a previous union bound. Second, $F \hat{+} F$ must be sufficiently large to pay for the number of choices for the generalised arithmetic progression P.

As we can count generalised arithmetic progressions with dimension d and size s by choosing the $a_0, a_1, \dots, a_d$ and $\ell_1, \dots, \ell_d$, there are at most $(ns)^{d+1}$ of them. Temporarily ignoring the number of choices for the fingerprint inside the progression (which we already dealt with in (3.2)), this amounts to requiring that F satisfies

$$(ns)^{d+1}(1-p)^{|F\hat{+}F|} \to 0$$

as $n \to \infty$. While these conditions may seem too strict – small F with large $F \hat{+} F$ for every A – this is exactly what we are able to do.

To state the actual theorem, we need to relate the dimension of the progression to some notion of dimension for A. The notion that we use is the Freĭman dimension of A, but to state its definition, we must first define what are Freĭman homomorphisms and Freĭman isomorphisms. A Freĭman homomorphism is a function $\phi : A \to Y$ such that for every $a_1, a_2, b_1, b_2 \in A$,

$$a_1 + b_1 = a_2 + b_2 \quad \text{implies} \quad \phi(a_1) + \phi(b_1) = \phi(a_2) + \phi(b_2).$$

A Freĭman isomorphism is a bijection ϕ such that both ϕ and its inverse ϕ^{-1} are Freĭman homomorphisms. The Freĭman dimension of A, $\dim_{\mathrm{F}}(A)$, is then defined to be the largest $d \in \mathbb{N}$ for which there is a full rank subset of $\mathbb{Z}^d$ that is Freĭman isomorphic to A[1].

It is furthermore useful to define the robustness of the Freĭman dimension of A: we say that A has ε-robust Freĭman dimension d if $\dim_{\mathrm{F}}(A) \geqslant d$ and there is no $A' \subseteq A$ such that $|A'| \geqslant (1-\varepsilon)|A|$ and $\dim_{\mathrm{F}}(A') < d$. In words, this means that the Freĭman dimension of A is at least d even if we remove an ε proportion of its elements.

With these definitions, we can state the fingerprint theorem that we prove.

Theorem 3.1. *Let n be a large enough prime and let $k, d \in \mathbb{N}$. For every $0 < \gamma < 1/2$, there exists $C = C(\gamma) > 0$ such that the following holds for all $m \geqslant (d+1)k/2$ and $C/k < \varepsilon < \gamma$. For each d-dimensional generalised arithmetic progression $P \subseteq \mathbb{Z}_n$, there exists a collection $\mathcal{F} = \mathcal{F}_{k,m,\varepsilon}(P)$ of subsets of P satisfying:*

(1) For every $F \in \mathcal{F}$, we have

$$|F| \leqslant C\varepsilon^{-1}\sqrt{m \log m} \quad \text{and} \quad |F \hat{+} F| \geqslant \frac{(1-\gamma)(d+1)k}{2}. \tag{3.3}$$

(2) For all $A \in \binom{P}{k}$ with $|A \hat{+} A| \leqslant m$ and ε-robust Freĭman dimension d, there exists $F \in \mathcal{F}$ such that $F \subseteq A$.

We will deduce this theorem from the following supersaturation result.

Theorem 3.2. *For every $0 < \gamma < 1$, there exists a constant $c = c(\gamma) > 0$ such that, for every sufficiently large set $A \subseteq \mathbb{Z}_n$, every $d \in \mathbb{N}$ and every $0 < \varepsilon < \gamma$, the following holds. If A has ε-robust Freĭman dimension d and $Y \subseteq A + A$ satisfies*

$$\big|\{(a_1, a_2) \in A^2 : a_1 + a_2 \notin Y\}\big| \leqslant c\varepsilon|A|^2, \tag{3.4}$$

[1]We say that $A \subseteq \mathbb{R}^d$ has full rank if $\mathrm{rank}(A) = d$; recall that for $A \subseteq \mathbb{R}^d$, we write $\mathrm{rank}(A)$ for the minimum dimension of an affine subspace that contains A.

then $|Y| \geqslant (1-\gamma)(d+1)|A|/2$.

In words, whenever $Y \approx A + A$ in the sense of (3.4), then it also (almost) satisfies the lower bound given by Freĭman's lemma (Lemma 3.3), up to a factor of $1/2$. It is therefore useful to recall Freĭman's lemma: The statement of the lemma is then:

Lemma 3.3 (Freĭman [61])**.** *Let* $A \subseteq \mathbb{R}^d$ *be a finite set of full rank. Then,*

$$|A + A| \geqslant (d+1)|A| - \binom{d+1}{2}.$$

Once we have Theorem 3.2, the proof of Theorem 3.1 is substantially simpler than usual for a similarly strong container theorem. In fact, it is also the only part of the proof that we could omit by using the container theorem for sumsets [26, Theorem 4.2]. The self-contained proof is so simple that we include it here in the overview to motivate the usefulness of Theorem 3.2.

Proof of Theorem 3.1 assuming Theorem 3.2. Our aim is to show that, for every $A \in \binom{P}{k}$ with $|A \hat{+} A| \leqslant m$ and ε-robust Freĭman dimension d, there exists a subset $F \subseteq A$ such that (3.3) holds. We will show that a q-random subset F of A satisfies the first inequality by a suitable choice of q, and the second one via an application of Theorem 3.2 with $Y = F \hat{+} F$.

In order to apply Theorem 3.2 to $Y = F \hat{+} F$, we need to show that Y satisfies (3.4). To this end, we will first prove that a random choice of F makes it unlikely for $F \hat{+} F$ to miss any y such that

$$\rho_{A\hat{+}A}(y) \geqslant \frac{2\varepsilon k^2}{Cm}, \tag{3.5}$$

where

$$\rho_{A\hat{+}A}(y) = \big|\{(a_1, a_2) \in A^2 : a_1 \neq a_2,\, a_1 + a_2 = y\}\big|$$

is the number of pairs of distinct elements of A that sum to a given $y \in A \hat{+} A$. Henceforth, we will refer to the y satisfying (3.5) as the popular elements of $A \hat{+} A$. We will take a q-random subset where either

$$q = \frac{C\sqrt{m \log m}}{2\varepsilon k} \tag{3.6}$$

if the right-hand side is at most 1, and $q = 1$ otherwise (a trivial case which we will ignore).

Notice that we can upper bound the number of missing pairs by:

$$\big|\{(a_1, a_2) \in A^2 : a_1 + a_2 \notin F \hat{+} F\}\big| \leqslant |A| + \sum_{y \in (A\hat{+}A)\setminus(F\hat{+}F)} \rho_{A\hat{+}A}(y) \tag{3.7}$$

where the term $|A|$ comes from the pairs (a, a) for $a \in A$. Once we have proved that $F \hat{+} F$ contains all popular $y \in A \hat{+} A$, we will have an upper bound on $\rho_{A\hat{+}A}(y)$ for every $y \in (A \hat{+} A) \setminus (F \hat{+} F)$. Inserting this into (3.7), we deduce that the number of missing pairs is at most

$$|A| + \sum_{y \in (A\hat{+}A)\setminus(F\hat{+}F)} \rho_{A\hat{+}A}(y) < |A| + \frac{2\varepsilon k^2}{Cm}|A \hat{+} A| \leqslant c\varepsilon |A|^2, \tag{3.8}$$

if we take $C \geqslant 3/c$, where $c = c(\gamma)$ is the constant in Theorem 3.2, since $|A| = k$, $|A \hat{+} A| \leqslant m$

and $\varepsilon k > C$. The upper bound in (3.8) is what we need to apply Theorem 3.2; doing so gives

$$|F \hat{+} F| \geqslant \frac{(1-\gamma)(\dim_{\mathrm{F}}(A)+1)k}{2},$$

from which we can use our assumption that $\dim_{\mathrm{F}}(A) \geqslant d$ to complete the proof.

It therefore only remains to prove our claim that with positive probability F contains all popular elements of $A \hat{+} A$ while also being sufficiently small. Notice that our choice of F as a q-random subset of A implies, for each $y \in A \hat{+} A$, that

$$\mathbb{P}(y \notin F \hat{+} F) = (1-q^2)^{\rho_{A\hat{+}A}(y)/2} \tag{3.9}$$

because (i) each pair $(a_1, a_2) \in A^2$ that satisfies $a_1 + a_2 = y$ and $a_1 \neq a_2$ is counted twice in $\rho_{A\hat{+}A}(y)$ and (ii) the probability that such a pair is chosen in F is q^2.

Take B_F to be the random variable counting the number of $y \in A \hat{+} A$ such that

$$\rho_{A\hat{+}A}(y) \geqslant \frac{2\varepsilon k^2}{Cm} \quad \text{and} \quad y \notin F \hat{+} F.$$

By linearity of expectation and (3.9), we have

$$\mathbb{E}[B_F] = \sum_{\substack{y \in A\hat{+}A \\ \rho_{A\hat{+}A}(y) \geqslant 2\varepsilon k^2/Cm}} \mathbb{P}(y \notin F \hat{+} F) \leqslant m(1-q^2)^{\varepsilon k^2/Cm} \leqslant m \exp\left(-\frac{\varepsilon q^2 k^2}{Cm}\right),$$

where we used $|A \hat{+} A| \leqslant m$ to bound the number of terms in the sum.

Since $\varepsilon q^2 k^2/(Cm) = (C/4\varepsilon)\log m$ by our choice (3.6) of q, it follows by Markov's inequality that

$$\mathbb{P}(B_F > 0) \leqslant m^{-1},$$

where we also used that $\varepsilon < 1$ and $C \geqslant 8$. Using Chernoff's inequality to bound the probability that $|F| > 2qk$, we deduce that

$$\mathbb{P}(|F| > 2qk) + \mathbb{P}(B_F > 0) < 1,$$

which proves that there exists a fingerprint $F \subseteq A$ satisfying (3.3), as required. □

Before moving on with the overview, let us briefly discuss the robustness property in Theorem 3.2. This condition may at first seem unnatural, but the following simple construction shows that some form of robustness is necessary: take A to be the union of $d-1$ random points with a progression P, and define $Y = P + P$. We have simultaneously with high probability that $\dim_{\mathrm{F}}(A) = d$, $|Y| \approx 2|A|$ and the sum of almost all pairs of elements of A are in Y.

With Theorem 3.1 in hand, we can now check that for the family $\mathcal{F}$ of all sets satisfying (3.3), our requirements for the fingerprints F are satisfied. Recall that what we need from the size of the sumset is

$$(ns)^{d+1}(1-p)^{|F\hat{+}F|} \to 0,$$

where $s = |P|$ and $d = \dim(P)$. Modern formulations of Freĭman's theorem tells us that we can

take $s \leqslant \exp(\sigma^{1+o(1)})k$, which in our range of σ and k corresponds to $n^{o(1)}$. However, we can only apply Theorem 3.1 to sets $A \subseteq P$ such that $\dim_{\mathrm{F}}(A) \geqslant \dim(P)$ (ignoring the robustness for the moment). Standard formulations of Freĭman's theorem only guarantee that $\dim(P)$ is at most $\sigma[A]$, which would not be good enough for us.

Fortunately, Chang [35] proved a version of Freĭman's theorem that guarantees that $\dim(P)$ is at most $\dim_{\mathrm{F}}(A)$, at the cost of a weaker bound on the size of P as compared to the more recent results of Sanders [115] and Raghavan [106]. The impact of the suboptimal size of the progression is a slight reduction in the range of p for which our proof works.

Theorem 3.4 (Chang [35], see [74, Proposition 1.3])**.** *There exists $C' > 0$ such that for all finite subsets $A \subseteq \mathbb{Z}$ with $|A| \geqslant 2$ and $\sigma[A] \leqslant \sigma$, there is a d-dimensional generalised arithmetic progression P such that $A \subseteq P$,*

$$|P| \leqslant \exp\bigl(C'\sigma^2(\log\sigma)^3\bigr)|A|$$

and $d \leqslant \dim_{\mathrm{F}}(A)$.

Now, we can use the lower bound on $|F \mathbin{\hat{+}} F|$ given by Theorem 3.1 to obtain

$$(1-p)^{|F\hat{+}F|} \leqslant \exp\bigl(-(1-\gamma)(d+1)kp/2\bigr),$$

which, choosing $k = (2+o(1))\log_{1/(1-p)} n$ (a little larger than $(2\log n)/p$), is at most

$$\exp\bigl(-(1+\gamma)(d+1)\log n\bigr). \tag{3.10}$$

Replacing this back in the previous equation, and recalling that $s = n^{o(1)}$, thus yields

$$(ns)^{d+1}(1-p)^{|F\hat{+}F|} \leqslant n^{-\gamma} \to 0.$$

The attentive reader may have noticed that Chang's theorem is stated for subsets of $\mathbb{Z}$ instead of $\mathbb{Z}_n$. To use it for subsets of $\mathbb{Z}_n$, we will use instead a version of Green–Ruzsa's theorem (Freĭman's theorem for Abelian groups) due to Cwalina and Schoen [46]. It does not bound $\dim(P) \leqslant \dim_{\mathrm{F}}(A)$ directly, though, but it yields a proper progression, which we can combine with a theorem by Tao and Vu [124] to obtain what we need (see Proposition 3.20 and [98, Lemma 4.1]).

Moreover, it is not true that every A has ε-robust Freĭman dimension equal to $\dim_{\mathrm{F}}(A)$. This is not a problem, however, since it is straightforward to prove (see Proposition 3.26) that every set A has a large enough subset A' with ε-robust Freĭman dimension $d' \in \mathbb{N}$.

Finally, let us check that the size of F given to us by Theorem 3.1 is compatible with the range of p in Theorem 1.3. To do so, we need to show that, as $n \to \infty$,

$$\binom{s}{|F|}(1-p)^{|F\hat{+}F|} \to 0.$$

As we already know from (3.10) that the second term is at most n^{-d}, we need

$$\binom{s}{|F|} \leqslant s^{|F|} \leqslant \exp\left(C\varepsilon^{-1}(m\log m)^{1/2}\log s\right) = n^{o(1)}. \tag{3.11}$$

The second inequality in (3.11) follows from (3.3) in Theorem 3.1, while the third is a consequence of our choice of k and the bound on s given by Theorem 3.4. Indeed, we have

$$C\varepsilon^{-1}(m\log m)^{1/2}\log s \ll k^{3/4} = o(\log n), \tag{3.12}$$

because $m \leqslant k^{1+c}$ for $A \in \mathcal{A}_1$, which implies that $\log s \leqslant k^{3c}$, and we can take $\varepsilon = k^{-2c}$ and C to be a constant. Our choice of $c = 1/40$ is therefore more than sufficient to prove (3.12).

At this point, one might ask why we decided to keep track of the constant C up to the final computation. Note that it is crucial that the value of C does not increase too quickly with d growing, since otherwise (3.12) would not hold for large d. Recall that in the proof of Theorem 3.1, we took $C \approx c^{-1}$. The constant c is essentially given to us by our supersaturation result and its value is tied to how many translates we need in our approximate bound for Freĭman's lemma (Theorem 1.4).

To see where the dependency of c on the number of translates comes from, consider the contrapositive of Theorem 3.2: if the set Y is small, then it misses many pairs $(a_1, a_2) \in A^2$. By Theorem 1.4, we can find a small $T \subseteq A$ such that

$$|A + T| - |Y| \geqslant \gamma(d+1)|A|.$$

The pigeonhole principle then ensures us that there is some $x \in T$ satisfying

$$\big|(A + x) \setminus Y\big| \geqslant \frac{\gamma(d+1)|A|}{|T|} = c|A|$$

for $c = \gamma(d+1)/|T|$. Now, if we add the $c|A|$ pairs determined by $(A+x)\setminus Y$ to our collection of missing pairs $(a_1, a_2) \in A^2$ such that $a_1 + a_2 \notin Y$, remove x from A and repeat this procedure t times, we would have at least $c|A|t$ missing pairs in total. Recalling that A has ε-robust Freĭman dimension d, we can take $t = \varepsilon|A|$ while maintaining $\dim_{\mathrm{F}}(A) \geqslant d$, and hence obtain $c\varepsilon|A|^2$ missing pairs, as required.

The above sketch proof of Theorem 3.2 shows that to prove our supersaturation result with c being an absolute constant, we really need the size of T to be a linear function of d, as in Theorem 1.4. In fact, for (3.12), a bound of the form $d^{O(1)}$ would suffice.

Prior to our work, Jing and Mudgal [80] proved the following theorem that is closely related to Theorem 1.4. It obtains the correct leading constant, i.e. without the $(1-5\gamma)/2$ or $1/6$ loss in our bound, at the expense of a worse relationship between the number of translates and the dimension of the set:

Theorem 3.5 ([80, Theorem 1.2])**.** *Given $d \in \mathbb{N}$, there exists a constant $C = C(d) > 0$ such that, for every finite set of full rank $A \subseteq \mathbb{R}^d$, there exists $T \subseteq A$ satisfying $|T| \leqslant C$ and*

$$|A + T| \geqslant (d+1)|A| - 5(d+1)^3.$$

Theorem 3.5 is part of a recent line of work [22, 56] that relies on a beautiful theorem of Bollobás, Leader and Tiba [22, Theorem 8] to obtain sumset lower bounds via few-translates. Unfortunately, the proof of this theorem uses Shao's [120] almost-all Balog–Szemerédi–Gowers theorem, and, as a consequence, the dependency of the number of translates C on the dimension d is of tower-type [22, 120]. Nevertheless, we use their result to prove Theorem 3.2 in the range $d = O(1)$.

A related result, which avoids super-polynomial dependencies between its parameters, is the following elegant theorem proved by Fox, Luo, Pham and Zhou [56]. Its proof relies on a clever path counting argument akin to Gowers' proof of the Balog–Szemerédi theorem. Note that we state a specialized version of their much more general result.

Theorem 3.6 ([56, Theorem 1.1])**.** *There exists $c > 0$ such that the following holds. Let $A \subseteq \mathbb{R}^d$ with $|A| = k$. For every $s \in \{1, \dots, k\}$, there exists $T \subseteq A$ such that $|T| \leqslant s$ and*

$$|A + T| \geqslant c \min\left\{\sigma[A]^{1/3}, s\right\}|A|. \tag{3.13}$$

This result would work for our needs if we could ensure $\sigma[A] > c^{-1}d^3$, but we cannot. However, Fox, Luo, Pham and Zhou also exhibit a construction [56, Proposition 2.4] which shows that we cannot even replace $\sigma[A]^{1/3}$ in (3.13) by $\sigma[A]^{1/1.29}$; this means that our requirement for the rank of the set in Theorem 1.4 is essential.

The only missing part in this overview is a proof of Theorem 1.4 itself. Instead of sketching it, we complete the picture with the (short) proof of its weaker version in the next section.

3.2 A simple proof of a weaker Freĭman's lemma via few translates

This section is dedicated to proving a weaker form of Theorem 1.4. Although it is not strong enough to prove the upper bound in the main theorem of this chapter, it does imply a weaker version where the constant 2 in (3.1) is replaced by a 6. The deduction of this weaker result is the same as that of Theorem 1.3 (see Sections 3.6 and 3.8) simply replacing Theorem 1.4 by Theorem 3.7.

Theorem 3.7. *Let $d, r \in \mathbb{N}$. If $A \subseteq \mathbb{R}^d$ is a finite set with $\operatorname{rank}(A) \geqslant r$, then there exists a set $T \subseteq A$ such that $|T| \leqslant r/2 + 1$ and*

$$|A + T| \geqslant \frac{r|A|}{6}.$$

The idea of the proof is to add elements of A to T one by one, picking in each step a new translate that adds a substantial number of new elements to the sumset. This suggests a greedy argument, picking $x \in A \setminus T$ that increases the size of the sumset the most. However, it is not clear how to show that we can make substantial progress for a sufficient number of steps. Previous arguments, such as the one by Fox, Luo, Pham and Zhou [56, Theorem 1.1], show that if progress stops, then A must have small doubling; as we must handle polynomially large $\sigma[A]$, such strategies do not work in our case. Instead, we adopt a variation of this idea that

incorporates the geometry of the set, allowing us to reach conclusions that do not rely on the doubling of the set and depend only on its rank.

Roughly speaking, in the proof of Proposition 3.8 below, we will show that we can add a new element to T so that it increases the size of $A+T$ by a factor proportional to $|A \setminus \operatorname{span}(T)|$. Notice that, as long as $|T| < \operatorname{rank}(A)$, we can take a non-trivial step.

Now, if we have enough steps that add $|A|/3$ elements to the sumset, then after $r/2$ steps we will have both the bound for the sumset and for $|T|$ in Theorem 3.7. Otherwise, as we will show that every step makes at least $|A \setminus \operatorname{span}(T)|/2$ "progress", it follows that, after the last "good" step, we must have $|A \cap \operatorname{span}(T)| \geqslant |A|/3$. In this scenario, we define

$$A^* = A \cap \operatorname{span}(T) \qquad \text{and} \qquad A' = A \setminus \operatorname{span}(T),$$

and observe that

$$\operatorname{rank}(A') \geqslant \operatorname{rank}(A) - \operatorname{rank}(T) \geqslant r - r/2 = r/2, \tag{3.14}$$

since $|T| \leqslant r/2$. At this point, we now discard our old T and greedily choose elements of a new set of translates $Z \subseteq A'$ each increasing the rank of $A^* \cup Z$ by one. Each new element that we add yields a disjoint, translated copy of A^* in the sumset. We then obtain

$$|A + Z| \geqslant |A^* + Z| = \frac{r|A^*|}{2} \geqslant \frac{r|A|}{6},$$

because, by (3.14), we can greedily add $r/2$ elements to Z, and $|A^*| \geqslant |A|/3$.

We proceed by formalizing the notion that either the greedy argument suffices, or a significant part of A lies in the span of the elements already in T. The version we state below is more general than we need to prove Theorem 3.7 because we will reuse it when proving Theorem 1.4.

Proposition 3.8. *Let $d, r \in \mathbb{N}$, let $\gamma > 0$ and let $A \subseteq \mathbb{R}^d$ be a finite set with $\operatorname{rank}(A) \geqslant r$. If $T \subseteq \mathbb{R}^d$ satisfies $0 \in T$, $\operatorname{rank}(T) < r$ and $|A \cap \operatorname{span}(T)| \leqslant \gamma|A|$, then there exists an element $a_* \in A$ such that*

$$\big|A + \big(T \cup \{a_*\}\big)\big| \geqslant |A + T| + \frac{(1-\gamma)|A|}{2}.$$

The key idea here is to find a co-dimension 1 hyperplane $\mathbf{h}$ which intersects A exactly in $A \cap \operatorname{span}(T)$. We can find such a hyperplane because A is finite and $\operatorname{rank}(T) < \operatorname{rank}(A)$. Note that the two open half-spaces defined by $\mathbf{h}$ induce a partition $A' = A_+ \cup A_-$. Without loss of generality, we assume that $|A_+| \geqslant |A_-|$, and our goal is now to add, for each point in A_+, a new element to the sumset. To achieve this, we let u be a normal vector of $\mathbf{h}$, choose $y_+ \in A_+$ to maximise $\langle y_+, u\rangle$, and observe that $y_+ + A_+$ is disjoint from $A + T$.

Proof. First, choose a vector $u \in \mathbb{R}^d$ to be the normal of the hyperplane $\mathbf{h}$ discussed above. It should satisfy the following properties

(1) $u \neq 0$,

(2) $\langle z, u\rangle = 0$, for all $z \in \operatorname{span}(T)$, and

(3) $\langle z, u\rangle \neq 0$, for all $z \in A \setminus \operatorname{span}(T)$.

There is a u satisfying items (1) and (2) since $0 \in T$ and $\operatorname{rank}(T) < r$. We can furthermore find a u for which item (3) also holds because there are only finitely many elements in $A \setminus \operatorname{span}(T)$, as A itself is finite.

We partition $A = A_+ \cup A_* \cup A_-$ according to the position of its elements relative to the hyperplane defined by $\langle x, u\rangle = 0$, i.e.

$$\begin{aligned} A_+ &= \{x \in A : \langle x, u\rangle > 0\}, \\ A_* &= \{x \in A : \langle x, u\rangle = 0\}, \\ A_- &= \{x \in A : \langle x, u\rangle < 0\}. \end{aligned}$$

Assume without loss of generality that $|A_+| \geqslant |A_-|$ by swapping the sign of u if necessary, and take $y_+ \in A_+$ to be a maximiser of $f(y) = \langle y, u\rangle$.

We claim that $A_+ + y_+$ is disjoint from $A + T$. In fact, if we let $a \in A_+, b \in A$ and $c \in T$, then:

$$\begin{aligned} \langle a + y_+, u\rangle &> \langle y_+, u\rangle, && \text{as } a \in A_+, \\ &\geqslant \langle b, u\rangle, && \text{by the maximality of } y_+, \\ &= \langle b + c, u\rangle, && \text{as we chose } u \text{ with } c \perp u. \end{aligned}$$

Therefore, if $|A_+| \geqslant (1-\gamma)|A|/2$, we can pick $a_* = y_+$ and prove the proposition:

$$|A + (T \cup \{a_*\})| \geqslant |A + T| + |A_+ + a_*| \geqslant |A + T| + \frac{(1-\gamma)|A|}{2}.$$

Otherwise, since we took $|A_+| \geqslant |A_-|$ and $A_+ \cup A_* \cup A_-$ partition A, we have

$$|A_*| = |A| - |A_-| - |A_+| > \gamma|A|. \tag{3.15}$$

By our choice of u and the definition of A_*, we have $A_* = A \cap \operatorname{span}(T)$, which, by (3.15), contradicts our assumption that $|A \cap \operatorname{span}(T)| \leqslant \gamma|A|$. □

The 1-dimensional perspective we took in this proof suggests that instead of adding a single maximiser y_+ to T at each step, we should pick both the maximiser y_+ and the minimiser y_-. Unfortunately, if we add $\{y_+, y_-\}$ to T, then we could increase $\operatorname{rank}(T)$ by two instead of one, causing the greedy argument to run for only half as many steps.

We need the following simple lemma to handle the case when $|A \cap \operatorname{span}(T)|$ is large.

Lemma 3.9. *Let $d, r, s \in \mathbb{N}$, let $\gamma > 0$ and let $A \subseteq \mathbb{R}^d$ be a finite set with $\operatorname{rank}(A) \geqslant r$. If $A_* \subseteq A$ satisfies $\operatorname{rank}(A_*) < s$, then there exists a set $Z \subseteq A \setminus A_*$ such that*

$$|A_* + Z| \geqslant (r - s)|A_*|$$

and $|Z| \leqslant r - s$.

Proof. First, note that for any set $A \subseteq \mathbb{R}^d$ such that $\operatorname{rank}(A) < s$, we have $\operatorname{rank}(\operatorname{span}(A)) \leqslant s$. That is, taking the span of a set that does not contain 0 may increase its rank by 1.

We will construct $Z = \{z_1, \dots, z_{r-s}\} \subseteq A \setminus A_*$ one element at a time, also defining

$$Z_i = \{z_1, \dots, z_i\} \quad \text{and} \quad W_i = \operatorname{span}(A_* \cup Z_i)$$

for $i \in \{0, \dots, r-s\}$. Notice that if $i < r-s$, then $\operatorname{rank}(W_i) \leqslant s + i < r$, by our assumption on A^* and the definition of W_i. Therefore, there exists $z_{i+1} \in A \setminus W_i$, as $\operatorname{rank}(A) \geqslant r$. Since $A^* \subseteq W_i$, this implies that $A^* + z_{i+1}$ and W_i are disjoint, and hence that $A^* + z_{i+1}$ is disjoint from $A^* + Z_i$. This readily implies the lemma because

$$|A_* + Z| = \sum_{i=1}^{r-s} |A_* + z_i| = (r-s)|A_*|,$$

where in the first equality we repeatedly used that $A_* + z_{i+1}$ and $A_* + Z_i$ are disjoint. □

The proof of Theorem 3.7 now follows easily from Proposition 3.8 and Lemma 3.9:

Proof of Theorem 3.7. We may assume that $0 \in A$; notice that this translation does not change $\operatorname{rank}(A)$. First, we construct sets $T_i = \{a_0, \dots, a_i\} \subseteq A$ by choosing $a_0 = 0$ and a_{i+1} as given by Proposition 3.8. In details, let t be the first index for which

$$|A \cap \operatorname{span}(T_t)| > \frac{|A|}{3}. \tag{3.16}$$

Note that while (3.16) does not hold, we have $\operatorname{rank}(T_i) < \operatorname{rank}(A)$ since $T_i \subseteq A$. We may therefore apply Proposition 3.8 to T_i with $\gamma = 1/3$ to define a_i, which then implies that

$$|A + T_i| \geqslant \frac{|T_i|\,|A|}{3} = \frac{(i+1)|A|}{3} \tag{3.17}$$

for all $i \leqslant t$. Hence, if $t \geqslant r/2$, then, by (3.17), we have

$$|A + T_t| \geqslant \frac{(t+1)|A|}{3} \geqslant \frac{r|A|}{6},$$

and taking $T = T_t$ concludes the proof. We may therefore assume that $t < r/2$.

In this case, we want to apply Lemma 3.9 with $A_* = A \cap \operatorname{span}(T_t)$ and $s = r/2$. Note that, as $\operatorname{rank}(T_t) \leqslant t$ and $0 \in T_t$, we have

$$\operatorname{rank}(A_*) \leqslant \operatorname{rank}(\operatorname{span}(T_t)) \leqslant t < r/2,$$

where in the last inequality we used our assumption that $t < r/2$. This application yields a set $Z \subseteq A \setminus A_*$ such that $|Z| \leqslant r - s = r/2$ and

$$\begin{aligned} |A + Z| &\geqslant |A_* + Z| && \text{because } A_* \subseteq A, \\ &\geqslant \frac{r|A_*|}{2} && \text{by Lemma 3.9}, \\ &\geqslant \frac{r|A|}{6} && \text{since } |A_*| \geqslant |A|/3 \text{ by (3.16)}. \end{aligned}$$

Taking $T = Z$ concludes the proof. □

3.3 Improving Freĭman's lemma via few-translates

To obtain the sharp upper bound for $\alpha(G_p)$, the bound $|A+T| \geqslant r|A|/6$ is not enough; we need at least $|A+T| \geqslant (1-5\gamma)(r+1)|A|/2$ for small $\gamma > 0$. In this section, we describe the overall approach and state the intermediate results we require to improve the bound. Once we have stated these results, we will show that assuming them is sufficient to prove Theorem 1.4 – this will motivate their statements, since neither they nor their proofs are straightforward. We will prove that these results hold in subsequent sections.

To discuss the methods that we will use to improve the bound, it will be convenient to first make some definitions.

Definition 3.10. For finite sets $A, W \subseteq \mathbb{R}^d$, we define

$$Z(A,W) = \Pi_{W^\perp}(A),$$

where $\Pi_U(V)$ denotes the orthogonal projection of V onto U. We also partition $\mathbb{R}^d$ into equivalence classes in that projection, denoting these by

$$[z]_W = \big\{x \in \mathbb{R}^d : \Pi_{W^\perp}(x) = z\big\},$$

and partition A into equivalence classes in the same way:

$$[\![z]\!]_{W,A} = [z]_W \cap A.$$

It will be convenient to omit the dependency of those definitions on A and W whenever these sets are clear from context, leaving us with the notation $Z, [z], [\![z]\!]$. We also refer to $[z]$ as a "fibre", and say that a fibre $[z]$ is "empty" if $z \notin Z$. Finally, we remark that those are only used here in Chapter 3, so as not to be confused with $[n] = \{1, \dots, n\}$ used elsewhere in this thesis (note that $[n]$ is always used with integers n, and $[z]$ used with vectors $z \in \mathbb{R}^d$).

The start of the proof follows the same idea as in Section 3.2: to obtain T_{i+1} from T_i, at each step we add the element provided by Proposition 3.8. We do this for t steps, where t is the first index for which $|A \cap \operatorname{span}(T_t)| > \gamma|A|$. If $t \geqslant r$, then, by Proposition 3.8, we have $|A+T_t| \geqslant (1-\gamma)(t+1)|A|/2$, and we are done. Otherwise, we define

$$T^* = T_{t-1} \qquad \text{and} \qquad W = \operatorname{span}(T_t), \tag{3.18}$$

and we look into the set of non-empty fibres $Z = Z(A,W)$. Notice that W is neither $\mathbb{R}^d$ nor $\{0\}$ since $0 < t < r$. The rest of the argument is divided into two different cases.

If there are many distinct non-empty fibres in $Z = \{z_1, \dots, z_m\}$, then we use the following generalisation of Lemma 3.9. Whereas for that previous result we needed one point per dimension, in Proposition 3.11 we take one point $y_i \in [\![z_i]\!]$ per non-empty fibre to be our set T, and show that such a choice yields disjoint translates $y_i + A_*$ for $A_* = A \cap W$.

Proposition 3.11. *Let $d, r \in \mathbb{N}$, let $\gamma > 0$, and let $A \subseteq \mathbb{R}^d$ be a finite set with $\operatorname{rank}(A) \geqslant r$. If $W \subseteq \mathbb{R}^d$ is such that $|A \cap W| \geqslant \gamma|A|$ and $|Z| \geqslant (r+1)/\gamma$, where $Z = Z(A,W)$, then there*

exists $T \subseteq A$ *such that*

$$|A+T| \geqslant (r+1)|A|$$

and $|T| = (r+1)/\gamma$.

Proof. Define $m = (r+1)/\gamma$. Take $\{z_1, \dots, z_m\} \subseteq Z$ and, for each z_i, pick some arbitrary $y_i \in [\![z_i]\!]$. Note that $A \cap W \subseteq [\![0]\!]$. As we have chosen $y_i \in [z_i]$, we have $[\![0]\!] + y_i \subseteq [z_i]$. Doing the same with $i \neq j$ shows that

$$\big([\![0]\!] + y_i\big) \cap \big([\![0]\!] + y_j\big) \subseteq [z_i] \cap [z_j] = \emptyset. \tag{3.19}$$

We can therefore conclude that taking $T = \{y_1, \dots, y_m\} \subseteq A$ satisfies

$$\begin{aligned}
|A+T| &\geqslant \big|[\![0]\!] + T\big| && \text{as } [\![0]\!] \subseteq A \text{ by definition,} \\
&\geqslant \sum_{i=1}^{m} \big|[\![0]\!] + y_i\big| && \text{by (3.19),} \\
&\geqslant m\gamma|A| && \text{as } \big|[\![0]\!]\big| \geqslant \gamma|A|, \text{ by assumption,} \\
&= (r+1)|A|,
\end{aligned}$$

as required. □

The case when there are few non-empty fibres, i.e. $Z = Z(A, W)$ is small, is more complex. Letting $r_W = \operatorname{rank}(W)$ and recalling (3.18), we have

$$|A + T^*| \geqslant (1-\gamma)\frac{(r_W+1)|A|}{2}.$$

Therefore, to obtain a final set of translates T such that $|A+T| \geqslant (1-5\gamma)(r+1)|A|/2$, we need to find a set T' that roughly satisfies

$$|A + T'| \geqslant \frac{(r-r_W)|A|}{2}. \tag{3.20}$$

To combine T^* and T', though, we must take care not to count overlaps between the sumsets $A+T'$ and $A+T^*$ more than once.

The content of the next proposition shows that choosing T' as discussed above is possible. It says that we can choose a T' which almost attains (3.20) while avoiding not only $A+T^*$, but the whole of $A + [\![0]\!]$ – recall that $T^* \subseteq W \cap A \subseteq [\![0]\!]$.

Proposition 3.12. *Let* $d, r, r_W \in \mathbb{N}$, $\eta > 0$, *and let* $A \subseteq \mathbb{R}^d$ *be a finite set with* $\operatorname{rank}(A) \geqslant r$. *Let also* $W \subseteq \mathbb{R}^d$ *be a subspace with dimension* r_W, *and* $Z = Z(A, W)$. *If* $\big|[\![z]\!]\big| \leqslant \eta|A|$ *for every* $z \in Z \setminus \{0\}$, *then there exists* $T' \subseteq A$ *such that*

$$\Big|\big(A+T'\big) \setminus \big(A + [\![0]\!]\big)\Big| \geqslant \frac{r-r_W}{2}\Big(|A| - \big|[\![0]\!]\big|\Big) - \eta|Z||A| \tag{3.21}$$

and $|T'| \leqslant |Z|$.

To gain some intuition for why Proposition 3.12 is true, consider the following. Applying

Freĭman's lemma in the projected world $W^\perp$, if we could obtain a lower bound depending on $|A|$, instead of $|Z|$ – which is what we actually get – then we would be done. Notice, however, that this naive application considers every non-empty fibre $[\![z]\!]$ as a single element to avoid repeated counts. That is, for each $z_1 + z_2 \in Z + Z$, this approach counts only $x + y$ for a single choice of $x \in [\![z_1]\!]$ and $y \in T' \cap [\![z_2]\!]$.

In the proof of Proposition 3.12, we will show that we can instead consider $[\![z_1]\!] + y$ and not overcount. To achieve that, we assign to each $z \in Z$ a weight that encodes the size of the corresponding fibre $[\![z]\!]$, and incorporate this weight into the proof of Freĭman's lemma in the projected world. Although considering $[\![z_1]\!] + y$ is still not enough to replace $|Z|$ with $|A|$ in the bound, we can crucially choose $z_1, z_2 \in Z$ whichever way we want, as long as $y \in [\![z_2]\!]$. We therefore choose the representation that maximises $\big|[\![z_1]\!]\big|$ and show that this modification is enough to obtain the bound in Proposition 3.12.

The above discussion overlooks the removal of the zero fibre from the sumset, i.e. it gives a lower bound for $|A + T'|$ instead of one for $\big|(A + T') \setminus (A + [\![0]\!])\big|$. To incorporate it, we must take into account our choice maximizing $\big|[\![z_1]\!]\big|$, which imposes the (technical) requirement that every non-zero fibre has size at most $\eta|A|$. As this is not always the case, in order to apply Proposition 3.12, we will change W slightly to make it so: we will "add" to W the large non-zero fibres until we have $\big|[\![z]\!]\big| \leqslant \eta|A|$ for all remaining $z \in Z$. As we will do this economically, using this new W will not hurt our bound too much.

Nevertheless, combining T' and T^* had another unanticipated cost: it incurred a negative $(r - r_W)\big|[\![0]\!]\big|/2$ term in the bound of Proposition 3.12. The following proposition shows that a simple change suffices to offset this loss: we add two extra points from each non-empty fibre to our final T.

Proposition 3.13. *Let $d \in \mathbb{N}$, and let $A \subseteq \mathbb{R}^d$ be a finite set. For every $W \subseteq \mathbb{R}^d$ and $u^* \in \mathbb{R}^d$, there is $T'' \subseteq A$ satisfying $|T''| \leqslant 2|Z|$ and*

$$\Big|\big([\![0]\!] + T''\big) \setminus \big(A + [\![0]\!]_*\big)\Big| \geqslant |Z|\Big(\big|[\![0]\!]\big| - \big|[\![0]\!]_*\big|\Big),$$

where $Z = Z(A, W)$ and $[\![0]\!]_* = \{x \in [\![0]\!] : \langle x, u^*\rangle = 0\}$.

It is not immediate that Proposition 3.13 is really enough to make up for what we lost in Proposition 3.12, as we are removing $A + [\![0]\!]_*$ instead of $A + T^*$, and $[\![0]\!]_*$ is determined by W and u^*. What we need is that, for our choice of u^*, both $\big|[\![0]\!]_*\big| \leqslant \gamma|A|$ and $T_{t-1} = T^* \subseteq [\![0]\!]_*$. This condition, combined with the following observation, is enough to prove Theorem 1.4. The proof of the observation is a simple (but slightly tedious) manipulation of set relations.

Observation 3.14. *For every $T', T'' \subseteq A$, we have*

$$\big|\big(A + \hat{T}\big) \setminus \big(A + [\![0]\!]_*\big)\big| \geqslant \big|\big(A + T'\big) \setminus \big(A + [\![0]\!]\big)\big| + \big|\big([\![0]\!] + T''\big) \setminus \big(A + [\![0]\!]_*\big)\big| \tag{3.22}$$

where $\hat{T} = T' \cup T''$.

Proof. Note first that $[\![0]\!] + T''$ is a subset of $A + \hat{T}$: this follows easily from

$$[\![0]\!] \subseteq A \quad \text{and} \quad T'' \subseteq \hat{T}.$$

We therefore have

$$\left|A+\hat{T}\right| \geqslant |S| + \left|[\![0]\!]+T''\right|$$

where $S = \left(A+T'\right) \setminus \left([\![0]\!]+T''\right)$, and moreover,

$$\left|(A+\hat{T}) \setminus (A+[\![0]\!]_*)\right| \geqslant \left|S \setminus (A+[\![0]\!]_*)\right| + \left|([\![0]\!]+T'') \setminus (A+[\![0]\!]_*)\right|. \tag{3.23}$$

As the last term in (3.23) already matches what appears in (3.22), the proof is reduced to showing that

$$\left|S \setminus (A+[\![0]\!]_*)\right| \geqslant \left|\left(A+T'\right) \setminus \left(A+[\![0]\!]\right)\right|,$$

where, recall,

$$S \setminus (A+[\![0]\!]_*) = \left(A+T'\right) \setminus \left(([\![0]\!]+T'') \cup (A+[\![0]\!]_*)\right).$$

It is enough, therefore, to prove that

$$[\![0]\!]+T'' \subseteq [\![0]\!]+A \quad \text{and} \quad A+[\![0]\!]_* \subseteq A+[\![0]\!].$$

Both inclusions are trivial, as $T'' \subseteq A$ by assumption, and $[\![0]\!]_* \subseteq [\![0]\!]$ by definition. □

We are now ready to put the pieces together and prove Theorem 1.4.

Proof of Theorem 1.4 assuming Propositions 3.12 and 3.13. By translating if necessary, we may assume that $0 \in A$; note that this does not change $\operatorname{rank}(A)$. As in the proof of the weaker version, Theorem 3.7, we start by defining $T_0 = \{0\}$: observe that $\operatorname{rank}(T_0) = 0$ and

$$|A+0| = |A|. \tag{3.24}$$

The next steps consist of taking $T_{i+1} = T_i \cup \{a_{i+1}\}$, where $a_{i+1} \in A$ is defined to be the x^* given by Proposition 3.8 applied to T_i. We can do so as long as $|A \cap \operatorname{span}(T_i)| < \gamma|A|$ and $i < r$, because $|T_i| = i+1$ implies that $\operatorname{rank}(T_i) \leqslant i$. We also stop this process for the first t such that

$$|A \cap \operatorname{span}(T_t)| \geqslant \gamma|A|.$$

By (3.24) and our choice of a_{i+1} via Proposition 3.8, we have, for all $i \leqslant t$,

$$|A+T_i| \geqslant |A| + (1-\gamma)\frac{i|A|}{2} \geqslant (1-\gamma)\frac{(i+2)|A|}{2}, \tag{3.25}$$

and $|T_i| \leqslant i+1$.

Now, if $t \geqslant r$, then we can take $T = T_t$, and we have completed the proof by (3.25). Otherwise, we may assume that $t < r$ and define $W_1 = \operatorname{span}(T_t)$. We would like to use W_1 for the rest of the proof, but Proposition 3.12 requires that all non-zero fibres have size at most $\eta|A|$. To continue, then, we need to define a subspace W such that for every $z \in Z(A,W) \setminus \{0\}$ and given $\eta > 0$,

$$\left|[\![z]\!]_{W,A}\right| \leqslant \eta|A|.$$

We will do so by iteratively projecting these large fibres onto the 0 fibre until none remain.

With foresight, we set $\eta = \gamma^2$. Formally, our process is:

1. Start with $\ell = 1$ and $W_1 = \mathrm{span}(T_t)$.
2. If there exists $z_\ell \in Z(W_\ell, A) \setminus \{0\}$ such that $\big|[\![z_\ell]\!]\big| \geqslant \eta|A|$, then we let
$$W_{\ell+1} = \mathrm{span}(W_\ell \cup \{z_\ell\}).$$
3. If there is no such z_ℓ, we stop with output W_ℓ.

A simple and important observation is that $\ell \leqslant 1/\eta$, since $\eta < \gamma$ and

$$|A| \geqslant |W_\ell \cap A| \geqslant \ell\eta|A|.$$

Noting that $\dim(W_1) \leqslant t$, since $|T_t| \leqslant t+1$ and $0 \in T_t$, it follows that

$$r_W := \dim(W_\ell) \leqslant \dim(W_1) + 1/\eta \leqslant t + 1/\eta. \tag{3.26}$$

Take W to be the output of the above process. Moreover, let $Z = Z(A, W)$ and divide the rest of the proof into two cases depending on $|Z|$. The first case is if $|Z| \geqslant (r+1)/\gamma$. Here, we claim that applying Proposition 3.11 completes the proof. The set $T \subseteq A$ provided by this application satisfies

$$|A+T| \geqslant (r+1)|A| \geqslant (1-5\gamma)\frac{(r+1)|A|}{2},$$

and

$$|T| = \frac{r+1}{\gamma} \leqslant \frac{4(r+1)}{\gamma},$$

as required.

It remains to deal with the other case, and we may thus assume that $|Z| < (r+1)/\gamma$. In this scenario, our set T will be the union of three sets: the second to last set T_{t-1}, the set T' given by Proposition 3.12, and T'', the output of Proposition 3.13 for a suitable choice of $u^* \in \mathbb{R}^d$. Note that it consists of few translates:

$$|T| \leqslant |T_{t-1}| + |T'| + |T''| \leqslant t + |Z| + 2|Z| \leqslant \frac{4(r+1)}{\gamma},$$

where the last inequality follows from our assumptions that $t < r$ and $|Z| < (r+1)/\gamma$. It remains to show that $A+T$ has the appropriate size.

First, we separate the contributions of T_{t-1} and $\hat{T} = T' \cup T''$ to the sumset

$$|A+T| \geqslant |A+T_{t-1}| + \big|\big(A+\hat{T}\big) \setminus \big(A+T_{t-1}\big)\big|, \tag{3.27}$$

and we would like to apply Observation 3.14 to bound the second term in the right-hand side. To do this, we need to show that $T_{t-1} \subseteq [\![0]\!]_*$ for some $u^* \in \mathbb{R}^d$, as that would imply

$$\big(A+\hat{T}\big) \setminus \big(A+[\![0]\!]_*\big) \subseteq \big(A+\hat{T}\big) \setminus \big(A+T_{t-1}\big). \tag{3.28}$$

Besides $T_{t-1} \subseteq [\![0]\!]_*$, our choice of u^* will also need to satisfy

$$\big|[\![0]\!]_*\big| < \gamma|A|. \tag{3.29}$$

To achieve that, we define the subspace $W_0 = \operatorname{span}(T_{t-1})$ and claim that we can pick $u^* \in \mathbb{R}^d$ such that

$$[\![0]\!]_* = \{x \in [\![0]\!] : \langle x, u^* \rangle = 0\} = A \cap W_0. \tag{3.30}$$

As we have stopped the greedy process at t, we must have $|A \cap W_0| < \gamma|A|$, so this choice also satisfies (3.29). We can choose $u^* \in W_0^{\perp}$ satisfying (3.30) because A is finite and

$$\dim(W_0) \leqslant t < r \leqslant d,$$

since $|T_{t-1}| \leqslant t$ and $0 \in T_{t-1}$. Notice that this choice of u^* mimics that of u in the proof of Proposition 3.8.

With this choice of u^*, we have $T_{t-1} \subseteq [\![0]\!]_*$, as required, and so combining (3.28) with Observation 3.14 yields, in (3.27),

$$|A + T| \geqslant |A + T_{t-1}| + \big|\big(A + T'\big) \setminus \big(A + [\![0]\!]\big)\big| + \big|\big([\![0]\!] + T''\big) \setminus \big(A + [\![0]\!]_*\big)\big| \tag{3.31}$$

for $T = T_{t-1} \cup T' \cup T''$, and T' and T'' as given by Proposition 3.12 and Proposition 3.13, respectively. As T' originates from Proposition 3.12, we have

$$\big|\big(A + T'\big) \setminus \big(A + [\![0]\!]\big)\big| \geqslant \frac{(r - r_W)}{2}\big(|A| - \big|[\![0]\!]\big|\big) - \eta|Z||A|, \tag{3.32}$$

where, recall, $r_W \leqslant t + 1/\eta$, by (3.26). Moreover, by Proposition 3.13, we have

$$\big|\big([\![0]\!] + T''\big) \setminus \big(A + [\![0]\!]_*\big)\big| \geqslant |Z|\big(\big|[\![0]\!]\big| - \big|[\![0]\!]_*\big|\big),$$

which, by (3.29), implies that

$$\big|\big([\![0]\!] + T''\big) \setminus \big(A + [\![0]\!]_*\big)\big| \geqslant |Z|\big(\big|[\![0]\!]\big| - \gamma|A|\big). \tag{3.33}$$

Recall that, by (3.25), we have

$$|A + T_{t-1}| \geqslant (1 - \gamma)\frac{(t+1)|A|}{2},$$

which, replaced alongside (3.32) and (3.33) in (3.31), yields

$$\begin{aligned} |A + T| \geqslant{}& (1 - \gamma)\frac{(t+1)}{2}|A| \\ &+ \frac{(r - r_W)}{2}\Big(|A| - \big|[\![0]\!]\big|\Big) \qquad (3.34) \\ &- \eta|Z||A| \\ &+ |Z|\Big(\big|[\![0]\!]\big| - \gamma|A|\Big). \qquad (3.35) \end{aligned}$$

The rest of the proof is dedicated to showing that this expression is at least the bound we need.

To this end, observe that

$$|Z| \geqslant \operatorname{rank}(Z) \geqslant \operatorname{rank}(A) - \operatorname{rank}(W) \geqslant r - r_W. \tag{3.36}$$

Another important observation is that

$$\big|[\![0]\!]\big| \geqslant |A \cap W_1| \geqslant \gamma|A|, \tag{3.37}$$

since $A \cap W_1 \subseteq [\![0]\!]$.

It follows from (3.36) and (3.37) that the sum of (3.34) and (3.35) is at least

$$\frac{r - r_W}{2}\Big(|A| - \big|[\![0]\!]\big|\Big) + |Z|\Big(\big|[\![0]\!]\big| - \gamma|A|\Big) \geqslant (1-\gamma)\frac{(r - r_W)}{2}|A|. \tag{3.38}$$

Replacing (3.38) in our lower bound for the size of $A + T$, we obtain

$$|A + T| \geqslant (1-\gamma)\frac{(t+1)}{2}|A| + (1-\gamma)\frac{(r - r_W)}{2}|A| - \eta|Z||A|$$

and using that $r_W \leqslant t + 1/\eta$ by (3.26), yields

$$|A + T| \geqslant (1-\gamma)\frac{(r+1)}{2}|A| - \frac{(1-\gamma)}{2\eta}|A| - \eta|Z||A|. \tag{3.39}$$

It is now enough to determine that

$$(1-\gamma)\frac{1}{2\eta} + \eta|Z| \leqslant 2\gamma(r+1) \tag{3.40}$$

as replacing it in (3.39) gives the desired bound:

$$|A + T| \geqslant (1 - 5\gamma)\frac{r+1}{2}|A|.$$

To obtain (3.40), observe that

$$\eta|Z| \leqslant \gamma(r+1) \tag{3.41}$$

follows from our choice of $\eta = \gamma^2$ and our assumption that, in the current case, $|Z| \leqslant (r+1)/\gamma$. Moreover, the following holds

$$\frac{1}{2\eta} < \gamma(r+1) \tag{3.42}$$

since $\gamma^3 r \geqslant 1$ by assumption. The proof is complete by substituting (3.41) and (3.42) in (3.40). □

3.4 Proposition 3.13: offsetting the loss of the zero fibre

As we remarked in the previous section, we offset the loss of the zero fibre by adding two carefully selected points from each non-empty fibre to the final set of translates. These points are the

maximiser and minimiser of the linear function $x \mapsto \langle x, u^* \rangle$ in each fibre. Let $Z = \{z_1, \dots, z_m\}$. For each $z_i \in Z$, we choose $y_i^+, y_i^- \in [\![z_i]\!]$ to be, respectively, a maximiser and a minimiser of $y \mapsto \langle y, u^* \rangle$ in $[\![z_i]\!]$, and set $Y_i := \{y_i^+, y_i^-\}$.

Recall that $Z = Z(A, W)$ is now defined as the projection of A onto $W^\perp$, where A and W come from the statement of the proposition. Similarly to the proof of Proposition 3.8, we will define "positive" and "negative" parts of $[\![0]\!]$ as

$$[\![0]\!]_+ = \big\{x \in [\![0]\!] : \langle x, u^* \rangle > 0\big\} \quad \text{and} \quad [\![0]\!]_- = \big\{x \in [\![0]\!] : \langle x, u^* \rangle < 0\big\},$$

which complement the "null part" $[\![0]\!]_* = \{x \in [\![0]\!] : \langle x, u^* \rangle = 0\}$ defined in the statement and complete a partition of the zero fibre.

Each pair of minimiser and maximiser we put in the set of translates will add to the sumset a translated copy of the sets $[\![0]\!]_+$ and $[\![0]\!]_-$. As they form a partition of $[\![0]\!] \setminus [\![0]\!]_*$, this will correspond to adding $\big|[\![0]\!]\big| - \big|[\![0]\!]_*\big|$ elements to the sumset for each of the $m = |Z|$ non-empty fibres. Showing this only requires the following simple geometric observations. The first says that the translates $[\![0]\!] + Y_i$ and $[\![0]\!] + Y_j$ are disjoint if Y_i and Y_j lie in distinct fibres.

Observation 3.15. *For all $i, j \in \{1, \dots, m\}$, if $i \neq j$, then*

$$\big([\![0]\!] + Y_i\big) \cap \big([\![0]\!] + Y_j\big) = \emptyset.$$

Proof. Note that $[\![0]\!] + Y_i \subseteq [z_i]$, and recall from (3.19) that $[z_i] \cap [z_j] = \emptyset$. □

The second observation says that the positive part of $[\![0]\!]$ translated by the fibre maximiser y_i^+ is disjoint from its negative part translated by the corresponding fibre minimiser. Its proof is essentially contained in the proof of Proposition 3.8.

Observation 3.16. *For all $i \in \{1, \dots, m\}$, we have*

$$\big(y_i^+ + [\![0]\!]_+\big) \cap \big(y_i^- + [\![0]\!]_-\big) = \emptyset. \tag{3.43}$$

Proof. Note that for any $a \in [\![0]\!]_+, b \in [\![0]\!]_-$, we have

$$\langle y_i^+ + a, u^* \rangle > \langle y_i^+, u^* \rangle \geqslant \langle y_i^-, u^* \rangle > \langle y_i^- + b, u^* \rangle.$$

Hence, $y_i^+ + a \neq y_i^- + b$ for every $a \in [\![0]\!]_+$ and $b \in [\![0]\!]_-$, which implies (3.43). □

The final observation says that the original set translated by the "null part" of $[\![0]\!]$ does not intersect the positive part of $[\![0]\!]$ translated by the fibre maximiser y_i^+, and that the analogous statement holds for the negative part and the fibre minimiser y_i^-.

Observation 3.17. *For all $i \in \{1, \dots, m\}$, we have*

$$\big(A + [\![0]\!]_*\big) \cap \big(y_i^+ + [\![0]\!]_+\big) = \big(A + [\![0]\!]_*\big) \cap \big(y_i^- + [\![0]\!]_-\big) = \emptyset.$$

Proof. Recall that $A = \bigcup_{j=1}^m [\![z_j]\!]$ is a partition. Whenever $[\![z_j]\!]$ and y_i^+ are in distinct fibres, we

have that $[\![z_j]\!] + [\![0]\!]_*$ and $y_i^+ + [\![0]\!]_+$ are disjoint by (3.19),

$$[\![z_j]\!] + [\![0]\!]_* \subseteq [z_j] \quad \text{and} \quad y_i^+ + [\![0]\!]_+ \subseteq [z_i].$$

We now consider the case when they are in the same fibre. Take $a \in [\![0]\!]_+$, $c \in [\![0]\!]_*$ and $y \in [\![z_i]\!]$. Note that $\langle a, u^* \rangle > 0$ since $a \in [\![0]\!]_+$, and $\langle c, u^* \rangle = 0$ as $c \in [\![0]\!]_*$. Then $c + y \neq a + y_i^+$, because

$$\langle a + y_i^+, u^* \rangle > \langle c + y_i^+, u^* \rangle \geqslant \langle c + y, u^* \rangle,$$

and this completes the proof in the y_i^+ case. The proof in the y_i^- case is analogous. □

We are now ready to prove Proposition 3.13.

Proof of Proposition 3.13. Recall that $m = |Z|$, where $Z = Z(A, W)$, and $Y_i = \{y_i^+, y_i^-\}$ consists of a maximiser and minimiser of $y \mapsto \langle y, u^* \rangle$ in $[\![z_i]\!]$ for each $z_i \in Z$. Y_i is well-defined because Z is the set of non-empty fibres, although it may be a singleton for example if $\big|[\![z_i]\!]\big| = 1$. We define T'' to be the following set with (trivially) at most $2m$ elements

$$T'' = \bigcup_{i=1}^{m} Y_i.$$

To avoid cumbersome notation, we will first prove the weaker inequality

$$\big|[\![0]\!] + T''\big| \geqslant \sum_{i=1}^{m} \big|y_i^+ + [\![0]\!]_+\big| + \big|y_i^- + [\![0]\!]_-\big|, \tag{3.44}$$

and then show that the same steps can be applied removing the set $A + [\![0]\!]_*$ in the left-hand side to obtain the desired bound.

As T'' is the union of the Y_i, we use Observation 3.15 to obtain

$$\big|[\![0]\!] + T''\big| = \sum_{i=1}^{m} \big|[\![0]\!] + Y_i\big|. \tag{3.45}$$

Now, because $[\![0]\!]_+$, $[\![0]\!]_*$ and $[\![0]\!]_-$ partition $[\![0]\!]$, we may decompose, for each Y_i, the term in (3.45) as

$$\big|[\![0]\!] + Y_i\big| = \big|\big([\![0]\!]_+ + Y_i\big) \cup \big([\![0]\!]_* + Y_i\big) \cup \big([\![0]\!]_- + Y_i\big)\big|, \tag{3.46}$$

which, ignoring the term corresponding to $[\![0]\!]_*$ and the "mixed-sign" terms, $[\![0]\!]_- + y_+$ and $[\![0]\!]_+ + y_-$, yields

$$\big|[\![0]\!] + Y_i\big| \geqslant \big|\big([\![0]\!]_+ + y_i^+\big) \cup \big([\![0]\!]_- + y_i^-\big)\big|. \tag{3.47}$$

We can now apply Observation 3.16, which implies that the right-hand side of (3.47) is at least

$$\big|\big([\![0]\!]_+ + y_i^+\big) \cup \big([\![0]\!]_- + y_i^-\big)\big| = \big|[\![0]\!]_+ + y_i^+\big| + \big|[\![0]\!]_- + y_i^-\big|, \tag{3.48}$$

hence establishing (3.44) via (3.45).

To obtain the desired bound, we now show that the same steps can be applied removing the set $A + [\![0]\!]_*$ in the left-hand side of (3.44). First, (3.45) holds if we remove $A + [\![0]\!]_*$ from the set in the left-hand side and those in the sum in the right-hand side because removing a set cannot create intersections in disjoint sets. We then repeat the steps in (3.46) and (3.47) – they are also possible regardless of the set removal – and, upon reaching (3.48), we again use that disjointness is preserved under set removal. This gives us

$$\Big|\big([\![0]\!] + T''\big) \setminus \big(A + [\![0]\!]_*\big)\Big| \geqslant \sum_{i=1}^{m} \Big|\big(y_i^+ + [\![0]\!]_+\big) \setminus \big(A + [\![0]\!]_*\big)\Big| + \Big|\big(y_i^- + [\![0]\!]_-\big) \setminus \big(A + [\![0]\!]_*\big)\Big|.$$

We are now in a position to use Observation 3.17 to deduce that the set we removed is disjoint from the ones in the right-hand side above, and recover the lower bound prior to the removal

$$\Big|\big([\![0]\!] + T''\big) \setminus \big(A + [\![0]\!]_*\big)\Big| \geqslant \sum_{i=1}^{m} \big|y_i^+ + [\![0]\!]_+\big| + \big|y_i^- + [\![0]\!]_-\big| = m\Big(\big|[\![0]\!]_+\big| + \big|[\![0]\!]_-\big|\Big).$$

Again using that $[\![0]\!]_- \cup [\![0]\!]_* \cup [\![0]\!]_+$ is a partition of $[\![0]\!]$, we get

$$m\Big(\big|[\![0]\!]_+\big| + \big|[\![0]\!]_-\big|\Big) = m\Big(\big|[\![0]\!]\big| - \big|[\![0]\!]_*\big|\Big),$$

as required to complete the proof. □

3.5 Weighted Freĭman's lemma: proof of Proposition 3.12

The final piece we need to prove Theorem 1.4 is detailing how to choose $T' \subseteq A$ such that $|T'| \leqslant |Z|$ and (3.21) holds, and how to prove a variant of Freĭman's lemma where the size of the original, unprojected set A appears in the lower bound, instead of simply $|Z|$.

We will first prove a weaker, insufficient statement to make the reader comfortable with the notation and ideas in the proof of Proposition 3.12. Our goal here is to show that we can choose $T' \subseteq A$ such that $|T'| \leqslant |Z|$ and[2]

$$|A + T'| \geqslant \sum_{w \in Z+Z} \max_{z \in (w-Z)\cap Z} \big|[\![z]\!]\big|. \tag{3.49}$$

Before proceeding, it will be useful to let $Z = \{z_1, \dots, z_m\}$. Here (and in the proof itself), we will take

$$T' = \{y_1, \dots, y_m\},$$

where, for each $i \in \{1, \dots, m\}$, $y_i \in [\![z_i]\!]$ is arbitrary. It is immediate that the size of T' is appropriate, that is,

$$|T'| \leqslant |Z|,$$

so we must show that it also satisfies (3.21).

[2]Notice that $(w - Z) \cap Z = \{z \in Z : \exists z' \in Z \text{ such that } z + z' = w\}$.

Having defined T', we start this warm-up by partitioning $A+T'$ into fibres of $Z+Z$:

$$A+T' = \bigcup_{w\in Z+Z} [w]\cap(A+T'). \tag{3.50}$$

As (3.50) defines a partition, we also have

$$\left|\bigcup_{w\in Z+Z}\Big([w]\cap(A+T')\Big)\right| = \sum_{w\in Z+Z}\left|\Big([w]\cap(A+T')\Big)\right|.$$

Note also that if $w=z_i+z_j\in Z+Z$, then

$$[\![z_i]\!]+y_j \subseteq \big([w]\cap(A+T')\big) \tag{3.51}$$

since $[\![z_i]\!]\subseteq A$ and $y_j\in[\![z_j]\!]\cap T'$.

Our approach will be to count only the elements in a single (translated) fibre $[\![z_i]\!]+y_j$ and ignore the rest of $[z_i+z_j]\cap(A+T')$ – by (3.50) and (3.51), this is a valid lower bound for $|A+T'|$. As the union in (3.50) ranges over all $w\in Z+Z$, we can pick any possible representation z_i+z_j for w. Our choice will be the one for which $\big|[\![z_i]\!]\big|$ is as large as possible, resulting in

$$\sum_{w\in Z+Z}\left|\Big([w]\cap(A+T')\Big)\right| \geqslant \sum_{w\in Z+Z}\max_{z\in(w-Z)\cap Z}\big|[\![z]\!]\big|,$$

since $\big|[\![z]\!]+y\big| = \big|[\![z]\!]\big|$. This completes the proof of (3.49).

In the proof of Proposition 3.12, the statement that is analogous to (3.49) is

$$\Big|(A+T')\setminus(A+[\![0]\!])\Big| \geqslant \sum_{w\in(Z^*+Z^*)\setminus Z}\max_{z\in(w-Z^*)\cap Z^*}\big|[\![z]\!]\big|, \tag{3.52}$$

where $Z^*=Z\setminus\{0\}$ – in words, we "remove the 0 from Z" before doing the sumset. The proof of (3.52) is essentially the same as the warm-up above, but that is still not enough. We must have a good lower bound for its right-hand side in order to complete a proof of Proposition 3.12.

Lemma 3.18, below, is the lower bound we need for the right-hand side of (3.52): when applying it to prove Proposition 3.12, we will set $U=Z^*$ and $f(z)=\big|[\![z]\!]\big|$. The proof of the lemma is similar to traditional proofs of Freĭman's lemma, but we must take into account the weight $f(u)$ of each $u\in U$ when selecting vertices of $\mathrm{conv}(U)$, the convex hull of U, instead of choosing an arbitrary vertex[3].

Lemma 3.18. *Let $d\in\mathbb{N}$. For every finite $U\subseteq\mathbb{R}^d$ and $f:U\to\mathbb{R}_+$, we have*

$$\sum_{w\in U+U}\max_{u\in(w-U)\cap U} f(u) \geqslant \frac{\mathrm{rank}(U)+1}{2}\sum_{u\in U} f(u). \tag{3.53}$$

We will also need a simple observation to repeat one of the steps in the proof of Freĭman's lemma and prove Lemma 3.18.

[3]Hence the name "weighted Freĭman's lemma".

Observation 3.19. *Let $U \subseteq \mathbb{R}^d$ be a finite set. If $v \in U$ is a vertex of $\operatorname{conv}(U)$ and $U' = U \setminus \{v\}$, then there exists a hyperplane $\mathbf{h}$ such that v and $U' \setminus \mathbf{h}$ are on different sides of $\mathbf{h}$. Moreover, $|\mathbf{h} \cap U'| \geqslant \operatorname{rank}(U')$.*

Proof. Write $\operatorname{conv}(U') = \bigcap_{\mathbf{h} \in \mathbf{H}} \mathbf{h}^-$, where $\mathbf{H}$ is the collection of hyperplanes supporting the facets of $\operatorname{conv}(U')$ and $\mathbf{h}^-$ denotes a closed half-space defined by $\mathbf{h}$. Since $v \notin \operatorname{conv}(U')$, there exists $\mathbf{h} \in \mathbf{H}$ such that $v \notin \mathbf{h}^-$, so v and $U' \setminus \mathbf{h}$ are on different sides of $\mathbf{h}$. The second part of the statement follows from $\mathbf{h}$ intersecting a facet of $\operatorname{conv}(U')$ and the fact that every facet contains at least $\operatorname{rank}(U')$ vertices. □

Now, we can prove Lemma 3.18.

Proof of Lemma 3.18. We prove the lemma by induction on $|U|$. In the base case, we have an empty set, so the left-hand side of (3.53) is equal to 0, as is the right-hand side. We can therefore assume that for every $U' \subsetneq U$, we have (3.53).

For the induction step, we proceed similarly to the standard proof of Freĭman's lemma. The main difference is that instead of choosing an arbitrary element from V, the vertices of $\operatorname{conv}(U)$, we must choose a vertex considering the weight function f. To be precise, we choose a vertex $v \in V$ such that

$$f(v) \leqslant \frac{1}{s+1} \sum_{u \in U} f(u), \tag{3.54}$$

where $s = \operatorname{rank}(U)$, noting that such a choice exists by the pigeonhole principle,

$$(s+1) \min_{v \in V} f(v) \leqslant |V| \min_{v \in V} f(v) \leqslant \sum_{v \in V} f(v) \leqslant \sum_{u \in U} f(u).$$

We fix one such $v \in V$, define $U' = U \setminus \{v\}$ and divide the remainder of the proof based on whether $\operatorname{rank}(U') = s$.

Consider first the case $\operatorname{rank}(U') = s$. The induction hypothesis implies that

$$\sum_{w \in U'+U'} \max_{u \in (w-U) \cap U} f(u) \geqslant \sum_{w \in U'+U'} \max_{u \in (w-U') \cap U'} f(u) \geqslant \frac{s+1}{2} \sum_{u' \in U'} f(u') \tag{3.55}$$

since $U' \subseteq U$. Therefore, by (3.55), we can accomplish our goal by finding a set

$$S \subseteq (U+U) \setminus (U'+U')$$

such that

$$\sum_{w \in S} \max_{u \in (w-U) \cap U} f(u) \geqslant \frac{s+1}{2} f(v). \tag{3.56}$$

In order to define our candidate set for S, let $\mathbf{h}$ be the hyperplane given by Observation 3.19 for U and v, and take $\overline{\mathrm{N}}(v) = \big(\mathbf{h} \cap \operatorname{conv}(U')\big) \cup \{v\}$. We claim that $\overline{\mathrm{N}}(v) + v$ is a suitable choice for S because it is disjoint from $U' + U'$ and satisfies (3.56).

To see that $\overline{\mathrm{N}}(v) + v$ and $U' + U'$ are disjoint, first notice that $2v \notin U' + U'$ follows from v being a vertex of $\operatorname{conv}(U)$. For the remaining elements of $\overline{\mathrm{N}}(v)$, i.e. $v' \in \mathbf{h} \cap \operatorname{conv}(U')$, $v' + v$ is

not in $U' + U'$ because $(v' + v)/2$ is a midpoint of the segment connecting v and v', and this midpoint clearly lies outside $\mathrm{conv}(U')$ by our choice of $\mathbf{h}$.

It remains to show that (3.56) holds with $S = \overline{\mathrm{N}}(v) + v$. Observe that if $w = v' + v$ for some $v' \in \overline{\mathrm{N}}(v)$, then $v \in (w - U) \cap U$. Hence,

$$\sum_{w \in \overline{\mathrm{N}}(v)+v} \max_{u \in (w-U)\cap U} f(u) \geqslant \sum_{w \in \overline{\mathrm{N}}(v)+v} f(v) = \left|\overline{\mathrm{N}}(v)\right| f(v) \geqslant (s+1) f(v), \tag{3.57}$$

where in the last inequality we used that

$$|\overline{\mathrm{N}}(v)| = |\mathbf{h} \cap U'| + 1 \geqslant \mathrm{rank}(U') + 1 = s + 1$$

which is due to our choice of $\mathbf{h}$ by Observation 3.19 and our assumption that $\mathrm{rank}(U') = s$. Combining (3.55) and (3.57), this completes the induction step when $\mathrm{rank}(U') = \mathrm{rank}(U)$.

We may therefore assume that $\mathrm{rank}(U') = s-1$; that is, the removal of the vertex v decreases the rank of U. It will be useful for this case to recall (3.54), our criterion for the choice of v, in the following (trivially) equivalent form:

$$\sum_{u \in U} f(u) \geqslant (s+1) f(v). \tag{3.58}$$

Applying our induction hypothesis to U' yields

$$\sum_{w \in U'+U'} \max_{u \in (w-U')\cap U'} f(u) \geqslant \frac{s}{2} \sum_{u' \in U'} f(u'), \tag{3.59}$$

so, to deduce the bound (3.53), it is enough to add to (3.59) the term

$$\Big(\frac{1}{2} \sum_{u' \in U'} f(u')\Big) + \frac{s+1}{2} f(v) \tag{3.60}$$

using elements of $(U + U) \setminus (U' + U')$.

In order to add (3.60) to (3.59), we argue that $U' + U'$, $\{2v\}$ and $U' + v$ are disjoint. This follows from our assumption that $\mathrm{rank}(U') < \mathrm{rank}(U)$, which implies that $v \notin \mathcal{U}$, where $\mathcal{U}$ is an affine subspace with dimension $\mathrm{rank}(U')$ containing U'. We can therefore conclude that

$$\sum_{w \in U+U} \max_{u \in (w-U)\cap U} f(u) \geqslant \frac{s}{2} \sum_{u' \in U'} f(u') + f(v) + \sum_{w \in U'+v} \max_{u \in (w-U)\cap U} f(u). \tag{3.61}$$

Our first observation towards bounding the right-hand side of (3.61) is that

$$\sum_{w \in U'+v} \max_{u \in (w-U)\cap U} f(u) \geqslant \sum_{u' \in U'} f(u')$$

because if $w = u' + v$ for some $u' \in U'$, then $u' \in (w - U) \cap U$. It will therefore suffice to show that

$$f(v) + \sum_{u' \in U'} f(u') \geqslant \Big(\frac{1}{2} \sum_{u' \in U'} f(u')\Big) + \frac{s+1}{2} f(v),$$

or, equivalently,

$$2f(v) + \sum_{u' \in U'} f(u') \geqslant (s+1)f(v).$$

We now use that our choice of v satisfies (3.58), which implies that

$$2f(v) + \sum_{u' \in U'} f(u') \geqslant \sum_{u \in U} f(u) \geqslant (s+1)f(v),$$

and hence completes the proof of the induction step. □

With Lemma 3.18, we are now ready to prove Proposition 3.12.

Proof of Proposition 3.12. Our first claim is that finding a $T' \subseteq A$ such that $|T'| \leqslant |Z|$ and

$$\Big|(A+T') \setminus (A+[\![0]\!])\Big| \geqslant \sum_{w \in (Z^*+Z^*)\setminus Z} \max_{z \in (w-Z^*)\cap Z^*} \big|[\![z]\!]\big|, \tag{3.62}$$

where $Z^* = Z \setminus \{0\}$, is enough to complete the proof. To see that, apply Lemma 3.18 with $U = Z^*$ and $f(z) = \big|[\![z]\!]\big|$ to get the lower bound

$$\sum_{w \in Z^*+Z^*} \max_{z \in (w-Z^*)\cap Z^*} \big|[\![z]\!]\big| \geqslant \frac{\operatorname{rank}(Z^*)+1}{2} \sum_{z \in Z^*} \big|[\![z]\!]\big| \geqslant \frac{r-r_W}{2}\big(|A| - \big|[\![0]\!]\big|\big), \tag{3.63}$$

where in the last equality we have used that

$$Z^* = Z \setminus \{0\}, \quad A = \bigcup_{z \in Z} [\![z]\!] \quad \text{and} \quad \operatorname{rank}(Z^*) \geqslant r - r_W - 1,$$

by our assumptions that $Z = \Pi_{W^\perp}(A)$, $\operatorname{rank}(A) \geqslant r$ and $r_W = \operatorname{rank}(W)$.

However, the left-hand side of (3.63) is considering elements $w \in Z$ that are not in the sum in (3.62). This is not an issue because

$$\sum_{w \in Z} \max_{z \in (w-Z^*)\cap Z^*} \big|[\![z]\!]\big| \leqslant \eta |Z||A| \tag{3.64}$$

follows from our assumption that $\big|[\![z]\!]\big| \leqslant \eta|A|$ for all $z \in Z^*$. We obtain the claim that T' as specified finishes the proof by substituting (3.64) and (3.63) into (3.62):

$$\sum_{w \in (Z+Z)\setminus Z} \max_{z \in (w-Z^*)\cap Z^*} \big|[\![z]\!]\big| \geqslant \frac{r-r_W}{2}\big(|A| - \big|[\![0]\!]\big|\big) - \eta|Z||A|.$$

As it now suffices to find $T' \subseteq A$ such that $|T'| \leqslant |Z|$ and (3.62) holds, we simply need to repeat the proof of (3.49) given in the warm-up, but with the set $A + [\![0]\!]$ removed. That is, let $Z^* = \{z_1, \dots, z_m\}$, and define

$$T' = \{y_1, \dots, y_m\}$$

where each y_i is an arbitrary element of $[\![z_i]\!]$ for $i \in \{1, \dots, m\}$. The first step to show that T'

satisfies (3.62) is partitioning $A+T'$ into fibres of $Z+Z$ to obtain

$$\Big|\big(A+T'\big)\setminus\big(A+[\![0]\!]\big)\Big| = \sum_{w\in Z+Z}\Big|\big([w]\cap(A+T')\big)\setminus\big(A+[\![0]\!]\big)\Big|.$$

In order to handle the set removal, we claim that if $w\notin Z$, then the sets $[w]$ and $A+[\![0]\!]$ are disjoint, and therefore

$$\sum_{w\in Z+Z}\Big|\big([w]\cap(A+T')\big)\setminus\big(A+[\![0]\!]\big)\Big| \geqslant \sum_{w\in(Z+Z)\setminus Z}\big|[w]\cap(A+T')\big|. \tag{3.65}$$

To see this, simply note that if $a\in A$ and $a'\in[\![0]\!]$, then

$$\Pi_{W^\perp}(a+a') = \Pi_{W^\perp}(a)+\Pi_{W^\perp}(a')\in Z+0=Z$$

by the definitions of $[\![0]\!]$ and Z, and therefore $\Pi_{W^\perp}(A+[\![0]\!])\subseteq Z$.

Having established (3.65), we can proceed (almost) like in the warm-up. Restricting our attention to $Z^*+Z^*\subseteq Z+Z$, observe that every $w\in Z^*+Z^*$ can be written as

$$w=z+z' \quad\text{where}\quad z,z'\in(w-Z^*)\cap Z^*. \tag{3.66}$$

For each $w\in Z^*+Z^*$, then, we pick the pair $(z,z')\in(Z^*)^2$ satisfying (3.66) that maximises $\big|[\![z]\!]\big|$. Let y be the (unique) element of $T'\cap[z']$, and note that, as $[\![z]\!]+y\subseteq[w]\cap(A+T')$, we can conclude that

$$\big|[w]\cap(A+T')\big| \geqslant \max_{z\in(w-Z^*)\cap Z^*}\big|[\![z]\!]\big|. \tag{3.67}$$

Replacing (3.67) into (3.65) yields that (3.62) holds for T'

$$\sum_{w\in(Z+Z)\setminus Z}\big|[w]\cap(A+T')\big| \geqslant \sum_{w\in(Z^*+Z^*)\setminus Z}\max_{z\in(w-Z^*)\cap Z^*}\big|[\![z]\!]\big|,$$

and the proof therefore follows from our first claim. □

3.6 The supersaturation result

This section is dedicated to the proof of Theorem 3.2, restated for convenience:

Theorem 3.2. *For every $0<\gamma<1$, there exists a constant $c=c(\gamma)>0$ such that, for every sufficiently large set $A\subseteq\mathbb{Z}_n$, every $d\in\mathbb{N}$ and every $0<\varepsilon<\gamma$, the following holds. If A has ε-robust Freĭman dimension d and $Y\subseteq A+A$ satisfies*

$$\big|\{(a_1,a_2)\in A^2 : a_1+a_2\notin Y\}\big| \leqslant c\varepsilon|A|^2,$$

then $|Y|\geqslant(1-\gamma)(d+1)|A|/2$.

Now that we have Theorem 1.4, proving Theorem 3.2 is simple. Assume that Y is small, that is, $|Y|\leqslant(1-\gamma)(d+1)|A|/2$; our goal is to show that it misses many pairs of A^2. If we can find

$\varepsilon|A|$ elements $a^{(i)} \in A$ such that Y contains at most a $1-c$ proportion of $A+a^{(i)}$, we are done. To do that, we will use Theorem 1.4 to find a small set $T \subseteq A$ such that $|A+T|-|Y| \gtrsim \gamma(d+1)|A|$. By the pigeonhole principle, it follows that there exists an $a^{(i)} \in T$ such that Y misses many elements of $A+a^{(i)}$. We can then remove $a^{(i)}$ from A and repeat the process until the dimension of A drops below its original value (removing the $a^{(i)}$ is a simple way to avoid using the same element twice while keeping our working set large). As we assumed that A has ε-robust Freĭman dimension d, this can only happen after we have removed $\varepsilon|A|$ translates.

This is essentially the proof except for the fact that Theorem 1.4 requires $A \subseteq \mathbb{R}^d$, and the set in Theorem 3.2 is a subset of $\mathbb{Z}_n$. To handle this, we use Freĭman isomorphisms, relying on the fact that if ϕ is a Freĭman isomorphism, then $|A_1 + A_2| = |\phi(A_1) + \phi(A_2)|$.

Proof of Theorem 3.2. Assume that[4] $|Y| \leqslant (1-\gamma')(d+1)|A|/2$. We claim that there exists a sequence of distinct elements $a^{(1)}, \dots, a^{(t)} \in A$, where $t = \varepsilon|A|/4$, such that for each $i \in \{0, \dots, t-1\}$, the set $A_i = A \setminus \{a^{(1)}, \dots, a^{(i)}\}$ satisfies the following two properties:

$$d_i = \dim_{\mathrm{F}}(A_i) \geqslant d \quad \text{and} \quad \left|(A_i + a^{(i+1)}) \setminus Y\right| \geqslant 4c|A|. \tag{3.68}$$

The first property holds because, for all $i \leqslant t$,

$$|A_i| \geqslant |A| - i \geqslant |A| - t = \left(1 - \frac{\varepsilon}{4}\right)|A|, \tag{3.69}$$

so we still have $\dim_{\mathrm{F}}(A_i) \geqslant d$, since A has ε-robust Freĭman dimension d.

To prove that A_i satisfies the second property in (3.68), we will show how to select each $a^{(i+1)}$. That is, assume that we have distinct translates $\{a^{(1)}, \dots, a^{(i)}\}$ such that the set A_i satisfies (3.68). Since $\dim_{\mathrm{F}}(A_i) = d_i \geqslant d$, there exists a Freĭman isomorphism $\phi_i : A_i \to A_i'$ such that $A_i' \subseteq \mathbb{R}^{d_i}$ has full rank. If $\gamma' > 2^6 d^{-1/3}$, then we can apply Theorem 1.4 to the set A_i' with $r = d$ and $\gamma = 2^{-6}\gamma'$, to obtain a set $T_i' \subseteq A_i'$ such that

$$|T_i'| \leqslant \frac{2^6(d+1)}{\gamma'} \quad \text{and} \quad |A_i' + T_i'| \geqslant \left(1 - \frac{\gamma'}{4}\right)\frac{(d+1)|A_i'|}{2}.$$

Otherwise, we have $\gamma' \leqslant 2^6 d^{-1/3}$, i.e. $d \leqslant 2^{18}\gamma'^{-3}$. In this case, we can[5] apply Theorem 3.5 and obtain $T_i' = \{a_1, \dots, a_C\} \subseteq A_i'$ such that

$$|A_i' + T_i'| \geqslant (d+1)|A_i'| - 5(d+1)^3 \geqslant \frac{(d+1)|A_i'|}{2},$$

for some constant $C = C(\gamma')$, since A_i satisfies (3.69), A is sufficiently large and $d \leqslant 2^{18}\gamma'^{-3}$. Therefore, it follows that we have a set of translates T_i' of size at most $2^{-6}\gamma'(d+1)/c$ for some constant $c = c(\gamma') > 0$, in either case.

[4] We use γ' instead of γ because its value is (slightly) less than the γ in the application of Theorem 1.4.

[5] As it is stated, Theorem 3.5 can only be used for $d = d_i$, but that could result in too many translates. To circumvent this issue, we can randomly project the set to $\mathbb{R}^d$, and apply Theorem 3.5 to the projected set instead, using the randomness to avoid collisions.

As ϕ_i is a Freĭman isomorphism, we know that the preimage $T_i = \phi_i^{-1}(T_i')$ satisfies

$$|A_i + T_i| \geqslant \left(1 - \frac{\gamma'}{4}\right) \frac{(d+1)|A_i|}{2} = \left(1 - \frac{\gamma'}{4}\right) \frac{(d+1)(|A| - i)}{2}$$

and $|T_i| \leqslant 2^{-6}\gamma'(d+1)/c$. Since $i \leqslant t = \varepsilon|A|/4 < \gamma'|A|/4$, this is at least

$$|A_i + T_i| \geqslant \left(1 - \frac{\gamma'}{4}\right)^2 \frac{(d+1)|A|}{2} \geqslant \left(1 - \frac{\gamma'}{2}\right) \frac{(d+1)|A|}{2}.$$

Recalling our assumption $|Y| \leqslant (1-\gamma')(d+1)|A|/2$, it follows that

$$|(A_i + T_i) \setminus Y| \geqslant |A_i + T_i| - |Y| \geqslant \frac{\gamma'(d+1)|A|}{4}.$$

Thus, by the pigeonhole principle, there exists $a^{(i+1)} \in T_i$ such that

$$|(A_i + a^{(i+1)}) \setminus Y| \geqslant \frac{\gamma'(d+1)|A|}{4} \left(\frac{\gamma'(d+1)}{2^6 c}\right)^{-1} > 4c|A|,$$

since $|T_i| \leqslant 2^{-6}\gamma'(d+1)/c$ and $T_i \subseteq A_i$. Repeating this selection for each $i \in \{0, \dots, t-1\}$ yields the sets A_i and $\{a^{(1)}, \dots, a^{(t)}\}$ satisfying (3.68).

Now that we have the sets A_i, note that each $A + a^{(i)}$ contributes $4c|A|$ ordered pairs whose sum are not in Y. This gives us a total of

$$4c|A|t \geqslant c\varepsilon|A|^2$$

such pairs, because $t = \varepsilon|A|/4$. □

3.7 Chang's theorem for $\mathbb{Z}_n$

This section is devoted to the proof of Proposition 3.20, the variant of Chang's theorem that we will use in the proof of Theorem 1.3.

Proposition 3.20. *There exists $C > 0$ such that the following holds. Let $n \in \mathbb{N}$ be a prime, and let $\kappa \geqslant 2$. If $A \subseteq \mathbb{Z}_n$ satisfies*

$$\sigma[A] \leqslant \kappa \qquad \text{and} \qquad C\kappa^3(\log \kappa)^2 < |A| < \exp(-C\kappa^4(\log \kappa)^2)n, \tag{3.70}$$

then there is a generalised arithmetic progression $P \subseteq \mathbb{Z}_n$ such that

$$A \subseteq P, \qquad |P| \leqslant \exp(C\kappa^4(\log \kappa)^2)|A| \qquad \text{and} \qquad \dim(P) \leqslant \dim_{\mathrm{F}}(A). \tag{3.71}$$

We rely on two theorems to prove Proposition 3.20: Cwalina and Schoen's strengthening of the Green–Ruzsa theorem (Freĭman's theorem for general Abelian groups) [46] and the discrete John's theorem of Tao and Vu [124]. To state these two results, we need some auxiliary definitions. As in [46], define a coset progression to be a set of the form $P + H$, where H is a

subgroup of G and P is a generalised arithmetic progression. Moreover, the dimension of $P+H$ is $\dim(P)$, we say that a coset progression is proper if (a) P is proper, and (b) $|P+H| = |P|\,|H|$. We say that $P+H$ is 2-proper if $(P+P)+H$ is proper.

Theorem 3.21. *There exists $C' > 0$ such that the following holds. Let G be an Abelian group, and let $A \subseteq G$ be a finite set with $\sigma[A] \leqslant \kappa$. Either there exists a proper coset progression $P+H$ such that*

$$A \subseteq P+H, \quad \dim(P+H) \leqslant 2\kappa+1 \quad \text{and} \quad |P+H| \leqslant \exp(C'\kappa^4(\log(\kappa+2))^2)|A|,$$

or A is fully contained in at most $C'\kappa^3(\log\kappa)^2$ cosets, whose total cardinality is bounded by $\exp(C'\kappa^4(\log(\kappa+2))^2)|A|$, of some subgroup of G.

As our group of interest is $\mathbb{Z}_n$, a proper coset progression $P+H$ is either all of $\mathbb{Z}_n$, or simply a proper generalised arithmetic progression $P+0=P$. We also need a lemma by the same authors to obtain, from a d-dimensional coset progression $P+H$, a 2-proper coset progression that contains $P+H$, has dimension at most d, and whose size is not much larger than $|P+H|$.

Lemma 3.22. *There exists $C' > 0$ such that the following holds. Suppose that $P+H$ is a d-dimensional coset progression in an Abelian group G. Then there exists a 2-proper coset progression $P'+H'$ such that*

$$P+H \subseteq P'+H', \qquad \dim(P'+H') \leqslant d, \qquad \text{and} \qquad |P'+H'| \leqslant d^{C'd^2}|P+H|.$$

These two results enable us to give an overview of the (somewhat standard) proof of Proposition 3.20. First, we apply Theorem 3.21 to A. Using (3.70) and Lemma 3.22, we will show that this application yields a small 2-proper generalised arithmetic progression $P \subseteq \mathbb{Z}_n$ such that $A \subseteq P$. However, the bound on $d = \dim(P)$ is not strong enough for our purposes, so we use it instead to define a Freĭman isomorphism $\phi : P \to P_d \subseteq \mathbb{Z}^d$ in the natural way.

Restricting P_d to an $r = \dim_F(A)$ dimensional subspace, then, yields an r-dimensional *convex* progression, which is a set of the form $\mathcal{C} = K \cap \mathbb{Z}^d$. In this definition, $K \subseteq \mathbb{R}^d$ is a convex set of rank $r \leqslant d$. At this point, we use Tao and Vu's discrete John's theorem [124] (Theorem 3.23 below) to obtain a small generalised arithmetic progression $P' \subseteq \mathbb{Z}^d$ such that $\mathcal{C} \subseteq P'$ and $\dim(P') \leqslant r$. We complete the proof by taking the progression $\phi^{-1}(P') \subseteq \mathbb{Z}_n$.

The above sketch is not quite right for two reasons. The first is that ϕ^{-1} may not be defined for every element in P', but the linearity of the mapping allows us to extend it to all of $\mathbb{Z}^d$. The second reason is that the statement of Theorem 3.23 requires the convex progression to be (centrally) symmetric, i.e. $\mathcal{C} = -\mathcal{C}$. We amend this by applying it instead to $\mathcal{C}-\mathcal{C}$, bounding its size by $|P-P| \leqslant 2^d|P|$, and its dimension by that of the same d'-dimensional subspace that contains $\mathcal{C}$.

Theorem 3.23. *There exists $C' > 0$ such that the following holds. Let $d, r \in \mathbb{N}$ with $r \leqslant d$. For any r-dimensional symmetric convex progression $\mathcal{C} \subseteq \mathbb{Z}^d$, there exists a generalised arithmetic progression $P \subseteq \mathbb{Z}^d$ such that*

$$\mathcal{C} \subseteq P, \qquad |P| \leqslant r^{C'r^2}|\mathcal{C}| \qquad \text{and} \qquad \dim(P) \leqslant r = \dim(\mathcal{C}).$$

The bound for the size of the generalised arithmetic progression in Theorem 3.23 has since been improved [18, 78], but the one by Tao and Vu [124] suffices for our purposes. Their statement crucially defines the convex progression on an arbitrary lattice Λ over $\mathbb{R}^d$, rather than $\mathbb{Z}^d$. This allows us to avoid the full rank requirement $r = d$ in their original result, and take $r \leqslant d$ as in Theorem 3.23 by first setting $\Lambda = \mathbf{h} \cap \mathbb{Z}^d$, where $\mathbf{h}$ is an r-dimensional subspace containing $\mathcal{C}$. With this lattice, we then take a suitable linear isomorphism f and apply their original result to $\mathcal{C}' = f(\mathcal{C})$ with $\Lambda' = f(\Lambda)$ and $\mathbb{R}^r = f(\mathbf{h})$.

Using these results, we can now prove Proposition 3.20.

Proof of Proposition 3.20. As all the subgroups of $\mathbb{Z}_n$ are trivial when n is a prime, we claim that we cannot obtain the second case when applying Theorem 3.21 to A. Observe that none of the cosets obtained in the second outcome can be $\mathbb{Z}_n$ itself. The reason for that is that a single one of those cosets would, by (3.70), exceed the total cardinality bound given in Theorem 3.21 if $C > C'$ is sufficiently large:

$$|\mathbb{Z}_n| = n > \exp(C'\kappa^4(\log(\kappa+2))^2)|A|.$$

The claim follows by noting that the cosets also cannot all be of the form $a \in \mathbb{Z}_n$, since their total number is bounded by $C'\kappa^3(\log\kappa)^2$, and, by (3.70), this is insufficient to cover A.

We have shown that we always get the first outcome of Theorem 3.21 under our assumptions, so applying it to A yields a proper coset progression $P'' + H \subseteq \mathbb{Z}_n$ such that

$$A \subseteq P'' + H, \qquad \dim(P'') \leqslant 2\kappa + 1 \qquad \text{and} \qquad |P'' + H| \leqslant \exp(C\kappa^4(\log\kappa)^2)|A|,$$

where we used that $\kappa \geqslant 2$ and that $C > C'$ is appropriately large. Applying now Lemma 3.22 to $P'' + H$ gives a 2-proper coset progression $P' + H' \subseteq \mathbb{Z}_n$ such that

$$A \subseteq P' + H', \quad d = \dim(P') \leqslant 2\kappa + 1 \quad \text{and} \quad |P' + H'| \leqslant \exp(C\kappa^4(\log\kappa)^2)|A|, \tag{3.72}$$

again by $C > C'$. Moreover, $H' = \{0\}$ because $H' = \mathbb{Z}_n$ is impossible by (3.70):

$$\exp(C\kappa^4(\log\kappa)^2)|A| < n = |P' + \mathbb{Z}_n|,$$

so we will write simply P' for $P' + H'$.

Let $\phi : P' \to P'_d \subseteq \mathbb{Z}^d$ be the function defined by

$$\phi(y) = (w_1^{(y)}, \dots, w_d^{(y)}) \quad \text{where} \quad y = a_0 + \sum_{i=1}^{d} w_i^{(y)} a_i$$

and

$$P' = \left\{ a_0 + \sum_{i=1}^{d} w_i a_i : w_i \in \mathbb{Z},\ 0 \leqslant w_i < \ell_i \right\}.$$

Since P' is 2-proper, we know that ϕ is a Freĭman isomorphism; this and $A \subseteq P'$ imply that so is $\phi|_A : A \to \phi(A)$. Let $A_d = \phi(A)$ and assume without loss of generality (by translating ϕ if necessary) that $0 \in A_d$. If we define $\mathbf{h}$ to be the minimal subspace that contains A_d,

then $\mathcal{C}' = \mathbf{h} \cap P'_d$ is a convex progression and $\mathcal{C} = \mathcal{C}' - \mathcal{C}'$ is a symmetric convex progression satisfying $\mathcal{C}' \subseteq \mathcal{C}$. We can therefore apply Theorem 3.23 to $\mathcal{C}$ and obtain a generalised arithmetic progression $P_d \subseteq \mathbb{Z}^d$ which, we claim, satisfies

$$A_d \subseteq P_d, \qquad \dim(P_d) \leqslant \dim_F(A) \qquad \text{and} \qquad |P_d| \leqslant \exp(C\kappa^4(\log\kappa)^2)|A|. \tag{3.73}$$

The containment $A_d \subseteq P_d$ follows from $A_d \subseteq \mathcal{C}' \subseteq \mathcal{C} \subseteq P_d$, where the last step is given by Theorem 3.23. In order to bound the dimension of P_d, let $d_{\mathcal{C}} = \dim(\mathcal{C})$ and observe that

$$\dim(P_d) \leqslant d_{\mathcal{C}} \leqslant \dim(\mathbf{h}) \leqslant \dim_F(A)$$

by Theorem 3.23, $\mathcal{C} \subseteq \mathbf{h}$ and the definition of $\mathbf{h}$ as the minimal subspace containing A_d. Finally, to obtain the claimed bound on the size of P_d, we first use that, by Theorem 3.23,

$$|P_d| \leqslant d_{\mathcal{C}}^{C'd_{\mathcal{C}}^2}|\mathcal{C}| \leqslant \exp(C\kappa^2\log\kappa)|\mathcal{C}| \tag{3.74}$$

where we used that $d_{\mathcal{C}} \leqslant d \leqslant 2\kappa + 1$ by (3.72) and that $C > C'$ is sufficiently large. The second step is showing that the size of $\mathcal{C}$ is at most

$$|\mathcal{C}| = |\mathcal{C}' - \mathcal{C}'| \leqslant |P'_d - P'_d| \leqslant 2^d|P'_d| \leqslant \exp(C\kappa^4(\log\kappa)^2)|A|, \tag{3.75}$$

which is due first to $\mathcal{C}' \subseteq P'_d$, and then by taking a slightly larger C than previous occurrences and using (3.72). Combining (3.74) and (3.75) yields the desired bound on $|P_d|$ for a yet slightly larger value of C, and completes the proof of (3.73).

It remains to define a generalised arithmetic progression $P \subseteq \mathbb{Z}_n$ from $P_d \subseteq \mathbb{Z}^d$. Observe that ϕ^{-1} is linear (or affine), so we can extend it to obtain a $\psi : \mathbb{Z}^d \to \mathbb{Z}_n$ which satisfies

$$\psi|_{A_d} = (\phi|_A)^{-1}. \tag{3.76}$$

Hence, $P = \psi(P_d)$ is a (not necessarily proper) generalised arithmetic progression such that

$$A \subseteq P, \qquad |P| \leqslant \exp(C\kappa^4(\log\kappa)^2)|A| \qquad \text{and} \qquad \dim(P) \leqslant \dim_F(A),$$

by (3.73), the linearity of ψ and (3.76). This is exactly (3.71), so the proof is complete. □

3.8 An upper bound for the independence number

In this section, we prove the upper bound part of Theorem 1.3:

Theorem 3.24. *Let n be a prime number and let $p = p(n)$ satisfy $p \geqslant (\log n)^{-1/80}$. The random Cayley sum graph G_p of $\mathbb{Z}_n$ satisfies*

$$\alpha(G_p) \leqslant \big(2 + o(1)\big)\log_{\frac{1}{1-p}} n$$

with high probability as $n \to \infty$.

Throughout, we fix a small enough $\delta > 0$ and $k = (2 + 4\delta) \log_{\frac{1}{1-p}} n$. We will also follow the outline presented in Section 3.1 and use the notation defined there. Each sub-collection $\mathcal{A}_i$ requires different techniques to bound the probability that $\alpha(G_p) > k$, so we handle each separately and show all three go to 0 as $n \to \infty$.

3.8.1 Bounding the probability over choices in $\mathcal{A}_1$

A brief recap: in this doubling range, there are far too many choices for $A \in \mathcal{A}_1$ for a union bound to work. The key observation is that we do not need to count every such A: if somehow going to a smaller subset $\tilde{A} \subseteq A$ reduces our choices considerably, this would also be enough, since the event $\{\tilde{A} \hat{+} \tilde{A} \subseteq S^c\}$ contains $\{A \hat{+} A \subseteq S^c\}$. As a matter of fact, we will use this idea twice.

The first time we use this idea is in order to replace each set $A \in \mathcal{A}_1$ by a large subset $A' \subseteq A$ whose Freĭman dimension is robust in the sense required by Theorem 3.2, our supersaturation result. We build the set A' from A by greedily removing small, "bad" subsets $B \subseteq A$ such that removing them from A reduces its Freĭman dimension. Each such removal reduces $\dim_F(A)$ by at least one, so we finish in at most $\dim_F(A) - 1$ steps. A step reduces the size of the working set by at most $\varepsilon|A|$, which would already give $|A'| \geqslant (1 - \dim_F(A)\varepsilon)|A|$. To get a bound that depends on σ instead of on $\dim_F(A)$, we use the following trivial consequence of Freĭman's lemma, Lemma 3.3, which we record for ease of reference later.

Observation 3.25. *For all $A \subseteq \mathbb{Z}_n$ with $\sigma[A] \leqslant \sigma$, we have $\dim_F(A) + 1 \leqslant 2\sigma$.*

Proof. Let $d = \dim_F(A)$, and take $A' \subseteq \mathbb{Z}^d$ of full rank to be the image of A under a Freĭman isomorphism ϕ. Applying Freĭman's lemma to A' yields

$$|A' + A'| \geqslant (d+1)|A'| - \binom{d+1}{2}$$

but $|A' + A'| \leqslant \sigma|A'|$ and $|A'| = |A|$ by our choice of ϕ. Dividing both sides by $|A|$ yields

$$\sigma \geqslant d + 1 - \frac{d+1}{2}$$

where we used $|A| \geqslant d$ to simplify the right-hand side. The proof follows by rearranging. □

The proof of the following proposition is just formalizing the sketch using Observation 3.25.

Proposition 3.26. *Let $A \subseteq \mathbb{Z}_n$ with $\sigma[A] \leqslant \sigma$ and let $\varepsilon < 1/2\sigma$. There exists $d \in \mathbb{N}$ and $A' \subseteq A$ such that $|A'| > (1 - 2\varepsilon\sigma)|A|$ and A' has ε-robust Freĭman dimension d.*

Proof. We start an iterative process with $A_0 = A$ and $i = 0$. While there exists $B \subseteq A_i$ such that

$$|B| \leqslant \varepsilon|A| \quad \text{and} \quad \dim_F(A_i \setminus B) < \dim_F(A_i),$$

we define $A_{i+1} = A_i \setminus B$. Let t be the number of steps in this process, and let $A' = A_t$.

First, notice that A' has ε-robust Freĭman dimension because there is no $B \subseteq A'$ to continue the process. Moreover, $t \leqslant \dim_F(A) - 1$ since each step reduces $\dim_F(A_i)$ by at least one. Since

$|A_i| \geqslant |A| - i\varepsilon|A|$, it follows that

$$|A'| \geqslant (1-\varepsilon t)|A| > (1-\varepsilon \dim_{\mathrm{F}}(A))|A| > (1-2\varepsilon\sigma)|A|,$$

where in the last inequality we used Observation 3.25. □

The second time we employ the idea of going to smaller subsets is when we use the fingerprints given by Theorem 3.1, whose statement we repeat here for convenience.

Theorem 3.1. *Let n be a large enough prime and let $k, d \in \mathbb{N}$. For every $0 < \gamma < 1/2$, there exists $C = C(\gamma) > 0$ such that the following holds for all $m \geqslant (d+1)k/2$ and $C/k < \varepsilon < \gamma$. For each d-dimensional generalised arithmetic progression $P \subseteq \mathbb{Z}_n$, there exists a collection $\mathcal{F} = \mathcal{F}_{k,m,\varepsilon}(P)$ of subsets of P satisfying:*

(1) For every $F \in \mathcal{F}$, we have

$$|F| \leqslant C\varepsilon^{-1}\sqrt{m \log m} \quad \text{and} \quad |F \hat{+} F| \geqslant \frac{(1-\gamma)(d+1)k}{2}. \tag{3.77}$$

(2) For all $A \in \binom{P}{k}$ with $|A \hat{+} A| \leqslant m$ and ε-robust Freĭman dimension d, there exists $F \in \mathcal{F}$ such that $F \subseteq A$.

As we remarked in the overview, applying the previous theorem to the (trivial) generalised arithmetic progression $\mathbb{Z}_n$ would result in too many fingerprints. To overcome this, we use Proposition 3.20, proved in the previous section. With this, we are now ready to prove that

$$\mathbb{P}(\exists\, A \in \mathcal{A}_1 : A \hat{+} A \subseteq S^{\mathsf{c}}) \to 0 \quad \text{as} \quad n \to \infty. \tag{3.78}$$

Proof of (3.78). Recall that $k = (2+4\delta)\log_{\frac{1}{1-p}} n$, and that for all $A \in \mathcal{A}_1$, we know that $|A| = k$ and $\sigma[A] \leqslant k^{1/40}$. Fixing $\varepsilon = k^{-1/20}$, we can apply Proposition 3.26 to A with ε and $\sigma = k^{1/40}$ to conclude that every such set contains an A' of size at least

$$|A'| \geqslant (1-2\varepsilon\sigma)k \geqslant (1-2k^{-1/40})k \tag{3.79}$$

with ε-robust Freĭman dimension d_A, for some $d_A \in \mathbb{N}$. Observe that (3.79) implies that the doubling of A' is at most 2σ:

$$\sigma[A'] = \frac{|A'+A'|}{|A'|} \leqslant \frac{|A+A|}{|A'|} \leqslant 2\sigma, \tag{3.80}$$

for all sufficiently large k. We fix[6] one A' for each $A \in \mathcal{A}_1$, and denote by $\mathcal{A}_1'$ the collection of all such A'.

As we remarked before, $A' \hat{+} A' \subseteq S^{\mathsf{c}}$ is implied by $A \hat{+} A \subseteq S^{\mathsf{c}}$ for each $A \in \mathcal{A}_1$. Therefore, we have the bound

$$\mathbb{P}\big(\exists\, A \in \mathcal{A}_1 : A \hat{+} A \subseteq S^{\mathsf{c}}\big) \leqslant \mathbb{P}\big(\exists\, A' \in \mathcal{A}_1' : A' \hat{+} A' \subseteq S^{\mathsf{c}}\big)$$

[6] Abusing notation to denote such a mapping via the $'$ symbol.

where, here and throughout, S is a p-random subset of $\mathbb{Z}_n$. Moreover, let $\mathcal{Y}(P)$ be the collection of subsets $Y \subseteq P$ with

$$(1-2k^{-1/40})k \leqslant |Y| \leqslant k \qquad \text{and} \qquad \sigma[Y] \leqslant 2\sigma \tag{3.81}$$

such that Y has ε-robust Freĭman dimension d_Y for some $d_Y \geqslant \dim(P)$. We claim that we can take another union bound:

$$\mathbb{P}(\exists\, A' \in \mathcal{A}_1' : A' \hat{+} A' \subseteq S^{\mathsf{c}}) \leqslant \sum_{P \in \mathcal{P}(\mathbb{Z}_n)} \mathbb{P}(\exists\, Y \in \mathcal{Y}(P) : Y \hat{+} Y \subseteq S^{\mathsf{c}}) \tag{3.82}$$

where $\mathcal{P}(\mathbb{Z}_n)$ is the collection of generalised arithmetic progressions $P \subseteq \mathbb{Z}_n$ such that

$$|P| \leqslant \exp(k^{1/5}).$$

In order to prove (3.82), it is enough to show that, for every $A' \in \mathcal{A}_1'$, there exists a generalised arithmetic progression $P \in \mathcal{P}(\mathbb{Z}_n)$ such that $A' \in \mathcal{Y}(P)$. We do so by applying Proposition 3.20 with $\kappa = 2\sigma$. From this application, we will obtain a generalised arithmetic progression $P \in \mathcal{P}(\mathbb{Z}_n)$ such that

$$\dim(P) \leqslant \dim_{\mathrm{F}}(A') = d_A \qquad \text{and} \qquad A' \subseteq P. \tag{3.83}$$

The property $P \in \mathcal{P}(\mathbb{Z}_n)$ follows from (3.71), as

$$|P| \leqslant \exp(C'\sigma^4(\log\sigma)^2)k \leqslant \exp\big(C'k^{1/10}(\log k)^2\big)k \leqslant \exp\big(k^{1/5}\big), \tag{3.84}$$

where we used that $\sigma = k^{1/40}$ in the second inequality, and in the last one we used that k is sufficiently large. We now confirm that we can apply Proposition 3.20 as we want, by verifying that every $A' \in \mathcal{A}_1'$ satisfies (3.70). The lower bound holds because

$$|A'| \geqslant (1-2k^{-1/40})k \geqslant k^{1/10} \geqslant C'\sigma^3(\log\sigma)^2,$$

by (3.79) and using that k is sufficiently large, whereas the upper bound in (3.70) follows from

$$\exp(C'\sigma^4(\log\sigma)^2)|A'| \leqslant \exp(k^{1/5}) \leqslant \exp((\log n)^{2/5}) < n,$$

where we used (3.84), and that $k \leqslant (\log n)^2$ and n is sufficiently large.

We now claim that $A' \in \mathcal{Y}(P)$, where $P \in \mathcal{P}(\mathbb{Z}_n)$ is the generalised arithmetic progression given by applying Proposition 3.20 to $A' \in \mathcal{A}_1'$. To see this, simply note that the conditions in (3.81) follow from (3.79) and (3.80), and the ε-robust Freĭman dimension bound $d_A \geqslant \dim(P)$ follows from (3.83) and Proposition 3.26, so this completes the proof of (3.82).

In order to bound the term in the right-hand side of (3.82), we analyse the contribution of each fixed $P \in \mathcal{P}(\mathbb{Z}_n)$. Notice that if $\mathcal{Y}(P)$ is empty, then the probability term is equal to 0, so we may assume that $\mathcal{Y}(P)$ is non-empty. Rather than directly taking a union bound over

choices of $Y \in \mathcal{Y}(P)$, our final union bound is over a collection of fingerprints $\mathcal{F}(P)$:

$$\begin{aligned}\mathbb{P}(\exists Y \in \mathcal{Y}(P) : Y \hat{+} Y \subseteq S^{\mathsf{c}}) &\leqslant \mathbb{P}(\exists F \in \mathcal{F}(P) : F \hat{+} F \subseteq S^{\mathsf{c}}) \\ &\leqslant |\mathcal{F}(P)| \max_{F \in \mathcal{F}(P)} \mathbb{P}(F \hat{+} F \subseteq S^{\mathsf{c}}). \end{aligned} \tag{3.85}$$

That is true as long as, for each $Y \in \mathcal{Y}(P)$, there exists $F \in \mathcal{F}(P)$ such that $F \subseteq Y$; again we are using that $F \hat{+} F \subseteq Y \hat{+} Y$. We claim that applying Theorem 3.1 to P with $\gamma = \gamma(\delta)$ (to be determined later) and $m = 2\sigma k$ yields such a collection of fingerprints $\mathcal{F}(P)$.

First, we define a candidate for $\mathcal{F}(P)$ which proves our claim, and later we show that we can construct this candidate. Our candidate for $\mathcal{F}(P)$ is defined as

$$\mathcal{F}(P) = \bigcup_{Y \in \mathcal{Y}(P)} \mathcal{F}_{|Y|,m,\varepsilon}(P) \tag{3.86}$$

where $\mathcal{F}_{|Y|,m,\varepsilon}(P)$ is the collection of fingerprints given by Theorem 3.1. For $\mathcal{F}(P)$ to prove our claim, we must thus show that, for each $Y \in \mathcal{Y}(P)$, there is an $F \in \mathcal{F}_{|Y|,m,\varepsilon}(P)$ such that $F \subseteq Y$. By property 2 of Theorem 3.1, it suffices for each Y to have ε-robust Freĭman dimension $d_Y \geqslant \dim(P)$ and $|Y + Y| \leqslant m = 2\sigma k$. These, however, hold for Y by (3.81), and so we have that $\mathcal{F}(P)$ is a valid candidate of fingerprints for P.

To confirm that we can apply Theorem 3.1 to P as in (3.86), we need to check that

$$m \geqslant \frac{|Y|(\dim(P)+1)}{2} \tag{3.87}$$

for all $Y \in \mathcal{Y}(P)$. (3.87) follows from

$$\frac{|Y|(\dim(P)+1)}{2} \leqslant \frac{|Y|(\dim_{\mathrm{F}}(Y)+1)}{2} \leqslant |Y+Y| \leqslant 2\sigma k = m$$

where the first inequality uses that $\dim(P) \leqslant \dim_{\mathrm{F}}(Y)$ by definition of $\mathcal{Y}(P)$, the second inequality relies on Observation 3.25 and the third one on (3.81). Therefore, (3.86) is a valid definition for $\mathcal{F}(P)$.

We proceed to give an upper bound to the right-hand side of (3.85). With the goal of first bounding the size of $\mathcal{F}(P)$, define $\Phi(P) = \max\{|F| : F \in \mathcal{F}(P)\}$ and note that, trivially,

$$|\mathcal{F}(P)| \leqslant \sum_{q=0}^{\Phi(P)} \binom{|P|}{q} \leqslant (|P|+1)^{\Phi(P)}.$$

Since, by (3.77) in Theorem 3.1, $|F| \leqslant C\varepsilon^{-1}\sqrt{m \log m}$ for all $F \in \mathcal{F}(P)$, we can use that $m = 2\sigma k$, $\varepsilon = k^{-1/20}$ and $\sigma = k^{1/40}$ to obtain

$$|\mathcal{F}(P)| \leqslant \exp(k^{3/5} \log |P|) \leqslant \exp(k^{4/5}) \tag{3.88}$$

where in the first inequality we used that k is sufficiently large, and in the last we used (3.84).

To obtain an upper bound on $\mathbb{P}(F \hat{+} F \subseteq S^{\mathsf{c}})$ for all $F \in \mathcal{F}(P)$, we use (3.77), the lower

bound on $|F \hat{+} F|$ given by Theorem 3.1:

$$|F \hat{+} F| \geqslant \frac{(1-\gamma)(\dim(P)+1)}{2} \min_{Y \in \mathcal{Y}(P)} |Y|.$$

Since $\gamma = \gamma(\delta)$ is a constant and $|Y| \geqslant (1-2\varepsilon\sigma)k$ for all $Y \in \mathcal{Y}(P)$ by (3.81), we obtain

$$\max_{F \in \mathcal{F}(P)} \mathbb{P}(F \hat{+} F \subseteq S^{\mathsf{c}}) \leqslant (1-p)^{(1-2\gamma)(\dim(P)+1)k/2}, \tag{3.89}$$

as k is sufficiently large. Replacing (3.88) and (3.89) back into (3.85) yields

$$\mathbb{P}(\exists\, Y \in \mathcal{Y}(P) : Y \hat{+} Y \subseteq S^{\mathsf{c}}) \leqslant \exp(k^{4/5})(1-p)^{(1-2\gamma)(\dim(P)+1)k/2}. \tag{3.90}$$

This will be our bound for each term in the right-hand side of (3.82).

Observe that (3.90) does not depend on the specific choice of P, only on its size and dimension. Recalling that $|P| \leqslant \exp(k^{1/5}) =: s$ for every $P \in \mathcal{P}(\mathbb{Z}_n)$ by (3.84), we can group terms in the right-hand side of (3.82) based on $d = \dim(P)$ to deduce that, by (3.90), we have

$$\mathbb{P}(\exists\, A \in \mathcal{A}_1 : A \hat{+} A \subseteq S^{\mathsf{c}}) \leqslant \sum_{d=1}^{\infty} s(ns)^{d+1} \exp(k^{4/5})(1-p)^{(1-2\gamma)(d+1)k/2}, \tag{3.91}$$

where we have bounded the number of d-dimensional generalised arithmetic progressions in $\mathbb{Z}_n$ with size at most s by $s(ns)^{d+1}$.

It is therefore enough to prove that the right-hand side of (3.91) goes to 0 as $n \to \infty$. Note that, as $k = (2+4\delta)\log_{\frac{1}{1-p}} n$, we have

$$(1-p)^{k/2} = n^{-(1+2\delta)},$$

which combined with a suitably small choice of $\gamma = \gamma(\delta)$, implies that

$$n^{d+1}(1-p)^{(1-2\gamma)(d+1)k/2} = n^{d+1-(1-2\gamma)(1+2\delta)(d+1)} \leqslant n^{-\delta(d+1)}. \tag{3.92}$$

The final observation is that it follows from $s = \exp(k^{1/5})$ that there is $1 > \nu > 0$ such that

$$s^{d+2} \exp(k^{4/5}) \leqslant \exp\left((d+1)(\log n)^{1-\nu}\right) \tag{3.93}$$

since $k \leqslant 3(\log n)/p$ for $\delta < 1/4$ and n is sufficiently large. The value of ν depends on the exponent of the $\log n$ term in our choice of p, and $\nu = 0.18$ works for $p \geqslant (\log n)^{-1/80}$.

Combining (3.92) with (3.93), we show, for sufficiently large n, that (3.91) is at most

$$\sum_{d=1}^{\infty} s(ns)^{d+1} \exp(k^{4/5})(1-p)^{(1-2\gamma)(d+1)k/2} \leqslant \sum_{d=1}^{\infty} n^{-\delta d/2}$$

which goes to 0 as $n \to \infty$, as we wanted to show. □

3.8.2 Bounding the probability over choices in $\mathcal{A}_2$

Our goal is to show that the term corresponding to choices in $\mathcal{A}_2$ tends to 0 as $n \to \infty$. We emphasize that, in this section, no new ideas, or even modification of previous, existing results, are needed. We just need to use the following result of Green [72].

Proposition 3.27 ([72], see [73, Proposition 6.1]). *For every $k \in \mathbb{N}$ and $m \geqslant 2k-1$, there exists $r \in \mathbb{N}$ with*

$$r \leqslant \min\{4m/k, k\} \quad \text{and} \quad r \leqslant \frac{2m}{k} + \frac{1}{k}\binom{r}{2}, \tag{3.94}$$

such that

$$\left|\mathcal{A}_2^{(m)}\right| \leqslant n^r k^{4k},$$

where $\mathcal{A}_2^{(m)} = \{A \in \mathcal{A}_2 : |A \hat{+} A| = m\}$.

With Proposition 3.27, we can take a union bound over each $A \in \mathcal{A}_2$ to obtain

$$\mathbb{P}\left(\exists\, A \in \mathcal{A}_2 : A \hat{+} A \subseteq S^{\mathsf{c}}\right) \leqslant \sum_{m=k^{1+1/40}}^{\delta k^2/10} \left|\mathcal{A}_2^{(m)}\right| (1-p)^m \to 0 \quad \text{as } n \to \infty, \tag{3.95}$$

and we prove the last step below.

Proof of (3.95). First, we bound the r in Proposition 3.27 using (3.94):

$$r \leqslant \frac{2m}{k} + \frac{1}{k}\left(\frac{4m}{k}\right)^2 \leqslant \frac{m}{k}(2+2\delta)$$

since sets $A \in \mathcal{A}_2^{(m)} \subseteq \mathcal{A}_2$ satisfy $m/k = \sigma[A] \leqslant \delta k/10$. Now, applying Proposition 3.27 to each $\mathcal{A}_2^{(m)}$, we obtain that

$$\left|\mathcal{A}_2^{(m)}\right| \leqslant n^r k^{4k} \leqslant n^{(2+2\delta)m/k + 5(\log n)^{1/50}\log\log n} \tag{3.96}$$

and the last inequality follows from $k \leqslant (\log n)^{51/50}$, which is implied by $p \geqslant (\log n)^{-1/80}$.

Notice that it follows from $k = (2+4\delta)\log_{\frac{1}{1-p}} n$ that

$$(1-p)^m = n^{-(2+4\delta)m/k}. \tag{3.97}$$

Together with (3.96) and the fact that $m/k = \sigma[A] \geqslant k^{1/40} \geqslant (\log n)^{1/40}$, (3.97) yields

$$\left|\mathcal{A}_2^{(m)}\right| (1-p)^m \leqslant n^{-\delta m/k}$$

from which the result follows by summing over all $k^{1+1/40} \leqslant m \leqslant \delta k^2/10$. □

3.8.3 Bounding the probability over choices in $\mathcal{A}_3$

Similarly to the previous case, we start with a union bound

$$\mathbb{P}\left(\exists\, A \in \mathcal{A}_3 : A \hat{+} A \subseteq S^{\mathsf{c}}\right) \leqslant \sum_{m=\delta k^2/10}^{k^2} \left|\mathcal{A}_3^{(m)}\right| (1-p)^m, \tag{3.98}$$

letting $\mathcal{A}_3^{(m)}$ be the sub-collection of $\mathcal{A}_3$ consisting of subsets $A \subseteq \mathbb{Z}_n$ with $|A \hat{+} A| = m$.

We will need the following slight strengthening of [73, Proposition 5.1].

Proposition 3.28. *Let $\delta > 0$ be sufficiently small and let $\eta > 0$. If $k \leqslant (\log n)^{2-\eta}$ and $m \geqslant \delta k^2/10$, then*

$$\left|\mathcal{A}_3^{(m)}\right| \leqslant n^{(2+\delta+o(1))m/k},$$

as $k \to \infty$, where $\mathcal{A}_3^{(m)} = \{A \in \mathcal{A}_3 : |A \hat{+} A| = m\}$.

It is easy to complete the proof of Theorem 1.3 with Proposition 3.28.

Proof that (3.98) $\to 0$ *as* $n \to \infty$. Observe that

$$k \leqslant \frac{3\log n}{p} \leqslant (\log n)^{41/40} \leqslant (\log n)^{2-\eta},$$

by our choice of $p \geqslant (\log n)^{-1/80}$. Hence, we can apply Proposition 3.28 to bound, for every $m \geqslant \delta k^2/10$, the size of $\mathcal{A}_3^{(m)}$ by

$$\left|\mathcal{A}_3^{(m)}\right| \leqslant n^{(2+\delta+o(1))m/k} \leqslant n^{(2+2\delta)m/k}$$

where the last inequality holds for large k. We can bound each term of (3.98) by:

$$\left|\mathcal{A}_3^{(m)}\right| (1-p)^m \leqslant n^{(2+2\delta)m/k} n^{-(2+4\delta)m/k} \leqslant n^{-2\delta m/k}.$$

Summing over $m \geqslant \delta k^2/10$ yields the desired result. □

The proof of Proposition 3.28 is essentially identical to that of [73, Proposition 5.1], and to obtain the statement above we only need to optimize the dependencies between the parameters. Before proceeding, we recall a key definition from [73]:

Definition 3.29. A set $\{a_1, \dots, a_d\} \subseteq \mathbb{Z}_n$ is M-dissociated if

$$\sum_{i=1}^{d} \lambda_i a_i \neq 0$$

for every $(\lambda_1, \dots, \lambda_d) \in \mathbb{Z}^d \setminus \{0\}$ such that $\sum_{i=1}^{d} |\lambda_i| \leqslant M$.

It is useful to understand this notion as an analogue of linear independence for elements of $\mathbb{Z}_n$ taking coefficients in $\mathbb{Z}$, with a restriction on the sum of their magnitudes. The first result that we need to optimize is Lemma 3.30, a slight strengthening of [73, Lemma 5.3].

Lemma 3.30. *For every sufficiently small $\delta > 0$ and $\eta > 0$, the following holds. Fix a large prime n and $M = (\log n)/(\log\log n)^4$. If $k \leqslant (\log n)^{2-\eta}$ and $m \geqslant \delta k^2/10$, then any M-dissociated subset of $A \in \mathcal{A}_3^{(m)}$ has size at most $\big(2+\delta+o(1)\big)m/k$, as $k \to \infty$.*

We will begin by giving an overview of the proof from [73], alongside some useful definitions of our own. Whenever there is $(\lambda_1, \dots, \lambda_d) \in \mathbb{Z}^d$ such that $x = \sum_{i=1}^d \lambda_i a_i$ and $\sum_{i=1}^d |\lambda_i| \leqslant M$, we will say that x is in the M-span of $\{a_1, \dots, a_d\}$. In that same situation, we will say that x can be written as an M-bounded combination of $\{a_1, \dots, a_d\}$.

The proof starts by fixing the value of $M = (\log n)/(\log\log n)^4$ and any M-dissociated set $D = \{a_1, \dots, a_d\} \subseteq A$. We define G' to be the graph with vertex set $\mathbb{Z}_n$ and edges between a pair of vertices if their difference can be written as $\pm a_i \pm a_j$ for some $i, j \in \{1, \dots, d\}$. The graph G' is useful because of the following observation:

Observation 3.31. *For each $a \in \mathbb{Z}_n$, let $V_a \subseteq V(G')$ denote its connected component in G'. Every element $b \in V_a$ can be expressed as $b = a + \sum_{j=1}^d \lambda_j a_j$ with $\sum_{j=1}^d |\lambda_j| \leqslant 2\,\mathrm{diam}(V_a)$.*

Proof. Take $b \in V_a$, and consider any shortest path P connecting a to b in G'. Observe that

$$b - a = \sum_{j=1}^{d} (\lambda_j^+ - \lambda_j^-) a_j$$

where λ_j^+, λ_j^- count respectively the number of times $a_j, -a_j$ appear as terms in P. Then,

$$\sum_{j=1}^{d} \big(|\lambda_j^+| + |\lambda_j^-|\big) \leqslant 2|P| \leqslant 2\,\mathrm{diam}(V_a)$$

where the last step follows from P being one of the shortest path in V_a. □

Having thus related the span of D with the diameter of connected components in G', we will decompose $G = G'[A]$ using the following lemma, also from [73], with $\Delta = k^{1/2}\log k$:

Lemma 3.32 ([73, Lemma 5.4])**.** *Let G be a graph with vertex set $V(G)$ and let $\Delta > 1$. There exists a partition $V(G) = A_* \cup A_1 \cup \cdots \cup A_t$ such that*

1. *$|A_*| \leqslant 32(v(G)/\Delta)^2$.*
2. *$e(A_i, A_j) = 0$ for every $i \neq j$.*
3. *The diameter of $G[A_i]$ is at most Δ for every $i \in \{1, \dots, t\}$.*

Applying Lemma 3.32 to G yields a partition $A = A_* \cup A_1 \cup \cdots \cup A_t$ such that

(i) if $A_i \neq A_j$, then $A_i + D$ and $A_j + D$ are disjoint, and

(ii) every A_i has a $y_i \in A_i$ such that $A_i - y_i$ is contained in the 2Δ-span of D.

To see that property (i) holds, notice that if $A_i + D$ and $A_j + D$ intersect, then there is an edge of G connecting A_i and A_j, contradicting item (2) of Lemma 3.32. Property (ii), on the other

hand, follows directly from an application of Observation 3.31 to A_i, the diameter of which is bounded by item (3) in Lemma 3.32.

To combine these two properties into a proof of Lemma 3.30, notice that the disjointness given by property (i) yields a lower bound for $A + A$ in terms of each A_i:

$$|A + A| \geqslant \sum_i |A_i + D|.$$

We now want a lower bound for each $|A_i + D|$ in terms of d. Towards that goal, we will define, for every $i \in \{1, \dots, t\}$, a Freĭman isomorphism $\phi_i : A_i \to A_i' \subseteq \mathbb{Z}^d$. Each will map (a translate of) D to a structured set, where structured only means that we have a lower bound for its sumset with any sufficiently small set. These isomorphisms are naturally defined by the 2Δ-bounded decomposition of $A_i - y_i$ given by property (ii).

Observation 3.33. *The function $\phi_i : A_i \to A_i' \subseteq \mathbb{Z}^d$ defined by $\phi_i(x) = (\lambda_1, \dots, \lambda_d)$, where $x - y_i = \sum_{j=1}^d \lambda_j a_j$ is a 2Δ-bounded decomposition of $x - y_i$, is a Freĭman isomorphism.*

Proof. The fact that $\phi^{-1} : A_i' \to A_i$ is a Freĭman homomorphism follows from its linearity, so we focus on the other direction of the implication. Take $a_1, a_2, a_3, a_4 \in A_i$ such that

$$a_1 + a_2 = a_3 + a_4, \tag{3.99}$$

and we want to show that $\phi_i(a_1) + \phi_i(a_2) = \phi_i(a_3) + \phi_i(a_4)$. For $\ell \in \{1, \dots, 4\}$, write

$$a_\ell - y_i = \sum_{i=1}^d a_i \lambda_i^{(\ell)}, \tag{3.100}$$

a 2Δ-bounded combination of D. We now replace (3.100) in (3.99) and rearrange to obtain

$$\sum_{j=1}^d a_j (\lambda_j^{(1)} + \lambda_j^{(2)} - \lambda_j^{(3)} - \lambda_j^{(4)}) = 0.$$

Using the fact that each (3.100) is 2Δ-bounded yields, as $k \to \infty$,

$$\sum_{\ell=1}^4 \sum_{j=1}^d |\lambda_j^{(\ell)}| \leqslant 4(2\Delta) \leqslant 8k^{1/2} \log k \leqslant (\log n)^{1-\eta/3} \leqslant M,$$

where the inequalities follow by the definitions/assumptions which we now recall

$$\Delta = k^{1/2} \log k, \qquad k \leqslant (\log n)^{2-\eta} \qquad \text{and} \qquad M = (\log n)/(\log \log n)^4.$$

However, D is M-dissociated, so we conclude that, for all $j \in \{1, \dots, d\}$,

$$(\lambda_j^{(1)} + \lambda_j^{(2)}) - (\lambda_j^{(3)} + \lambda_j^{(4)}) = 0,$$

which, by the definition of ϕ_i, is just another way to write $\phi_i(a_1) + \phi_i(a_2) = \phi_i(a_3) + \phi_i(a_4)$. □

Observe that $\phi_i(D + y_i) = \{e_1, \dots, e_d\}$, where $\{e_1, \dots, e_d\}$ is the canonical basis of $\mathbb{Z}^d$.

Hence, to obtain a lower bound for the size of $A_i + D$ in terms of d and complete the proof of Lemma 3.30, we will bound $|\phi_i(A_i) + \phi_i(D + y_i)|$ using the isoperimetric inequality of Wang and Wang [129], as stated in [73] (see also [20]).

Proposition 3.34 ([129], see [73, "The Isoperimetric Inequality"])**.** *For every $\gamma > 0$ and $C > 0$, there exists $d_0 = d_0(\gamma, C)$ such that the following holds for every $d \geqslant d_0$. If $S \subseteq \mathbb{Z}^d$ is a set of size at most Cd, then*

$$\big|S + \{e_1, \dots, e_d\}\big| \geqslant \left(\frac{1}{2} - \gamma\right) d|S|.$$

Proof of Lemma 3.30. Fix $A \in \mathcal{A}_3^{(m)}$ and let $D = \{a_1, \dots, a_d\}$ be any M-dissociated subset of A. Towards our aim of showing that $|D| = d \leqslant \big(2 + \delta + o(1)\big)m/k$, recall the definition of the graph G': its vertices are $\mathbb{Z}_n$ and there are edges between vertices $a, b \in \mathbb{Z}_n$ only when some of $\pm(a - b)$ can be written as either $a_i - a_j$ or $a_i + a_j$ for $i, j \in \{1, \dots, d\}$.

We apply Lemma 3.32 to $G = G'[A]$ with $\Delta = k^{1/2} \log k$ and obtain $A_* \cup A_1 \cup \dots \cup A_t$, a partition of A such that $e(A_i, A_j) = 0$ for all $i \neq j$,

$$\operatorname{diam}(G[A_i]) \leqslant k^{1/2} \log k$$

and

$$|A_*| \leqslant 32 \left(\frac{k}{k^{1/2} \log k}\right)^2 = o(k). \tag{3.101}$$

In order to obtain a lower bound for $A + A$, we use property (i) of G to show that

$$|A + A| \geqslant \sum_{i=1}^{t} |A_i + D|. \tag{3.102}$$

Now, we define, for each $i \in \{1, \dots, t\}$, the function ϕ_i; by property (ii) and Observation 3.33, we have that ϕ_i is a Freĭman isomorphism from A_i to $A_i' \subseteq \mathbb{Z}^d$. By the definition of Freĭman isomorphisms, we obtain

$$|A_i + D| = |A_i + (D + y_i)| = |A_i' + \phi_i(D + y_i)|. \tag{3.103}$$

Recall that the definition of ϕ_i implies that $\phi_i(D + y_i) = \{e_1, \dots, e_d\}$. Hence, to apply Proposition 3.34 to the right-hand side of (3.103), we need to bound $|A_i'|$ in terms of d and show that d is sufficiently large. We will show both of these by observing that if $k > 10d/\delta$, then we are already done, since

$$d \leqslant \frac{\delta k}{10} \leqslant \frac{m}{k} \leqslant \big(2 + \delta + o(1)\big)\frac{m}{k},$$

which is what we wanted to show. We can therefore assume that $d \geqslant \delta k/10$, which simultaneously implies that d is large when $k \to \infty$, and that $|A_i'| \leqslant k \leqslant 10d/\delta$. Applying Proposition 3.34 to A_i' with $\gamma = \delta^2$ and $C = 10/\delta$, we obtain

$$\big|A_i' + \{e_1, \dots, e_d\}\big| \geqslant \left(\frac{1}{2} - \delta^2\right) d|A_i'| \tag{3.104}$$

for each $i \in \{1, \dots, t\}$.

Replacing (3.104) in (3.102), and using (3.103), we conclude that

$$|A+A| \geqslant d\left(\frac{1}{2}-\delta^2\right)\sum_{i=1}^{t}|A_i'|.$$

Thus, using the fact that $|A_i'| = |A_i|$, which follows from each ϕ_i being a Freĭman isomorphism, and that $A = A_* \cup A_1 \cup \cdots A_t$ is a partition, we have

$$d\left(\frac{1}{2}-\delta^2\right)\sum_{i=1}^{t}|A_i'| = d\left(\frac{1}{2}-\delta^2\right)\big(|A|-|A_*|\big) = d\left(\frac{1}{2}-\delta^2\right)(1-o(1))k$$

where we used (3.101) to bound $|A_*|$. Finally, it follows from $|A+A| \leqslant m$ and δ being sufficiently small that $d \leqslant \big(2+\delta+o(1)\big)m/k$ as $k \to \infty$. □

With Lemma 3.30, we need only one more piece, the following lemma. It is a simple count of the number of choices for coefficients in a dissociated set, and we will use it to repeat the proof of Proposition 5.1 in Green and Morris [73] and obtain Proposition 3.28.

Lemma 3.35 ([73, Lemma 5.5])**.** *For every $M, d \in \mathbb{N}$, the number of choices for $\lambda_1, \ldots, \lambda_d \in \mathbb{Z}$ such that $\sum_{i=1}^{d}|\lambda_i| \leqslant M$ is at most $(4d)^M$.*

Proof of Proposition 3.28. Fix $M = (\log n)/(\log\log n)^4$ and take $A \in \mathcal{A}_3^{(m)}$. We will count the choices for A by first choosing a maximal M-dissociated subset $D = \{a_1, \ldots, a_d\} \subseteq A$ and then choosing the remaining elements of $A \setminus D$ using the properties of D.

First, we count the choices for D naively, and rely on Lemma 3.30 to bound its size. That is, we apply Lemma 3.30 to D, obtain $|D| \leqslant d \leqslant \big(2+\delta+o(1)\big)m/k$, and thus deduce that the number of choices for D is at most:

$$\sum_{t=1}^{d}\binom{n}{t} \leqslant (n+1)^d \leqslant n^{(2+\delta+o(1))m/k}. \tag{3.105}$$

The second step is counting the choices for $A \setminus D$, and we do so by counting the possible ways to write each of its elements as an M-bounded combination of D. Fix $a' \in A \setminus D$, and note that the maximality of D implies that there is $\Lambda = \{\lambda_0, \lambda_1, \ldots, \lambda_d\} \subseteq \mathbb{Z}$ such that

$$\lambda_0 a' + \sum_{j=1}^{d}\lambda_j a_j = 0$$

and the elements of Λ satisfy

$$\sum_{j=0}^{d}|\lambda_j| \leqslant M, \qquad \lambda_0 \neq 0 \qquad \text{and} \qquad \lambda_i \neq 0 \text{ for some } i > 0.$$

We can therefore use Lemma 3.35 to count the number of choices for Λ and observe that

$$(4d+4)^M \leqslant \exp(3M\log k) \leqslant \exp\left(\frac{\log n}{(\log\log n)^2}\right) \leqslant n^{1/\log k}$$

where the first inequality follows from the (trivial) observation that $d \leqslant k$, and the rest is a consequence of our choice of $M = (\log n)/(\log\log n)^4$ and our assumption that $k \leqslant (\log n)^{2-\eta}$.

We can now choose each of the $k-d$ elements of $A \setminus D$ by the above procedure, and obtain that there are at most

$$n^{(k-d)/(\log k)} = n^{o(k)} \tag{3.106}$$

such elements. Combining (3.105) and (3.106) with $m \geqslant \delta k^2/10$ thus yields

$$\left|\mathcal{A}_3^{(m)}\right| \leqslant n^{(2+\delta+o(1))m/k+o(k)} = n^{(2+\delta+o(1))m/k},$$

as required. □

3.9 The lower bound

In this section, we prove the lower bound in Theorem 1.3.

Theorem 3.36. *Let $\delta > 0$, $n \in \mathbb{N}$ be a prime number and let $p = p(n)$ satisfy $n^{-\delta/8} \leqslant p \leqslant 1/2$. The random Cayley sum graph G_p of $\mathbb{Z}_n$ satisfies*

$$\alpha(G_p) \geqslant (2-2\delta)\log_{\frac{1}{1-p}} n$$

with high probability as $n \to \infty$.

The proof of this result is significantly easier than the upper bound. In fact, using only the pseudorandom properties of G_p is enough to obtain $\alpha(G_p) \geqslant \big(1/2 + o(1)\big)(\log n)/p$ (see [2] and [4, Corollary 2.2]). To improve the leading constant to 2, we use both the randomness (as opposed to only the pseudorandomness) and the fact that we can restrict our attention to any sub-collection of potential independent sets in G_p.

More precisely, for each $k \in \mathbb{N}$, we define Z_k to be the random variable counting all independent k-sets in G_p with maximal doubling, that is

$$Z_k = \left|\left\{A \in \mathcal{Z}_k : A \hat{+} A \subseteq S^c\right\}\right|$$

where

$$\mathcal{Z}_k = \left\{A \subseteq \mathbb{Z}_n : |A| = k,\, |A \hat{+} A| = \binom{k}{2}\right\}.$$

If $Z_k > 0$, then $\alpha(G_p) \geqslant k$ regardless of the potential independent k-sets that Z_k overlooks.

In order to prove the lower bound, it is enough to show that $\mathrm{Var}(Z_k) = o(\mathbb{E}[Z_k]^2)$ since

$$\mathbb{P}(\alpha(G_p) \geqslant k) \geqslant \mathbb{P}(Z_k > 0) \geqslant 1 - \frac{\mathrm{Var}(Z_k)}{\mathbb{E}[Z_k]^2},$$

by Chebyshev's inequality. The first step is to estimate $\mathbb{E}[Z_k]$, which we do by showing that $\mathcal{Z}_k$ is large and using linearity of expectation.

Lemma 3.37. *For all $k \in \mathbb{N}$, we have*

$$|\mathcal{Z}_k| \geqslant \left(1 - \frac{k^4}{n}\right)\binom{n}{k}.$$

Proof. We will (equivalently) show that the corresponding fraction of $A \subseteq \mathbb{Z}_n$ with $|A| = k$ satisfy $|A \hat{+} A| = \binom{k}{2}$. Observe that if $|A \hat{+} A| < \binom{k}{2}$, then there are distinct $x_1, x_2, x_1', x_2' \in A$ such that

$$x_1 + x_2 = x_1' + x_2'.$$

Motivated by that observation, define

$$\mathcal{Q} = \big\{\{x_1, x_2, x_1', x_2'\} \subseteq \mathbb{Z}_n : x_1 + x_2 = x_1' + x_2', \text{ and } x_1, x_2, x_1', x_2' \text{ are distinct}\big\},$$

and let Y be a uniformly random k-set in $\mathbb{Z}_n$. Taking a union bound over $\mathcal{Q}$ yields

$$\mathbb{P}(Y \notin \mathcal{Z}_k) \leqslant \sum_{Q \in \mathcal{Q}} \mathbb{P}(Q \subseteq Y) \leqslant n^3 \left(\frac{k}{n}\right)^4 = \frac{k^4}{n},$$

where the second inequality is due to the choices of x_1, x_2 and x_1' determining $x_2' = x_1 + x_2 - x_1'$ for $\{x_1, x_2, x_1', x_2'\} \in \mathcal{Q}$. □

The second lemma that we need to prove Theorem 3.36 gives a bound on $\mathrm{Var}(Z_k)$.

Lemma 3.38. *For every $k \in \mathbb{N}$ and $p = p(n) \in (0,1)$, we have*

$$\mathrm{Var}(Z_k) \leqslant \mathbb{E}[Z_k] + \binom{n}{k} \sum_{s=1}^{k} k^{3s} \binom{n}{k-s} (1-p)^{2\binom{k}{2} - ks/2}.$$

The main step in the proof of Lemma 3.38 is to show, for every $A \in \mathcal{Z}_k$, that

$$\sum_{\substack{Y \in \mathcal{Z}_k \\ Y \sim A}} (1-p)^{|(A \hat{+} A) \cup (Y \hat{+} Y)|} \leqslant \sum_{s=1}^{k} k^{3s} \binom{n}{k-s} (1-p)^{2\binom{k}{2} - ks/2} \tag{3.107}$$

where $Y \sim A$ if

$$Y \neq A \qquad \text{and} \qquad (A \hat{+} A) \cap (Y \hat{+} Y) \neq \emptyset. \tag{3.108}$$

In order to do that, define

$$\mathbf{I}(A, Y) = \big\{Y' \subseteq Y : (A \hat{+} A) \cap (Y' \hat{+} Y') = \emptyset\big\},$$

and

$$\mathbf{I}^*(A, Y) = \big\{Y' \in \mathbf{I}(A, Y) : (y + Y') \cap (A \hat{+} A) \neq \emptyset \text{ for all } y \in Y \setminus Y'\big\}. \tag{3.109}$$

The definition of $\mathbf{I}^*(A, Y)$ is motivated by the following lemma, which gives an upper bound on $\big|(A \hat{+} A) \cap (Y \hat{+} Y)\big|$:

Lemma 3.39. *Let $k \in \mathbb{N}$ and $A, Y \in \mathcal{Z}_k$. If $Y \sim A$, then*

$$\left|(A \hat{+} A) \cap (Y \hat{+} Y)\right| \leqslant \frac{k(k-t)}{2},$$

where $t = \min\left\{|Y'| : Y' \in \mathbf{I}^*(A, Y)\right\}$.

To prove Lemma 3.39, we will need the following simple observation about graphs.

Observation 3.40. *Let G be a graph with k vertices. If all maximal independent sets of G have at least t vertices, then G has at most $k(k-t)/2$ edges.*

Proof. Take $v \in V(G)$ to be a vertex of maximum degree, and $I_v \subseteq V(G)$ to be a maximal independent set containing v. Then,

$$t \leqslant |I_v| \leqslant k - d(v) = k - \Delta(G).$$

Thus, $\Delta(G) \leqslant k - t$, and the claimed bound follows. □

To deduce Lemma 3.39, we will apply Observation 3.40 to $G_{A\hat{+}A}[Y]$, the Cayley sum graph over $A \hat{+} A$ restricted to the vertex set Y.

Proof of Lemma 3.39. We claim that

$$\left|(A \hat{+} A) \cap (Y \hat{+} Y)\right| = e(G_{A\hat{+}A}[Y]), \tag{3.110}$$

which not only implies that the collection $\mathbf{I}^*(A, Y)$ is exactly the collection of maximal independent sets in $G_{A\hat{+}A}[Y]$, but also reduces the proof to applying Observation 3.40 with t to $G_{A\hat{+}A}[Y]$. Therefore, we first establish (3.110), and then the characterization of $\mathbf{I}^*(A, Y)$.

To show (3.110), recall that, by the definition of the Cayley sum graph, $y_1y_2 \in \binom{Y}{2}$ is an edge of $G_{A\hat{+}A}[Y]$ if and only if $y_1 + y_2 \in A \hat{+} A$. We obtain equality by observing that each sum in $Y \hat{+} Y$ corresponds to exactly one pair $y_1y_2 \in \binom{Y}{2}$, since $|Y \hat{+} Y| = \binom{k}{2}$ by $Y \in \mathcal{Z}_k$.

It follows from (3.110) that $\mathbf{I}(A, Y)$ corresponds to independent sets in $G_{A\hat{+}A}[Y]$. Moreover, for every $Y^* \in \mathbf{I}^*(A, Y)$, we know that there is no $Y' \in \mathbf{I}(A, Y)$ such that $Y^* \subsetneq Y'$ by the condition in (3.109). Our definition of t then corresponds to all maximal independent sets in $G_{A\hat{+}A}[Y]$ having at least t vertices, so we can apply Observation 3.40 as desired. □

With Lemma 3.39 in hand, the proof of (3.107) now relies on an efficient count of $Y \in \mathcal{Z}_k$ with $Y \sim A$, for each fixed $A \in \mathcal{Z}_k$. Notice that $\mathbf{I}^*(A, Y)$ is not empty, as every graph has at least one maximal independent set, so we can fix a set $Y^* \in \mathbf{I}^*(A, Y)$ of minimum size. Our counting strategy is to consider the choices for elements in Y^* and then the choices for $Y \setminus Y^*$; in fact, a trivial count of all sets $Y^* \subseteq \mathbb{Z}_n$ suffices, so we only need to count the possible elements in $Y \setminus Y^*$ efficiently.

In order to bound the choices for $Y \setminus Y^*$, notice that for every $y \in Y \setminus Y^*$, we have $(y + Y^*) \cap (A \hat{+} A) \neq \emptyset$ by (3.109). It follows that there are $y^* \in Y^*$ and distinct $x_1, x_2 \in A$ such that $y + y^* = x_1 + x_2$, or, equivalently,

$$Y \setminus Y^* \subseteq A \hat{+} A - Y^*.$$

We can therefore choose the elements of $Y \setminus Y^*$ from a set of size at most

$$|A \hat{+} A - Y^*| \leqslant |A|^2 |Y^*| \leqslant k^3,$$

so there are at most k^{3s} choices for $Y \setminus Y^*$ if $s = |Y \setminus Y^*|$. Bounding the number of choices for Y^* by $\binom{n}{k-s}$ yields the bound in (3.107), except for the $(1-p)^{2\binom{k}{2}-ks/2}$ term, which we obtain by using Lemma 3.39.

We now have all the ingredients to prove Lemma 3.38.

Proof of Lemma 3.38. Observe that, via standard calculations, we have that

$$\operatorname{Var}(Z_k) \leqslant \mathbb{E}[Z_k] + \sum_{A \in \mathcal{Z}_k} \sum_{\substack{Y \in \mathcal{Z}_k \\ Y \sim A}} (1-p)^{|(A \hat{+} A) \cup (Y \hat{+} Y)|}, \tag{3.111}$$

where, recall, $Y \sim A$ was defined in (3.108). We therefore need to show that

$$\sum_{\substack{Y \in \mathcal{Z}_k \\ Y \sim A}} (1-p)^{|(A \hat{+} A) \cup (Y \hat{+} Y)|} \leqslant \sum_{s=1}^{k} k^{3s} \binom{n}{k-s} (1-p)^{2\binom{k}{2}-ks/2} \tag{3.112}$$

for each $A \in \mathcal{Z}_k$.

To prove (3.112), fix $A \in \mathcal{Z}_k$ and, for each $Y \in \mathcal{Z}_k$ with $Y \sim A$, choose a set $Y^* = Y^*(A, Y)$ of minimum size. If we group the sets Y by the size of their corresponding Y^*, we can count them by first enumerating the choices for Y^* and then the choices for $Y \setminus Y^*$. For fixed $s = |Y \setminus Y^*| = k - |Y^*|$, there are at most $\binom{n}{k-s}$ choices for Y^*, and there are at most k^{3s} choices for $Y \setminus Y^*$ since $Y \setminus Y^* \subseteq A \hat{+} A - Y^*$.

We then bound the size of the union in (3.112) by applying Lemma 3.39 to the pair (A, Y)

$$\big|(A \hat{+} A) \cap (Y \hat{+} Y)\big| \leqslant \frac{ks}{2}$$

which means that we can ignore the case $s = 0$ because it would contradict $Y \sim A$. A trivial inclusion-exclusion now yields the bound we need for the size of the union:

$$\big|(A \hat{+} A) \cup (Y \hat{+} Y)\big| = 2\binom{k}{2} - \big|(A \hat{+} A) \cap (Y \hat{+} Y)\big| \geqslant 2\binom{k}{2} - \frac{ks}{2}. \tag{3.113}$$

Replacing (3.113) and the above count of Y for fixed s into the left-hand side of (3.112) gives

$$\sum_{A \in \mathcal{Z}_k} \sum_{\substack{Y \in \mathcal{Z}_k \\ Y \sim A}} (1-p)^{|(A \hat{+} A) \cup (Y \hat{+} Y)|} \leqslant \sum_{A \in \mathcal{Z}_k} \sum_{s=1}^{k} k^{3s} \binom{n}{k-s} (1-p)^{2\binom{k}{2}-ks/2}.$$

Trivially bounding the number of choices for $A \in \mathcal{Z}_k$ by $\binom{n}{k}$ and plugging the result into (3.111) completes the proof. □

With Lemma 3.37 and Lemma 3.38, the proof of Theorem 3.36 is just checking that the bounds match those of the statement.

Proof of Theorem 3.36. Fix $p = p(n)$ satisfying $1/2 \geqslant p \geqslant n^{-\delta/8}$ for the given $\delta > 0$, and let $k = (2-2\delta)\log_{\frac{1}{1-p}} n$. It suffices to show that

$$\frac{\mathrm{Var}(Z_k)}{\mathbb{E}[Z_k]^2} \to 0 \qquad \text{as} \qquad n \to \infty. \tag{3.114}$$

First, we compute the expected value of Z_k using Lemma 3.37, the fact that $k = o(n^{1/4})$ by our choice of p, and linearity of expectation:

$$\mathbb{E}[Z_k] = \big(1 - o(1)\big)\binom{n}{k}(1-p)^{\binom{k}{2}} \geqslant \frac{1}{2}\binom{n}{k} n^{-(1-\delta)(k-1)} \to \infty \tag{3.115}$$

as $n \to \infty$, by our choice of k. Now, if we assume that

$$\mathbb{E}[Z_k] \geqslant \binom{n}{k} \sum_{s=1}^{k} k^{3s} \binom{n}{k-s} (1-p)^{2\binom{k}{2} - ks/2}, \tag{3.116}$$

then, by Lemma 3.38, we have $\mathrm{Var}(Z_k) \leqslant 2\mathbb{E}[Z_k]$ and

$$\frac{\mathrm{Var}(Z_k)}{\mathbb{E}[Z_k]^2} \leqslant \frac{2}{\mathbb{E}[Z_k]} \to 0 \qquad \text{as} \qquad n \to \infty,$$

by (3.115). We therefore assume that the converse of (3.116) holds.

Before we proceed, observe that applying the standard binomial inequality for $k \leqslant n/2$

$$\binom{n}{k-s} \leqslant \left(\frac{2k}{n}\right)^s \binom{n}{k}$$

to the right-hand side of (3.116) yields

$$\begin{aligned}\binom{n}{k} \sum_{s=1}^{k} k^{3s} \binom{n}{k-s} (1-p)^{2\binom{k}{2} - ks/2} &\leqslant \binom{n}{k}^2 (1-p)^{2\binom{k}{2}} \sum_{s=1}^{k} k^{3s} \left(\frac{2k}{n}\right)^s (1-p)^{-ks/2} \\ &\leqslant 4\mathbb{E}[Z_k]^2 \sum_{s=1}^{k} \left(\frac{2k^4}{n^\delta}\right)^s, \end{aligned} \tag{3.117}$$

where in the last inequality we used (3.115) with n sufficiently large and also

$$(1-p)^{-ks/2} = n^{(1-\delta)s}$$

because $k = (2-2\delta)\log_{\frac{1}{1-p}} n$.

By Lemma 3.38, our assumption that the converse of (3.116) holds, and (3.117), we have

$$\mathrm{Var}(Z_k) \leqslant 8\mathbb{E}[Z_k]^2 \sum_{s=1}^{k} \left(\frac{2k^4}{n^\delta}\right)^s. \tag{3.118}$$

Replacing (3.118) into (3.114), we conclude that the proof is complete if we show that

$$\sum_{s=1}^{k} \left(\frac{2k^4}{n^\delta} \right)^s \to 0 \qquad \text{as} \qquad n \to \infty.$$

This is easily seen to be true when $k = o(n^{\delta/4})$, which holds for our choice of $p \geqslant n^{-\delta/8}$. □

3.10 Follow-up work and open problems

Since we published [29], our methods were optimized by Nenadov [98] to obtain (3.1) for all

$$p \geqslant (\log n)^{-1/3} (\log \log n)^{O(1)}.$$

While he optimized the version of Freĭman's theorem that we use in Section 3.7, Proposition 3.20, the main novelty in his work is an improved version of Theorem 3.1 that not only removes the extra $\log |A|$, but also has a dependency on $\dim_{\mathrm{F}}(A)$ instead of $\sigma[A]$ in the size of the fingerprint. To obtain that improvement, he pioneered a technique which has applications in other problems where the asymmetric container theorem is useful [99, 101].

However, already when $p \leqslant (\log n)^{-1}$, there is an obstacle that prevents his approach or any other similar to ours from working. To understand the barrier, consider the following approximate summary of our strategy. We find a family $\mathcal{S} = \{F(A) + F(A) : A \in \binom{\mathbb{Z}_n}{k}\}$ of subsets of $\mathbb{Z}_n$ of size s (for some $s \in \mathbb{N}$) with the following two properties:

1. For each set $A \in \binom{\mathbb{Z}_n}{k}$, there exists $S \in \mathcal{S}$ with $S \subseteq A + A$.
2. The family is small, that is, $|\mathcal{S}| \leqslant (1-p)^{-s}$.

Observe that each $S \in \mathcal{S}$ is of the form $F(A) + F(A)$ and has size s, so we must trivially have $|F(A)| \geqslant \sqrt{s}$ for all $A \in \binom{\mathbb{Z}_n}{k}$. Naively counting every set $F \subseteq \mathbb{Z}_n$ with $|F| = \sqrt{s}$ when bounding the size of $\mathcal{S}$, we obtain

$$|\mathcal{S}| \geqslant \binom{n}{\sqrt{s}} \approx \exp(\sqrt{s} \log n). \tag{3.119}$$

In our proof, we show that we can choose F inside a small generalised arithmetic progression P, thus replacing the $\log n$ term in (3.119) with $\log |P|$. However, even if we could find such a set P with $|P| = O(\sqrt{s})$, we could not improve the bound in (3.119) beyond $\exp(\sqrt{s})$.

Combining this lower bound on the size of $\mathcal{S}$ with the upper bound that we require in property (2) gives

$$\exp(\sqrt{s}) \leqslant |\mathcal{S}| \leqslant (1-p)^{-s},$$

which implies that $s \geqslant p^{-2}$. Now, consider any $A \in \binom{\mathbb{Z}_n}{k}$ with $\sigma[A] = O(1)$: the corresponding set $S \in \mathcal{S}$ satisfies $S \subseteq A + A$ and therefore

$$p^{-2} \leqslant s = |S| \leqslant |A + A| \leqslant O(k).$$

As k is the upper bound we are trying to prove for $\alpha(G_p)$, it follows that the best we can hope

for this approach is, for some constant $C > 2$,

$$\alpha(G_p) \leqslant C \max\{p^{-2}, p^{-1} \log n\},$$

where the second term in the maximum is the lower bound that we proved in Section 3.9.

This reflected the state-of-the-art (up to polylogarithmic terms) for a famous conjecture due to Alon [5] that, for each group Γ of order n and $t \in [n]$, the typical independence number of G_S with $|S| = t$ is at most $\frac{n}{t}(\log n)^{O(1)}$. Alon [2, 5] originally established that, in this setting,

$$\alpha(G_S) = O\big(\min\{p^{-2}(\log n)^2, (n \log n/p)^{1/2}\}\big) \tag{3.120}$$

where $p = t/n$. The first term in (3.120) is obtained using a fingerprint lemma whose proof is essentially the same as that of our Theorem 3.1. On the other hand, the second term is obtained via an application of the expander mixing lemma [3] and a computation of the eigenvalues of the adjacency matrix of G_S using Fourier analysis.

Conlon, Fox, Pham and Yepremyan [42] obtained a lower order improvement for $\alpha(G_S)$ in general groups Γ by considering the more general setting of "random entangled graphs," and their method is also used to prove other conjectures of Alon about Cayley graphs. In the regime of lower doubling, they generalise the combinatorial fingerprints of Alon and ourselves (see [42, Theorem 1.1 and Lemma 2.7]), so the same barrier appears at p^{-2}.

In [29], we connected this barrier with a conjecture from Lovett, which was very recently solved by Alon and Pham [7] using Fourier analysis. They used that result to improve (3.120) to

$$\alpha(G_S) \leqslant p^{-3/2}(\log n)^{O(1)}$$

with high probability when Γ is Abelian. Their approach is similar to the approach that we borrow from Green, partitioning the possible independent sets A according to the index of a certain "popularity level" within the representations of $A + A$ as sums of elements in A, and taking a union bound over fingerprints (or covers in their terminology). Alon and Pham have also announced that, combining the techniques in [7] with our methods and several new ideas, they can further extend Theorem 1.3 to all

$$p \geqslant (\log n)^{-2}(\log\log n)^{O(1)}.$$

Another interesting open question is to determine the minimum number of translates necessary to obtain the correct leading constant of $d + 1$ in Freĭman's lemma. This is closely related to the question asked (and partially settled) by Bollobás, Leader and Tiba [22] on whether three translates suffice to obtain the Cauchy–Davenport lower bound for two sets $A, B \subseteq \mathbb{Z}_n$. Even though their question was recently resolved by Fox, Luo, Pham and Zhou [56] in the affirmative, it is not clear what is the truth in higher dimensions, even in the simpler case of $A = B$. We conjectured the following in [29]:

Conjecture 3.41 ([29, Conjecture 10.1])**.** *There exists $C > 0$ such that the following holds. For all $d \in \mathbb{N}$ and all finite sets $A \subseteq \mathbb{R}^d$ of full rank, A contains a subset T such that $|T| \leqslant Cd$ and*

$$|A + T| \geqslant (d+1)|A| - d^C.$$

The best negative result we have for this problem is given by the following simple construction. Let $d \geqslant 3$, and let $\{e_1, e_2, \ldots, e_d\}$ be the canonical basis of $\mathbb{R}^d$. Consider

$$A = P + \{0, e_1, \ldots, e_{d-1}\} \qquad \text{where} \qquad P = \{0, e_d, 2e_d, \ldots, ke_d\}$$

for some $k \in \mathbb{N}$. It is not hard to show that the following holds for this example: for every $\gamma > 0$, there is $c = c(\gamma) > 0$ such that if $T \subseteq A$ with $|T| < (2-\gamma)d$, then

$$|A + T| \leqslant (1-c)(d+1)|A|.$$

Chapter 4

Arithmetic progressions in sumsets: subsets of sparse random sets

In this chapter, we present the proof of Theorem 1.6. Recall that we denote by $[n]_p$ a p-random subset of $[n]$, that is, each element of $[n]$ is in the subset independently and with probability equal to p, and by $L(S)$ the length of the longest arithmetic progression in the set S.

Theorem 1.6. *There exist constants $c > 0$ and $C > 0$ such that, for any $0 < \beta \leqslant \alpha \leqslant 1$, $k \in \mathbb{N}$ and $p = p(n)$ satisfying*

$$k \leqslant \exp\left(c(\alpha\beta \log n)^{1/2} - \log\log n\right) \tag{4.1}$$

and

$$p \geqslant C\,\frac{k}{\beta}\left(\frac{(\log n)^3}{n}\right)^{1/2},$$

the following holds with high probability. Every pair of subsets $A, B \subseteq S$ with

$$|A| \geqslant \alpha|S| \qquad \text{and} \qquad |B| \geqslant \beta|S|,$$

where $S \sim [n]_p$, satisfies

$$L(A + B) \geqslant k.$$

In the proof of Theorem 1.6, we employ the container theorem for sumsets [26], a consequence of the asymmetric container lemma of Morris, Samotij and Saxton [96]. To build intuition for readers familiar with [75], it is instructive to highlight the similarities in both approaches, working in the case $A = B$ for simplicity. Hamel and Łaba rely on a "dense model theorem," which roughly says that the indicator of $A \subseteq [n]_p$ can be decomposed into a function which is dense in $\mathbb{Z}_n$, and a functional error term that is small in a suitable sense (see [68] for an introduction to such decompositions).

The container theorem for sumsets, on the other hand, ensures that each A is associated to a pair of sets (X, Y) in a small family, where moreover $A \subseteq X$ and $Y \subseteq A + A$. More importantly, when X is large, it can be seen as a "dense model" for A, because the container theorem ensures that Y contains most of the sums of the form $x_1 + x_2$ for $x_1, x_2 \in X$. This (purely combinatorial) property, together with the fact that $Y \subseteq A + A$, allows us to deduce facts about A based on the analysis of Y. Concretely in this case, we need to show that $X + X$

not only contains long APs when X is large, but does so "robustly," that is, that also every Y that is "close" to $X + X$ in the above sense contains a long AP.

To obtain such a result, we slightly modify the alternative proof of Green's theorem given by Croot, Łaba and Sisask [45], using almost periodicity of convolutions. This modification to their argument leads to a "robust" version of Green's theorem (see Theorem 4.3), and is the only part of our proof that requires any form of Fourier analysis. From this robust theorem, we deduce an intermediate result (Theorem 4.1) from which Theorem 1.6 easily follows, as we will show in the next section.

4.1 Containers for sets with sumsets free of long progressions

We can informally describe Theorem 4.1, the result which will imply Theorem 1.6, as stating that there is a "small" family $\mathcal{C}$ of triples $(X^{(1)}, X^{(2)}, Y)$ of sets such that (i) for all large $A, B \subseteq [n]$, there is $(X^{(1)}, X^{(2)}, Y) \in \mathcal{C}$ such that $A \subseteq X^{(1)}$, $B \subseteq X^{(2)}$ and $Y \subseteq A + B$, and (ii) the triples $(X^{(1)}, X^{(2)}, Y)$ are such that either at least one of $X^{(1)}$ and $X^{(2)}$ is "small", or Y contains "long" APs.

Theorem 4.1. *There exist constants $C', c > 0$ such that the following holds for all sufficiently large $n \in \mathbb{N}$. Let $0 < \beta \leqslant \alpha \leqslant 1$, and let $k \in \mathbb{N}$ satisfy*

$$k \leqslant \exp(c(\alpha\beta \log n)^{1/2} - \log\log n). \tag{4.2}$$

There exists a family $\mathcal{C} \subseteq 2^{[n]} \times 2^{[n]} \times 2^{[2n]}$ of size

$$|\mathcal{C}| \leqslant \exp\left(C' k \sqrt{n(\log n)^3}\right) \tag{4.3}$$

such that:

(i) For every $A, B \subseteq [n]$ with

$$\min\left\{|A|, |B|\right\} \geqslant C'\sqrt{n/\log n},$$

there exists $(X^{(1)}, X^{(2)}, Y) \in \mathcal{C}$ such that

$$A \subseteq X^{(1)}, \qquad B \subseteq X^{(2)} \qquad \text{and} \qquad Y \subseteq A + B. \tag{4.4}$$

(ii) For every $(X^{(1)}, X^{(2)}, Y) \in \mathcal{C}$, one of the following hold:

$$|X^{(1)}| \leqslant \frac{\alpha n}{8}, \qquad |X^{(2)}| \leqslant \frac{\beta n}{8}, \qquad \text{or} \qquad L(Y) \geqslant k.$$

We now give a proof of Theorem 1.6 under the assumption that Theorem 4.1 holds.

Proof of Theorem 1.6 assuming Theorem 4.1. Fix α, β, k and p as in the statement of Theorem 1.6. We want to show that, for $S \sim [n]_p$,

$$\mathbb{P}\left(\exists A, B \subseteq S : |A| \geqslant \alpha|S|,\ |B| \geqslant \beta|S|,\ L(A+B) < k\right) \to 0, \tag{4.5}$$

as $n \to \infty$. First, note that we can bound (4.5) by

$$\mathbb{P}\left(\exists A, B \subseteq [n]_p : |A| \geqslant \frac{\alpha np}{2},\ |B| \geqslant \frac{\beta np}{2},\ L(A+B) < k\right) + \mathbb{P}\Big(\big|[n]_p\big| \leqslant np/2\Big) \tag{4.6}$$

and applying Chernoff's inequality (see, e.g., [79, Corollary 2.3.4]) gives

$$\mathbb{P}\Big(\big|[n]_p\big| \leqslant np/2\Big) \leqslant \exp(-np/8). \tag{4.7}$$

We now turn our attention to the first term in (4.6).

Since our assumption (4.1) is identical to (4.2), we may apply Theorem 4.1 with α, β and k to obtain a family $\mathcal{C}$. In particular, it follows from Theorem 4.1(i) and (with room to spare)

$$\frac{\alpha np}{2} \geqslant \frac{\beta np}{2} \geqslant \frac{Ck}{2}\big(n(\log n)^3\big)^{1/2} \geqslant C'\left(\frac{n}{\log n}\right)^{1/2},$$

where the second to last inequality holds because

$$p \geqslant C\frac{k}{\beta}\left(\frac{(\log n)^3}{n}\right)^{1/2}.$$

We conclude that for every $A, B \subseteq [n]_p \subseteq [n]$ with $|A| \geqslant \alpha np/2$ and $|B| \geqslant \beta np/2$, the family $\mathcal{C}$ deterministically contains a triple $(X^{(1)}, X^{(2)}, Y)$ such that (4.4) holds, that is,

$$A \subseteq X^{(1)}, \qquad B \subseteq X^{(2)} \qquad \text{and} \qquad Y \subseteq A + B.$$

Writing $X_p^{(i)} = X^{(i)} \cap [n]_p$ for each $i \in \{1, 2\}$, and taking a union bound over the family $\mathcal{C}$, we obtain that the first term in (4.6) is at most

$$\sum_{(X^{(1)}, X^{(2)}, Y) \in \mathcal{C}} \mathbb{P}\left(\exists A \subseteq X_p^{(1)},\, B \subseteq X_p^{(2)} : |A| \geqslant \frac{\alpha np}{2},\, |B| \geqslant \frac{\beta np}{2},\, L(Y) < k\right),$$

where we have used that $L(Y) \leqslant L(A+B)$ follows trivially from $Y \subseteq A + B$, or

$$\sum_{(X^{(1)}, X^{(2)}, Y) \in \mathcal{C}} \mathbb{P}\left(|X_p^{(1)}| \geqslant \frac{\alpha np}{2},\, |X_p^{(2)}| \geqslant \frac{\beta np}{2},\, L(Y) < k\right) \tag{4.8}$$

since we must trivially have $|X_p^{(1)}| \geqslant |A|$ and $|X_p^{(2)}| \geqslant |B|$ when $A \subseteq X_p^{(1)}$ and $B \subseteq X_p^{(2)}$.

We now use Theorem 4.1(ii) to conclude that as the triples $(X^{(1)}, X^{(2)}, Y)$ in (4.8) must satisfy $L(Y) < k$, then either

$$|X^{(1)}| \leqslant \alpha n/8 \qquad \text{or} \qquad |X^{(2)}| \leqslant \beta n/8. \tag{4.9}$$

At this point, it follows from another application of Chernoff's inequality and (4.9) that

$$\mathbb{P}\Big(|X_p^{(1)}| \geqslant \alpha pn/2,\ |X_p^{(2)}| \geqslant \beta pn/2\Big) \leqslant \exp(-\beta np/12) \tag{4.10}$$

for all triples $(X^{(1)}, X^{(2)}, Y)$ satisfying (4.9), because $\beta \leqslant \alpha$ by assumption. Replacing (4.10) in (4.8), we obtain

$$\sum_{(X^{(1)},X^{(2)},Y)\in\mathcal{C}} \exp(-\beta np/12) \leqslant \exp(C'k\sqrt{n(\log n)^3} - \beta np/12) \tag{4.11}$$

where we used (4.3) to bound $|\mathcal{C}|$. Finally, using our lower bound for p,

$$p \geqslant C\frac{k}{\beta}\left(\frac{(\log n)^3}{n}\right)^{1/2},$$

and choosing the constant C to satisfy $C \geqslant 24C'$ is sufficient to make the right-hand side of (4.11) go to 0 as $n \to \infty$. This means that its sum with (4.7) also goes to 0, and that the proof is complete. □

4.2 Containers for sumsets with distinct summands

In this section, we prove Theorem 4.1 from Theorem 4.2, a slight generalization of the container theorem for sumsets of Campos [26, Theorem 4.2] due to Campos, Coulson, Serra and Wötzel [27]. Their statement is more general than Theorem 4.2 below, but we specialize it to $[n] \subseteq \mathbb{Z}$, taking $m = 2n$ in the original statement.

Theorem 4.2. *Let $\tau > 0$, and suppose that n, s_1 and s_2 are integers satisfying $s_2 \leqslant s_1$ and $\log n \leqslant s_1 \leqslant 2n \leqslant s_2^2 \log n$. There exists a family $\mathcal{C}_{s_1,s_2} \subseteq 2^{[n]} \times 2^{[n]} \times 2^{[2n]}$ of size*

$$|\mathcal{C}_{s_1,s_2}| \leqslant \exp\left(2^{20}\tau^{-2}(2n(\log n)^3)^{1/2}\right) \tag{4.12}$$

such that the following hold:

(a) For all $A, B \subseteq [n]$ with $|A| = s_1$ and $|B| = s_2$, there exists $(X^{(1)}, X^{(2)}, Y) \in \mathcal{C}_{s_1,s_2}$ such that

$$A \subseteq X^{(1)}, \qquad B \subseteq X^{(2)}, \qquad \text{and} \qquad Y \subseteq A + B.$$

(b) For all $(X^{(1)}, X^{(2)}, Y) \in \mathcal{C}_{s_1,s_2}$, either $\max\{|X^{(1)}|, |X^{(2)}|\} \leqslant 2n/\log n$, or

$$\big|\{(x_1, x_2) \in X^{(1)} \times X^{(2)} : x_1 + x_2 \notin Y\}\big| \leqslant \tau^2|X^{(1)}| \cdot |X^{(2)}|. \tag{4.13}$$

We will prove Theorem 4.1 from Theorem 4.2 by applying the latter for all pairs of integers $|A| = s_1$ and $|B| = s_2$, taking the final $\mathcal{C}$ to be the union of all the $\mathcal{C}_{s_1,s_2}$. This will immediately

imply that Theorem 4.1(i) holds (from Theorem 4.2(a)), but it is not clear how to deduce Theorem 4.1(ii) from Theorem 4.2(b). The main fact that we need is that, roughly speaking, the complicated looking condition (4.13) implies that $L(Y) \geqslant k$ if $\tau^2 = 1/2k$.

4.2.1 A robust version of Green's theorem

We show that (4.13) with $\tau^2 = 1/(2k)$ implies that $L(Y) \geqslant k$ by strengthening the statement of Green's theorem to show that for all $X^{(1)}, X^{(2)} \subseteq [n]$ with $|X^{(1)}| \gtrsim \alpha n$ and $|X^{(2)}| \gtrsim \beta n$, and all $Y \subseteq X^{(1)} + X^{(2)}$ satisfying (4.13), the set Y contains a k-AP.

Theorem 4.3. *There is a constant $c' > 0$ such that the following holds. Let $S, T \subseteq [n]$ satisfy*

$$|S| \geqslant \delta n \qquad \text{and} \qquad |T| \geqslant \gamma n$$

for some $0 < \gamma \leqslant \delta \leqslant 1$, and let $k \in \mathbb{N}$ be such that

$$k \leqslant \exp\left(c'(\delta\gamma \log n)^{1/2} - \log\log n\right).$$

If $Y \subseteq S + T$ and

$$|\{(x_1, x_2) \in S \times T : x_1 + x_2 \notin Y\}| \leqslant \frac{|S| \cdot |T|}{2k}, \tag{4.14}$$

then $L(Y) \geqslant k$.

To prove Theorem 4.3, we will follow closely the proof of Green's theorem given by Croot, Łaba and Sisask [45]. In particular, we will use the following consequence of their results, which can be obtained by applying [45, Lemma 2.2] to the Bohr set obtained by [45, Theorem 1.2]. Recall that the L^s-norm of a function f defined over the elements of a group Γ is

$$\|f\|_{L^s} = \left(\frac{1}{|\Gamma|}\sum_{x\in\Gamma} |f(x)|^s\right)^{1/s},$$

and that the convolution between two functions $f, g : \Gamma \to \mathbb{C}$ is defined as

$$f * g(x) = \frac{1}{|\Gamma|}\sum_{y\in\Gamma} f(y)g(x-y).$$

Proposition 4.4. *There exist absolute constants $C, c > 0$ for which the following holds. Let $0 < \delta, \gamma < 1$ and let $S, T \subseteq \mathbb{Z}_n$ be such that*

$$|S| \geqslant \delta n, \qquad \text{and} \qquad |T| \geqslant \gamma n.$$

For every $s \in \mathbb{N}$ and $0 < \varepsilon < 1$, there exists an arithmetic progression $P \subseteq \mathbb{Z}_n$ such that

$$|P| \geqslant \frac{c\varepsilon}{2\pi} n^{1/d},$$

where $d = Cs/\varepsilon^2$, and, for each $w \in P$,

$$\|1_S * 1_T(x + w) - 1_S * 1_T(x)\|_{L^s(x)} \leqslant \varepsilon(\delta\gamma)^{1/2} n. \tag{4.15}$$

The only modification to the proof of Croot, Łaba and Sisask that we make to prove Theorem 4.3 is that we approximate the ℓ_1-norm of the function $(1_S * 1_T) \cdot 1_Y$ by that of $1_S * 1_T$ using the triangle inequality. This results in a small extra term that we can bound by the missing pairs condition (4.14).

Proof of Theorem 4.3. In order to use Fourier analysis, we first embed $[n]$ into $\mathbb{Z}_q$ using a (trivial) Freĭman isomorphism ϕ, where $4n \leqslant q \leqslant 8n$ is a prime. Showing that $L(\phi(Y)) \geqslant k$ and then taking the k-AP $P' \subseteq \phi(Y)$, we can use that ϕ is a Freĭman isomorphism to deduce that $\phi^{-1}(P') = P$ is also a k-AP, and therefore $L(Y) \geqslant k$. The only caveat is that $|\phi(S)|/q$ and $|\phi(T)|/q$ differ by a factor between 4 and 8 from δ and γ, so the k-AP in $\phi(Y)$ is shorter than what is required by the statement of Theorem 4.3. To overcome that small issue, we adjust the value of c' in the statement by a small constant, so we can assume instead that n is a prime, and that $S, T \subseteq \mathbb{Z}_n$ rather than $S, T \subseteq [n]$ in the statement of Theorem 4.3.

Let $S, T \subseteq \mathbb{Z}_n$ and $k \in \mathbb{N}$ be as in the statement of Theorem 4.3, and assume, without loss of generality, that $|S| = \delta n$ and $|T| = \gamma n$. Choosing

$$s = C' \sqrt{\delta\gamma \log n} \qquad \text{and} \qquad \varepsilon = \frac{(\delta\gamma)^{1/2}}{(2e)} \tag{4.16}$$

suffices to ensure that P in the statement of Proposition 4.4 can have

$$|P| \geqslant \frac{c\varepsilon}{2\pi} n^{1/d} \geqslant \exp\left(c'(\delta\gamma \log n)^{1/2} - \log\log n\right) \geqslant k$$

if we choose the constant $C' < (4e^2 C c')^{-1}$, so let P be a k-AP satisfying (4.15). Now take

$$f(x) = (1_S * 1_T)(x) \cdot 1_Y(x),$$

and observe that it suffices to establish that there exists $x \in \mathbb{Z}_n$ such that

$$f(x + w) > 0, \tag{4.17}$$

for all $w \in P$ to conclude that $x + P \subseteq Y$, where clearly $x + P$ is also an arithmetic progression of length k.

Towards that goal, we will obtain a good upper bound for

$$\sum_{x \in \mathbb{Z}_n} \sup_{w \in P} |f(x + w) - 1_S * 1_T(x)|. \tag{4.18}$$

Our first step is to fix $x \in \mathbb{Z}_n$ and $w \in P$, let $u = x + w$, and apply the triangle inequality to conclude that

$$|f(u) - 1_S * 1_T(x)| \leqslant \big(|f(u) - 1_S * 1_T(u)| + |1_S * 1_T(u) - 1_S * 1_T(x)|\big) \tag{4.19}$$

for every fixed x and w. Replacing (4.19) in (4.18), we obtain

$$\sum_{x\in\mathbb{Z}_n} \sup_{w\in P} |f(x+w) - 1_S * 1_T(x)| \leqslant \sum_{u\in\mathbb{Z}_n} \big(|P|\cdot|f(u) - 1_S * 1_T(u)| + \sup_{w\in P} |1_S * 1_T(u) - 1_S * 1_T(u-w)|\big) \tag{4.20}$$

where we also bounded the sup by the sum, i.e.

$$\sum_{x\in\mathbb{Z}_n} \sup_{w\in P} |f(x+w) - 1_S * 1_T(x+w)| \leqslant \sum_{u\in\mathbb{Z}_n} |P|\cdot|f(u) - 1_S * 1_T(u)|.$$

The next step is obtaining upper bounds for the two terms in the right-hand side of (4.20). We handle the first term with (4.14), our bound on

$$\sum_{u\in\mathbb{Z}_n} |f(u) - 1_S * 1_T(u)| = |\{(x_1,x_2)\in S\times T : x_1+x_2\notin Y\}|,$$

concluding that

$$\sum_{u\in\mathbb{Z}_n} \big(|P|\cdot|f(u) - 1_S * 1_T(u)|\big) \leqslant |P|\cdot\frac{\delta\gamma n^2}{2k} \leqslant \frac{\delta\gamma n^2}{2}, \tag{4.21}$$

where we replaced $|S| = \delta n$, $|T| = \gamma n$ and $|P| = k$.

To bound the second term, we repeat the calculations in the proof of Croot, Łaba and Sisask to show that

$$\sum_{x\in\mathbb{Z}_n} \sup_{w\in P} |1_S * 1_T(x+w) - 1_S * 1_T(x)| \leqslant \frac{\delta\gamma n^2\,|P|^{1/s}}{2e}. \tag{4.22}$$

To make the computations easier to follow, define

$$g(x,w) = |1_S * 1_T(x+w) - 1_S * 1_T(x)| \qquad \text{and} \qquad g^*(x) = \sup_{w\in P} g(x,w).$$

Applying Hölder's inequality with the pair $\ell = (1-1/s)^{-1}$ and s, we conclude that

$$\sum_{x\in\mathbb{Z}_n} g^*(x) \leqslant n^{1-1/s}\left(\sum_{x\in\mathbb{Z}_n} |g^*(x)|^s\right)^{1/s}. \tag{4.23}$$

As $g^*(x)\geqslant g(x,w)\geqslant 0$, we can incur a $|P|^{1/s}$ term and swap the sup and the sum, that is,

$$\left(\sum_{x\in\mathbb{Z}_n} |g^*(x)|^s\right)^{1/s} \leqslant \left(\sum_{x\in\mathbb{Z}_n}\sum_{w\in P} g(x,w)^s\right)^{1/s} \leqslant |P|^{1/s}\sup_{w\in P}\left(\sum_{x\in\mathbb{Z}_n} g(x,w)^s\right)^{1/s}. \tag{4.24}$$

Replacing (4.24) in the right-hand side of (4.23), we have shown that the latter is at most

$$n^{1-1/s}\left(\sum_{x\in\mathbb{Z}_n} |g^*(x)|^s\right)^{1/s} \leqslant n\,|P|^{1/s}\sup_{w\in P}\|g(x,w)\|_{L^s(x)}. \tag{4.25}$$

Now, it follows from the definition of g that

$$\sup_{w\in P} \|g(x,w)\|_{L^s(x)} \leqslant \varepsilon(\delta\gamma)^{1/2}n, \tag{4.26}$$

where we used that P satisfies (4.15). Combining (4.25) and (4.26), we obtain

$$\sum_{x\in\mathbb{Z}_n} \sup_{w\in P} |1_S * 1_T(x+w) - 1_S * 1_T(x)| \leqslant \varepsilon(\delta\gamma)^{1/2}\, n^2\, |P|^{1/s},$$

which, by our choice of ε in (4.16), then yields (4.22):

$$\sum_{x\in\mathbb{Z}_n} \sup_{w\in P} |1_S * 1_T(x+w) - 1_S * 1_T(x)| \leqslant \frac{\delta\gamma n^2\, |P|^{1/s}}{2e}.$$

Substituting (4.21) and (4.22), we have that the right-hand side of (4.20) is at most

$$\delta\gamma n^2 \left(\frac{1}{2} + \frac{|P|^{1/s}}{2e}\right).$$

Observing that our choice of s satisfies $k = |P| < e^s$ by (4.16), we conclude that

$$\sum_{x\in\mathbb{Z}_n} \sup_{w\in P} |f(x+w) - 1_S * 1_T(x)| < \delta\gamma n^2 = \sum_{x\in\mathbb{Z}_n} 1_S * 1_T(x). \tag{4.27}$$

Since (4.27) can hold only if there is $x \in \mathbb{Z}_n$ such that, for all $w \in P$,

$$|f(x+w) - 1_S * 1_T(x)| < 1_S * 1_T(x)$$

we conclude that for this particular choice of x, $f(x+w) \neq 0$. But f is a non-negative function, and therefore

$$f(x+w) > 0,$$

which, by the reasoning in (4.17), completes the proof. □

4.2.2 Proof of Theorem 4.1

We will now combine Theorem 4.2 with Theorem 4.3 to prove Theorem 4.1. Together with the contents of Section 4.1, this will complete the proof of Theorem 1.6.

Proof of Theorem 4.1. Fix α, β and k as in the statement of Theorem 4.1 and let $\mathcal{C}_{s_1,s_2}$ be the family obtained by applying Theorem 4.2 with $\tau = (2k)^{-1/2}$ for some pair of integers s_1 and s_2. We define the family $\mathcal{C}$ as

$$\mathcal{C} = \bigcup_{s_1=s'}^{n} \bigcup_{s_2=s'}^{s_1} \mathcal{C}_{s_1,s_2} \tag{4.28}$$

where $s' = C\sqrt{n/\log n} \geqslant \log n$ is the lower bound for the sizes of the sets in Theorem 4.1(i).

Observe that (4.28) is valid, because

$$\log n \leqslant s' \leqslant s_2 \leqslant s_1 \leqslant 2n \leqslant (C\sqrt{n/\log n})^2 \log n = (s')^2 \log n \leqslant s_2^2 \log n,$$

so s_1 and s_2 satisfy the assumptions of Theorem 4.2. Next, we will prove that $\mathcal{C}$ satisfies the requirements to be our final container family. First, observe that $\mathcal{C}$ satisfies (4.3) because

$$|\mathcal{C}| \leqslant \sum_{s_1=s'}^{n} \sum_{s_2=s'}^{s_1} |\mathcal{C}_{s_1,s_2}| \leqslant n^2 \exp\left(2^{21} k (2n(\log n)^3)^{1/2}\right) \leqslant \exp\left(C' k \sqrt{n(\log n)^3}\right)$$

by (4.12), and our choices of $\tau = (2k)^{-1/2}$ and C' as a sufficiently large constant.

It remains to show that items (i) and (ii) in Theorem 4.1 hold for this definition of $\mathcal{C}$. For the former, first fix a pair of sets $A, B \subseteq [n]$ such that $\min\{|A|, |B|\} \geqslant C'\sqrt{n/\log n}$. By swapping A and B if necessary, we may assume that $|A| \geqslant |B|$, and then let $s_1 = |A|$ and $s_2 = |B|$; it is straightforward that $s' \leqslant s_2 \leqslant s_1 \leqslant n$. We therefore conclude by (4.28) that $\mathcal{C}_{s_1,s_2} \subseteq \mathcal{C}$, and item (i) follows from Theorem 4.2(a).

Now, fix a triple $(X^{(1)}, X^{(2)}, Y) \in \mathcal{C}$ with the goal of showing that it satisfies Theorem 4.1(ii). By (4.28) and Theorem 4.2(b), we have that either

$$\max\{|X^{(1)}|, |X^{(2)}|\} \leqslant 2n/\log n, \tag{4.29}$$

or

$$|\{(x_1, x_2) \in X^{(1)} \times X^{(2)} : x_1 + x_2 \notin Y\}| \leqslant \frac{|X^{(1)}| \cdot |X^{(2)}|}{2k}. \tag{4.30}$$

We claim that if the triple satisfies (4.29), then we are done because

$$2n/\log n \leqslant \alpha n/8. \tag{4.31}$$

Indeed, (4.31) holds because otherwise $k < 1$: if $\alpha \leqslant 16(\log n)^{-1}$, then

$$k \leqslant \exp(c(\alpha\beta \log n)^{1/2} - \log\log n) \leqslant \exp(4c - \log\log n) < 1$$

as long as n is greater than the constant $\exp(\exp(4c))$.

As we have handled the case $\max\{|X^{(1)}|, |X^{(2)}|\} \leqslant 2n/\log n$, we can assume that (4.30) holds. Further let $\delta = |X^{(1)}|/n$, and $\gamma = |X^{(2)}|/n$, and observe that if we have $\delta \leqslant \alpha/8$ or $\gamma \leqslant \beta/8$, then we are done because either $|X^{(1)}| \leqslant \alpha n/8$ or $|X^{(2)}| \leqslant \beta n/8$. Thus, we can assume that $\delta > \alpha/8$ and $\gamma > \beta/8$, which in turn implies that

$$k \leqslant \exp(c'(\delta\gamma \log n)^{1/2} - \log\log n) \tag{4.32}$$

by (4.2) if we take $c \leqslant c'/8$. It follows from (4.32) and the fact that Y satisfies (4.30) that we can apply Theorem 4.3 with k, $S = X^{(1)}$ and $T = X^{(2)}$, concluding that $L(Y) \geqslant k$. We have thus established Theorem 4.1(ii) and completed the proof. □

Bibliography

[1] H. L. Abbott. "Lower bounds for some Ramsey numbers". *Discrete Math.* 2, no. 4 (1972), 289–293.

[2] N. Alon. "Large sets in finite fields are sumsets". *J. Number Theory* 126, no. 1 (2007), 110–118.

[3] N. Alon and F. R. K. Chung. "Explicit construction of linear sized tolerant networks". In: *Proceedings of the First Japan Conference on Graph Theory and Applications (Hakone, 1986)*. Vol. 72. 1-3. 1988, 15–19.

[4] N. Alon, M. Krivelevich and B. Sudakov. "List coloring of random and pseudo-random graphs". *Combinatorica* 19, no. 4 (1999), 453–472.

[5] N. Alon. "The chromatic number of random Cayley graphs". *European J. Combin.* 34, no. 8 (2013), 1232–1243.

[6] N. Alon and A. Orlitsky. "Repeated communication and Ramsey graphs". *IEEE Trans. Inform. Theory* 41, no. 5 (1995), 1276–1289.

[7] N. Alon and H. T. Pham. "Random Cayley graphs and random sumsets" (2025). arXiv: `2509.02561 [math.CO]`.

[8] R. Alweiss, S. Lovett, K. Wu and J. Zhang. "Improved bounds for the sunflower lemma". *Ann. Math.* 194, no. 3 (2021), 795–815.

[9] L. Aragão, M. Campos, G. Dahia, R. Filipe and J. P. Marciano. "An exponential upper bound for induced Ramsey numbers" (2025). arXiv: `2509.22629 [math.CO]`.

[10] Y. Attwa, A. L. Vidal and P. Morris. "A note on multicolour Ramsey numbers and random sphere graphs" (2026). arXiv: `2602.02155 [math.CO]`.

[11] P. Balister, B. Bollobás, M. Campos, S. Griffiths, E. Hurley, R. Morris, J. Sahasrabudhe and M. Tiba. "Upper bounds for multicolour Ramsey numbers". *J. Amer. Math. Soc.* 39, no. 3 (2026), 765–780.

[12] P. Balister, B. Bollobás, R. Morris, J. Sahasrabudhe and M. Tiba. "Flat Littlewood polynomials exist". *Ann. Math.* 192, no. 3 (2020), 977–1004.

[13] J. Balogh, R. Morris and W. Samotij. "Independent sets in hypergraphs". *J. Amer. Math. Soc.* 28, no. 3 (2015), 669–709.

[14] J. Balogh, R. Morris and W. Samotij. "The method of hypergraph containers". In: *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018*. Vol. IV. World Sci. Publ., Hackensack, NJ, 2018, 3059–3092.

[15] J. Balogh and W. Samotij. "An efficient container lemma". *Discrete Anal.* (2020), Paper No. 17, 56pp.

[16] J. Beck. "On size Ramsey number of paths, trees, and circuits. I". *J. Graph Theory* 7, no. 1 (1983), 115–129.

[17] B. Bedert. "Large sum-free subsets of sets of integers via L^1-estimates for trigonometric series" (2025). arXiv: 2502.08624 [math.NT].

[18] S. L. Berg and M. Henk. "Discrete analogues of John's theorem". *Mosc. J. Comb. Number Theory* 8 (2019), 367–378.

[19] B. Bollobás. *Random graphs*. Academic Press, Inc., London, 1985, xvi+447.

[20] B. Bollobás and I. Leader. "Compressions and isoperimetric inequalities". *J. Combin. Theory Ser. A* 56, no. 1 (1991), 47–62.

[21] B. Bollobás, I. Leader and M. Tiba. "Large sumsets from medium-sized subsets" (2022). arXiv: 2206.09366 [math.CO].

[22] B. Bollobás, I. Leader and M. Tiba. "Large sumsets from small subsets". *Israel J. Math.* 268, no. 1 (2025), 253–314.

[23] J. Bourgain. "On arithmetic progressions in sums of sets of integers". In: *A tribute to Paul Erdős*. Cambridge Univ. Press, Cambridge, 1990, 105–109.

[24] C. Bowtell, R. Hancock and J. Hyde. "Proof of the Kohayakawa–Kreuter conjecture for the majority of cases" (2023). arXiv: 2307.16760 [math.CO].

[25] M. Bucić, T. Nguyen, A. Scott and P. Seymour. "Induced subgraph density. I. A loglog step towards Erdős–Hajnal". *Int. Math. Res. Not. IMRN* 2024, no. 12 (2024), 9991–10004.

[26] M. Campos. "On the number of sets with a given doubling constant". *Israel J. Math.* 236, no. 2 (2020), 711–726.

[27] M. Campos, M. Coulson, O. Serra and M. Wötzel. "The typical approximate structure of sets with bounded sumset". *SIAM J. Discrete Math.* 37, no. 3 (2023), 1386–1418.

[28] M. Campos, G. Dahia and Y. Kohayakawa. "Long arithmetic progressions in subsetsums of sparse random sets of integers" (To appear).

[29] M. Campos, G. Dahia and J. P. Marciano. "On the independence number of sparser random Cayley graphs". *J. Lond. Math. Soc.* 110, no. 6 (2024), Paper No. e70041, 54.

[30] M. Campos, S. Griffiths, R. Morris and J. Sahasrabudhe. "An exponential improvement for diagonal Ramsey". *Ann. Math.* 203, no. 3 (2026), 869–932.

[31] M. Campos, M. Jenssen, M. Michelen and J. Sahasrabudhe. "A new lower bound for sphere packing" (2023). arXiv: 2312.10026 [math.MG].

[32] M. Campos, M. Jenssen, M. Michelen and J. Sahasrabudhe. "A new lower bound for the Ramsey numbers $R(3,k)$" (2025). arXiv: 2505.13371 [math.CO].

[33] M. Campos and C. Pohoata. "An update on multicolor Ramsey lower bounds" (2026). arXiv: 2601.15183 [math.CO].

[34] M. Campos and W. Samotij. "Towards an optimal hypergraph container lemma". *Combinatorica* (to appear). arXiv: `2408.06617 [math.CO]`.

[35] M. Chang. "A polynomial bound in Freiman's theorem". *Duke Math. J.* 113, no. 3 (2002), 399–419.

[36] M. Christoph, A. Martinsson, R. Steiner and Y. Wigderson. "Resolution of the Kohayakawa–Kreuter conjecture". *Proc. Lond. Math. Soc.* 130, no. 1 (2025), Paper No. e70013, 34.

[37] C. Chvatál, V. Rödl, E. Szemerédi and W. T. Trotter Jr. "The Ramsey number of a graph with bounded maximum degree". *J. Combin. Theory Ser. B* 34, no. 3 (1983), 239–243.

[38] D. Conlon and W. T. Gowers. "Combinatorial theorems in sparse random sets". *Ann. Math.* 184, no. 2 (2016), 367–454.

[39] D. Conlon. "A new upper bound for diagonal Ramsey numbers". *Ann. Math.* 170, no. 2 (2009), 941–960.

[40] D. Conlon, D. Dellamonica, S. La Fleur, V. Rödl and M. Schacht. "A note on induced Ramsey numbers". In: *A Journey Through Discrete Mathematics*. Springer, Cham, 2017, 357–366.

[41] D. Conlon and A. Ferber. "Lower bounds for multicolor Ramsey numbers". *Adv. Math.* 378 (2021). Paper No. 107528, 5 pp.

[42] D. Conlon, J. Fox, H. T. Pham and L. Yepremyan. "On the clique number of random Cayley graphs and related topics" (2024). arXiv: `2412.21194 [math.CO]`.

[43] D. Conlon, J. Fox and B. Sudakov. "On two problems in graph Ramsey theory". *Combinatorica* 32, no. 5 (2012), 513–535.

[44] D. Conlon, J. Fox and B. Sudakov. "Recent developments in graph Ramsey theory". In: *Surveys in combinatorics 2015*. Vol. 424. London Math. Soc. Lecture Note Ser. Cambridge Univ. Press, Cambridge, 2015, 49–118.

[45] E. Croot, I. Łaba and O. Sisask. "Arithmetic progressions in sumsets and L^p-almost-periodicity". *Combin. Probab. Comput.* 22, no. 3 (2013), 351–365.

[46] K. Cwalina and T. Schoen. "A linear bound on the dimension in Green–Ruzsa's theorem". *J. Number Theory* 133, no. 4 (2013), 1262–1269.

[47] W. Deuber. "A generalization of Ramsey's theorem". In: *Infinite and Finite Sets*. Vol. 10. Colloq. Math. Soc. János Bolyai. Amsterdam-London: North-Holland, 1975, 323–332.

[48] S. Eberhard, B. Green and F. Manners. "Sets of integers with no large sum-free subset". *Ann. Math.* 180, no. 2 (2014), 621–652.

[49] P. Erdős. "Some remarks on the theory of graphs". *Bull. Amer. Math. Soc.* 53 (1947), 292–294.

[50] P. Erdős. "Extremal problems in number theory". In: *Proc. Sympos. Pure Math., Vol. VIII*. Amer. Math. Soc., Providence, RI, 1965, 181–189.

[51] P. Erdős. “Problems and results on finite and infinite graphs”. In: *Recent advances in graph theory (Proc. Second Czechoslovak Sympos., Prague, 1974)*. Academia, Prague, 1975, 183–192. (loose errata).

[52] P. Erdős. “On some problems in graph theory, combinatorial analysis and combinatorial number theory”. In: *Graph theory and combinatorics (Cambridge, 1983)*. Academic Press, London, 1984, 1–17.

[53] P. Erdős, A. Hajnal and L. Pósa. “Strong embeddings of graphs into colored graphs”. In: *Infinite and finite sets*. Vol. 10. Colloq. Math. Soc. János Bolyai. Amsterdam-London: North-Holland, 1975, 585–595.

[54] P. Erdős and G. Szekeres. “A combinatorial problem in geometry”. *Compositio Math.* 2 (1935), 463–470.

[55] K. Ford, B. Green, S. Konyagin, J. Maynard and T. Tao. “Long gaps between primes”. *J. Amer. Math. Soc.* 31, no. 1 (2018), 65–105.

[56] J. Fox, S. Luo, H. T. Pham and Y. Zhou. “Small subsets with large sumset: Beyond the Cauchy–Davenport bound”. en. *Comb. Probab. Comput.* 33 (2024), 1–21.

[57] J. Fox and B. Sudakov. “Induced Ramsey-type theorems”. *Adv. Math.* 219, no. 6 (2008), 1771–1800.

[58] J. Fox and B. Sudakov. “Density theorems for bipartite graphs and related Ramsey-type results”. *Combinatorica* 29, no. 2 (2009), 153–196.

[59] P. Frankl and V. Rödl. “Large triangle-free subgraphs in graphs without K_4”. *Graphs Combin.* 2, no. 2 (1986), 135–144.

[60] G. A. Freĭman. “The addition of finite sets. I”. *Izv. Vysš. Učebn. Zaved. Matematika* 1959, no. 6 (1959), 202–213.

[61] G. A. Freĭman. *Foundations of a structural theory of set addition*. Translations of Mathematical Monographs, Vol 37. Translated from the Russian. American Mathematical Society, Providence, RI, 1973, vii+108.

[62] E. Friedgut, E. Kuperwasser, W. Samotij and M. Schacht. “Sharp thresholds for Ramsey properties”. *Forum Math. Sigma* 14 (2026), Paper No. e32.

[63] H. Furstenberg. “Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions”. *Journal d’Analyse Mathématique* 31 (1977), 204–256.

[64] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory*. M. B. Porter Lectures. Princeton University Press, Princeton, NJ, 1981, xi+203.

[65] S. Glock, D. Kühn, A. Lo and D. Osthus. “The existence of designs via iterative absorption: hypergraph F-designs for arbitrary F”. *Mem. Amer. Math. Soc.* 284, no. 1406 (2023), v+131.

[66] W. T. Gowers. “Lower bounds of tower type for Szemerédi’s uniformity lemma”. *Geom. Funct. Anal.* 7, no. 2 (1997), 322–337.

[67] W. T. Gowers. “A new proof of Szemerédi’s theorem”. *Geom. Funct. Anal.* 11, no. 3 (2001), 465–588.

[68] W. T. Gowers. "Decompositions, approximate structure, transference, and the Hahn-Banach theorem". *Bull. Lond. Math. Soc.* 42, no. 4 (2010), 573–606.

[69] R. L. Graham and V. Rödl. "Numbers in Ramsey theory". In: *Surveys in combinatorics 1987 (New Cross, 1987)*. Vol. 123. London Math. Soc. Lecture Note Ser. Cambridge Univ. Press, Cambridge, 1987, 111–153.

[70] R. L. Graham, V. Rödl and A. Ruciński. "On bipartite graphs with linear Ramsey numbers". *Combinatorica* 21, no. 2 (2001), 199–209.

[71] B. Green. "Arithmetic progressions in sumsets". *Geom. Funct. Anal.* 12, no. 3 (2002), 584–597.

[72] B. Green. "Counting sets with small sumset, and the clique number of random Cayley graphs". *Combinatorica* 25, no. 3 (2005), 307–326.

[73] B. Green and R. Morris. "Counting sets with small sumset and applications". *Combinatorica* 36, no. 2 (2016), 129–159.

[74] B. Green and T. Tao. "Compressions, convex geometry and the Freiman–Bilu theorem". *Q. J. Math.* 57, no. 4 (2006), 495–504.

[75] M. Hamel and I. Łaba. "Arithmetic structures in random sets". *Integers* 8 (2008), A04, 21pp.

[76] T. E. Harris. "A lower bound for the critical probability in a certain percolation process". *Proc. Cambridge Philos. Soc.* 56 (1960), 13–20.

[77] Z. Hefty, P. Horn, D. King and F. Pfender. "Improving $R(3,k)$ in just two bites" (2025). arXiv: `2510.19718 [math.CO]`.

[78] P. van Hintum and P. Keevash. "Sharp bounds for a discrete John's theorem". *Combin. Probab. Comput.* 33 (2024), 1–3.

[79] S. Janson, T. Łuczak and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000, xii+333.

[80] Y. Jing and A. Mudgal. "Kemperman's inequality and Freiman's lemma via few translates" (2023). arXiv: `2307.03066 [math.CO]`.

[81] J. Kahn and G. Kalai. "Thresholds and expectation thresholds". *Combin. Probab. Comput.* 16, no. 3 (2007), 495–502.

[82] P. Keevash. "The existence of designs" (2014). arXiv: `1401.3665 [math.CO]`.

[83] J. H. Kim. "The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$". *Random Structures Algorithms* 7, no. 3 (1995), 173–207.

[84] B. Klartag. "Lattice packing of spheres in high dimensions using a stochastically evolving ellipsoid" (2025). arXiv: `2504.05042 [math.MG]`.

[85] B. Klartag and J. Lehec. "Affirmative resolution of Bourgain's slicing problem using Guan's bound". *Geom. Funct. Anal.* 35, no. 4 (2025), 1147–1168.

[86] Y. Kohayakawa, H. J. Prömel and V. Rödl. "Induced Ramsey numbers". *Combinatorica* 18, no. 3 (1998), 373–404.

[87] Y. Kohayakawa, T. Łuczak and V. Rödl. "Arithmetic progressions of length three in subsets of a random set". *Acta Arith.* 75, no. 2 (1996), 133–163.

[88] Y. Kohayakawa, V. Rödl, M. Schacht and E. Szemerédi. "Sparse partition universal graphs for graphs of bounded degree". *Adv. Math.* 226, no. 6 (2011), 5041–5065.

[89] S. V. Konyagin and V. F. Lev. "Combinatorics and linear algebra of Freiman's isomorphism". *Mathematika* 47, no. 1-2 (2000), 39–51.

[90] E. Kuperwasser, W. Samotij and Y. Wigderson. "On the Kohayakawa–Kreuter conjecture". *Math. Proc. Cambridge Philos. Soc.* 178, no. 3 (2025), 293–320.

[91] J. Leng, A. Sah and M. Sawhney. *Improved Bounds for Szemerédi's Theorem.* 2024. arXiv: 2402.17995 [math.CO].

[92] T. Łuczak and V. Rödl. "On induced Ramsey numbers for graphs with bounded maximum degree". *J. Combin. Theory Ser. B* 66, no. 2 (1996), 324–333.

[93] A. W. Marcus, D. A. Spielman and N. Srivastava. "Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem". *Ann. Math.* 182, no. 1 (2015), 327–350.

[94] S. Mattheus and J. Verstraete. "The asymptotics of $r(4,t)$". *Ann. Math.* 199, no. 2 (2024), 919–941.

[95] R. Miyazaki. "Arithmetic progressions in sumsets of random sets". MA thesis. Universidade de São Paulo, 2023.

[96] R. Morris, W. Samotij and D. Saxton. "An asymmetric container lemma and the structure of graphs with no induced 4-cycle". *J. Eur. Math. Soc. (JEMS)* 26, no. 5 (2024), 1655–1711.

[97] F. Mousset, R. Nenadov and W. Samotij. "Towards the Kohayakawa–Kreuter conjecture on asymmetric Ramsey properties". *Combin. Probab. Comput.* 29, no. 6 (2020), 943–955.

[98] R. Nenadov. "A remark on the independence number of sparse random Cayley sum graphs" (2025). arXiv: 2503.02100 [math.CO].

[99] R. Nenadov. "Improved bound on the number of cycle sets" (2025). arXiv: 2501.09904 [math.CO].

[100] R. Nenadov and A. Steger. "A short proof of the random Ramsey theorem". *Combin. Probab. Comput.* 25, no. 1 (2016), 130–144.

[101] R. Nenadov and L. Verlinde. "The Typical Structure of Sets with Bounded Sumset" (To appear).

[102] H. H. Nguyen. "On two-point configurations in a random set". *Integers* 9 (2009), A3, 41–45.

[103] T. Nguyen, A. Scott and P. Seymour. "Induced subgraphs density. IV. New graphs with the Erdős–Hajnal property" (2023). arXiv: 2307.06455 [math.CO].

[104] T. Nguyen, A. Scott and P. Seymour. "Induced subgraph density. VII. The five-vertex path". *Proc. Lond. Math. Soc.* 132, no. 3 (2026), Paper No. e70133, 21.

[105] J. Park and H. T. Pham. "A proof of the Kahn–Kalai conjecture". *J. Amer. Math. Soc.* 37, no. 1 (2024), 235–243.

[106] R. Raghavan. "Improved Bounds for the Freiman–Ruzsa Theorem" (2025). arXiv: 2512.11217 [math.NT].

[107] F. P. Ramsey. "On a Problem of Formal Logic". *Proc. London Math. Soc.* 30, no. 4 (1930), 264–286.

[108] V. Rödl. "The dimension of a graph and generalized Ramsey theorems". Master's thesis. Charles University, 1973.

[109] V. Rödl and A. Ruciński. "Threshold functions for Ramsey properties". *J. Amer. Math. Soc.* 8, no. 4 (1995), 917–942.

[110] V. Rödl and E. Szemerédi. "On size Ramsey numbers of graphs with bounded degree". *Combinatorica* 20, no. 2 (2000), 257–262.

[111] I. Z. Ruzsa. "Generalized arithmetical progressions and sumsets". *Acta Math. Hungar.* 65, no. 4 (1994), 379–388.

[112] I. Z. Ruzsa. "Arithmetic progressions in sumsets". *Acta Arith.* 60, no. 2 (1991), 191–202.

[113] A. Sah. "Diagonal Ramsey via effective quasirandomness". *Duke Math. J.* 172, no. 3 (2023), 545–567.

[114] W. Samotij. "Counting independent sets in graphs". *European J. Combin.* 48 (2015), 5–18.

[115] T. Sanders. "The structure theory of set addition revisited". *Bull. Amer. Math. Soc.* 50, no. 1 (2013), 93–127.

[116] W. Sawin. "An improved lower bound for multicolor Ramsey numbers and a problem of Erdős". *J. Combin. Theory Ser. A* 188 (2022). Paper No. 105579, 11 pp.

[117] D. Saxton and A. Thomason. "Hypergraph containers". *Invent. Math.* 201, no. 3 (2015), 925–992.

[118] M. Schacht. "Extremal results for random discrete structures". *Ann. Math.* 184, no. 2 (2016), 333–365.

[119] I. Schur. "Über die Kongruenz $x^m + y^m \equiv z^m \pmod{p}$". *Jahresbericht der Deutschen Mathematiker-Vereinigung* 25 (1916), 114–117.

[120] X. Shao. "On an almost all version of the Balog–Szemerédi–Gowers theorem". *Discrete Anal.* (2019). Paper No. 12, 18pp.

[121] J. Spencer. "Ramsey's theorem—a new lower bound". *J. Combin. Theory Ser. A* 18 (1975), 108–115.

[122] E. Szemerédi. "On sets of integers containing no k elements in arithmetic progression". *Acta Arith.* 27 (1975), 199–245.

[123] E. Szemerédi. "Regular partitions of graphs". In: *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*. Vol. 260. Colloq. Internat. CNRS. CNRS, Paris, 1978, 399–401.

[124] T. Tao and V. Vu. "John-type theorems for generalized arithmetic progressions and iterated sumsets". *Adv. Math.* 219 (2008), 428–449.

[125] A. Thomason. "Pseudo-Random Graphs". In: *Annals of Discrete Mathematics (33)*. Vol. 144. North-Holland Mathematics Studies. North-Holland, 1987, 307–331.

[126] A. Thomason. "An upper bound for some Ramsey numbers". *J. Graph Theory* 12, no. 4 (1988), 509–517.

[127] K. Tikhomirov. "On bounded degree graphs with large size-Ramsey numbers". *Combinatorica* 44, no. 1 (2024), 9–14.

[128] B. L. van der Waerden. "Beweis einer Baudetschen Vermutung". *Nieuw Archief voor Wiskunde*. 2nd ser. 15 (1927), 212–216.

[129] D. L. Wang and P. Wang. "Discrete isoperimetric problems". *SIAM J. Appl. Math.* 32, no. 4 (1977), 860–870.

[130] Y. Wigderson. "An improved lower bound on multicolor Ramsey numbers". *Proc. Amer. Math. Soc.* 149, no. 6 (2021), 2371–2374.