THE METHOD OF HYPERGRAPH CONTAINERS

ROBERT MORRIS

ABSTRACT. In this series of four lectures, we will give a fast-paced introduction to a recently-developed technique in probabilistic combinatorics, known as the *method of hypergraph containers*. This general technique can be applied in a wide range of settings; to give just a few examples, it has been used to study the following questions:

1. What is the largest H-free subgraph of $G(n,p)$?
2. How many sets $A \subset [n]$ contain no k-term arithmetic progression?
3. When does every r-colouring of $G(n,p)$ contain a monochromatic copy of H?
4. How many union-free families are contained in $\mathcal{P}(n)$?
5. What is the volume of the metric polytope?

The solutions of these problems (and many others) are based on the same fundamental principle: the objects in question exhibit a certain kind of 'clustering', which allows one to count them one cluster at a time, using (in each case) a suitable 'supersaturation' theorem.

Our plan is as follows: In Lecture 1, we will give a relatively gentle introduction to the method, focused on the example of triangle-free subgraphs of $G(n,p)$; in Lecture 2 we will state the general container lemma, and give several simple but important applications; finally, in Lectures 3 and 4, we will discuss some more advanced applications.

LECTURE 1: WHAT DO YOU DO WHEN THE 1ST MOMENT BLOWS UP?

In probabilistic combinatorics one is often faced with the following situation: you want to show that (with high probability) no member of some family of 'bad' events occurs, but the *expected* number of such events is large. Such situations often arise when there is positive correlation between the different bad events in your family, and the effect of such correlations can be difficult to bound. In these lectures we will discuss a recently-discovered method of dealing with certain situations of this type, whose basic idea can be summarised as follows:

"Independent sets in many 'natural' hypergraphs are 'clustered' together."

In this first lecture, we will illustrate this idea with a relatively simple, but important example.

Definition 1. The extremal number of a graph H with respect to the Erdős-Rényi random graph $G(n,p)$ is defined to be

$$\mathrm{ex}\big(G(n,p),H\big) := \max\big\{e(G) : G \subset G(n,p) \text{ and } H \not\subset G\big\}.$$

Our aim in this lecture is to prove the following theorem of Frankl and Rödl [32].

Theorem 2. *If $p \gg 1/\sqrt{n}$, then*

$$\mathrm{ex}\big(G(n,p),K_3\big) = \left(\frac{1}{4} + o(1)\right)pn^2 \tag{1}$$

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

To see the lower bound, simply consider the intersection of $G(n,p)$ with a copy of $K_{n/2,n/2}$; we will prove the upper bound. One natural first attempt at a proof of Theorem 2 would be to define a random variable

$$X_m := \big|\big\{G \subset G(n,p) : e(G) = m \text{ and } K_3 \not\subset G\big\}\big|,$$

and calculate the expected value of X_m. If $\mathbb{E}[X_m] = o(1)$, then it follows by Markov's inequality that $\text{ex}\big(G(n,p), K_3\big) \leqslant m$ with high probability. However, when we do the calculation we obtain instead

$$\mathbb{E}[X_m] \geqslant \binom{\text{ex}(n,K_3)}{m} p^m = \left(\big(1+o(1)\big)\frac{epn^2}{4m}\right)^m \gg 1$$

for all $m \leqslant p\binom{n}{2}$, so this approach fails.

What are we to do? Well, the reason the expected value of X_m blows up is that the triangle-free graphs with m edges are 'clustered' together, and this creates strong positive correlations between the events encoding their appearance in $G(n,p)$. If we can understand this clustering, we have a chance of grouping them into a relatively small number of 'bunches', and dealing with a whole bunch in a single step. To be more precise, we'd like to prove the following 'container' theorem.

Theorem 3 (The container theorem for triangle-free graphs)**.** *For each $n \in \mathbb{N}$, there exists a collection $\mathcal{G}$ of graphs on n vertices with the following properties:*

- *(a)* $|\mathcal{G}| \leqslant n^{O(n^{3/2})}$.
- *(b) Each $G \in \mathcal{G}$ contains $o(n^3)$ triangles.*
- *(c) Each triangle-free graph on n vertices is contained in some $G \in \mathcal{G}$.*

In words, this theorem says that there exists a relatively small collection of graphs $\mathcal{G}$, each of which contains few triangles, with the property that *every* triangle-free graph is a subgraph of some member of $\mathcal{G}$. To motivate the statement, let's begin by deducing from it a slightly weaker version of Theorem 2. To do so, we will need the following classical 'supersaturation' theorem.

Theorem 4 (Supersaturation for triangles)**.** *For every $\varepsilon > 0$, there exists $\delta > 0$ such that the following holds. If G is a graph on n vertices with*

$$e(G) \geqslant \left(\frac{1}{4} + \varepsilon\right) n^2$$

edges, then G has at least δn^3 triangles.

Theorem 4 is a straightforward consequence of Szemerédi's regularity lemma, and can also be proved by more elementary means, see the exercises. Note that it follows immediately from this theorem that a graph on n vertices with $o(n^3)$ triangles has at most $\big(1/4 + o(1)\big)n^2$ edges.

Now, suppose that $G(n,p)$ contains a triangle-free graph H with $m \geqslant (1/4+\varepsilon)pn^2$ edges. By Theorem 3, there exists a graph $G \in \mathcal{G}$ such that $H \subset G$, and since G contains $o(n^3)$ triangles, it follows that $e(G) \leqslant \big(1/4 + o(1)\big)n^2$. However, by Chernoff's inequality, the probability that such a graph G contains more than m edges of $G(n,p)$ is at most

$$\mathbb{P}\Big(\text{Bin}\big(e(G), p\big) \geqslant m\Big) \leqslant e^{-\Omega(pn^2)}.$$

Now, writing Y_m for the number of graphs $G \in \mathcal{G}$ that contain at least m edges of $G(n,p)$, and recalling that $|\mathcal{G}| \leqslant n^{O(n^{3/2})}$, we obtain

$$\mathbb{P}\big(Y_m \geqslant 1\big) \;\leqslant\; \mathbb{E}\big[Y_m\big] \;\leqslant\; \sum_{G \in \mathcal{G}} \mathbb{P}\Big(e\big(G \cap G(n,p)\big) \geqslant m\Big) \;\leqslant\; n^{O(n^{3/2})} e^{-\Omega(pn^2)} \;\to\; 0$$

if $p \gg \log n/\sqrt{n}$. We have thus proved (1) under a slightly stronger assumption on p.[1]

A container lemma for 3-uniform hypergraphs. In order to prove Theorem 3, we will need to work in a more general setting: that of 'almost linear' 3-uniform hypergraphs. To consider triangle-free graphs from this point of view, we will need the following simple but key definition.

Definition 5. The *hypergraph encoding triangles in K_n* is the 3-uniform hypergraph $\mathcal{H}$ with vertex set $V(\mathcal{H}) = E(K_n)$ and edge set

$$E(\mathcal{H}) = \Big\{ \big\{f_1, f_2, f_3\big\} \subset E(K_n) \,:\, \big\{f_1, f_2, f_3\big\} = E(K_3) \Big\}.$$

In words, the edges of $\mathcal{H}$ encode the triangles in K_n.

The next lemma is the key step in the proof of Theorem 3. Given a hypergraph $\mathcal{H}$, let us write $\Delta_\ell(\mathcal{H})$ for the maximum degree of a set of ℓ vertices of $\mathcal{H}$, that is

$$\Delta_\ell(\mathcal{H}) = \max\big\{ d_{\mathcal{H}}(A) \,:\, A \subset V(\mathcal{H}),\, |A| = \ell \big\},$$

where $d_{\mathcal{H}}(A) = \big|\big\{B \in E(\mathcal{H}) : A \subset B\big\}\big|$, and $\mathcal{I}(\mathcal{H})$ for the collection of independent sets of $\mathcal{H}$.

The Hypergraph Container Lemma for 3-uniform hypergraphs. *For every $c > 0$, there exists $\delta > 0$ such that the following holds. Let $\mathcal{H}$ be a 3-uniform hypergraph with average degree d, set $\tau := 1/\sqrt{d}$, and suppose that $\tau \leqslant \delta$, and that*

$$\Delta_1(\mathcal{H}) \leqslant c \cdot d \qquad \text{and} \qquad \Delta_2(\mathcal{H}) \leqslant c \cdot \sqrt{d}.$$

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

$$|\mathcal{C}| \leqslant \binom{v(\mathcal{H})}{\tau \cdot v(\mathcal{H})},$$

such that

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

In order to help us to understand the statement of this lemma, let's see why it implies Theorem 3.

Sketch proof of Theorem 3, assuming the hypergraph container lemma for 3-uniform hypergraphs. Let $\mathcal{H}$ be the hypergraph encoding triangles in K_n, and note that

$$v(\mathcal{H}) = \binom{n}{2}, \qquad \Delta_2(\mathcal{H}) = 1 \qquad \text{and} \qquad d_{\mathcal{H}}(v) = n - 1$$

for every $v \in V(\mathcal{H})$. We apply the container lemma to $\mathcal{H}$ with $c = 1$, and obtain a collection $\mathcal{C}$ of subsets of $E(K_n)$ (that is, graphs on n vertices), with the following properties:

[1] We will see later how to remove the unwanted factor of $\log n$.

(*a*) Every triangle-free graph is a subgraph of some $C \in \mathcal{C}$,

(*b*) Each $C \in \mathcal{C}$ has at most $(1-\delta)e(K_n)$ edges.

Moreover, since $\tau = \Theta(1/\sqrt{n})$, we have $|\mathcal{C}| \leqslant n^{O(n^{3/2})}$. Now, if each $C \in \mathcal{C}$ contains $o(n^3)$ triangles then we are done; otherwise, we choose $C \in \mathcal{C}$ with at least εn^3 triangles and apply the container lemma to the hypergraph $\mathcal{H}[C]$ induced by C. Note that $\mathcal{H}$ has average degree at least $3\varepsilon n$, since each triangle in C corresponds to an edge of $\mathcal{H}[C]$, so we can apply the lemma with $c = 1/\varepsilon$. Since $\tau = \Theta(1/\sqrt{n})$ and $v(\mathcal{H}) = |C| \leqslant n^2$, we produce at most $n^{O(n^{3/2})}$ new containers in each application. Moreover, since the containers shrink by a factor of $(1-\delta)$ in each step, after a bounded number of steps they will have size at most εn^2, and hence have at most εn^3 triangles, as required. $\square$

The observant reader will have noticed one or two things about this "proof". First, we cheated slightly: the bound on $|\mathcal{G}|$ depends on the maximum number of triangles allowed in a container.[2] Second, for the application we are currently interested in (bounding $\mathrm{ex}(G(n,p),K_3)$), we could have replaced condition (*b*) with "each $G \in \mathcal{G}$ contains $o(n^3)$ triangles or has at most $n^2/4$ edges".

The proof of the container lemma. As a warm-up for the proof of the hypergraph container lemma, let's consider first the somewhat simpler setting of graphs.

The Graph Container Lemma. *For every $c > 0$, there exists $\delta > 0$ such that the following holds. Let G be a graph with average degree d and maximum degree $\Delta(G) \leqslant c \cdot d$, and set $\tau := 2\delta/d$. There exists a collection $\mathcal{C}$ of subsets of $V(G)$, with*

$$|\mathcal{C}| \leqslant \binom{v(G)}{\lceil \tau \cdot v(G) \rceil},$$

such that

(*a*) *For every $I \in \mathcal{I}(G)$ there exists $C \in \mathcal{C}$ such that $I \subset C$,*

(*b*) *$|C| \leqslant (1-\delta)v(G)$ for every $C \in \mathcal{C}$.*

The proof of the graph container lemma is originally due to Kleitman and Winston [40]. The key idea is to encode each independent set $I \in \mathcal{I}(G)$ by a subset $S = S(I) \subset I$, which we will refer to as the *fingerprint* of I. The set S should be small, and should have the following crucial property: knowing only that $S(I) = S$ is sufficient to guarantee that I avoids a positive proportion of the vertices of G. In other words, every independent set $I \in \mathcal{I}(G)$ with $S(I) = S$ should be contained in a set $C = C(S)$ with $|C| \leqslant (1-\delta)v(G)$. We will call the set C the *container* of I.

We will construct S using a simple deterministic algorithm, which goes through the vertices of G one by one, recording whether or not they are a member of I. The order we use to query the vertices is given by the following definition.

Definition 6. The *max-degree order* of a (hyper)graph G is the ordering $(v_1, \ldots, v_n)$ of $V(G)$ such that for each $i \in [n]$, v_i is the[3] maximum degree vertex of $G[\{v_i, \ldots, v_n\}]$.

[2] More precisely, we proved that there exists a collection of $|\mathcal{G}| \leqslant \exp\left(C(\varepsilon) n^{3/2} \log n\right)$ containers, each containing at most εn^3 triangles, where $C(\varepsilon) \to \infty$ as $\varepsilon \to 0$.

[3] We break ties using some predefined (and arbitrary) ordering on $V(G)$.

We will prove the graph container lemma using the following algorithm.

The Graph Container Algorithm. Given a graph G and an independent set $I \in \mathcal{I}(G)$, we will maintain a partition $V(G) = A \cup S \cup X$, where A are the 'active' vertices, S is the current version of the fingerprint, and X is the set of 'excluded' vertices, which we already know are not in I. We start with $A = V(G)$ and $S = X = \emptyset$.

Now, while $|X| \leqslant \delta v(G)$, repeat the following steps:

1. Let v be the first vertex of I in the max-degree order on $G[A]$.
2. Move v into S, i.e., set $S := S \cup \{v\}$.
3. Move the neighbours of v into X, i.e., set $X := X \cup N(v)$.
4. Move the vertices which preceded v in the max-degree order on $G[A]$ into X, i.e., set $X := X \cup W$, where $W = \big\{u \in A : u < v \text{ in the max-degree order on } G[A]\big\}$.
5. Remove the new vertices of $S \cup X$ from A, i.e., set $A := V(G) \setminus (S \cup X)$.

Finally, set $A(I) := A$, $S(I) := S$ and $X(I) := X$.

To prove the graph container lemma, we will show that the algorithm above produces a set $S = S(I)$ with the desired properties.

Proof of the graph container lemma. For each $I \in \mathcal{I}(G)$, recall that $S(I)$ and $A(I)$ are the final fingerprint and active sets produced by the graph container algorithm. We claim that if $I, I' \in \mathcal{I}(G)$ are such that $S(I) = S(I')$, then $A(I) = A(I')$. Indeed, if we ever query a vertex and discover that it is in $I \triangle I'$, then we place it in only one of the sets $S(I)$ and $S(I')$, which implies that they are different. Thus, since we query the vertices of G in the same order until we find such a vertex, it follows that process is identical for the two independent sets, and thus the outcome is identical.

We may therefore define the container $C(S)$ of a set S to be the set $A(I) \cup S$ for any $I \in \mathcal{I}(G)$ such that $S(I) = S$, and

$$\mathcal{C} := \big\{C(S) : S = S(I) \text{ for some } I \in \mathcal{I}(G)\big\}.$$

Since the algorithm ended with $|X(I)| \geqslant \delta v(G)$, it follows that $|C| \leqslant (1-\delta)v(G)$ for every $C \in \mathcal{C}$, and since $X(I)$ is disjoint from I (since it consists of neighbours of vertices in I, and vertices which were queried and found not to be in I), it follows that $I \subset C(S(I))$ for every $I \in \mathcal{I}(G)$.

It remains to show that $|S(I)| \leqslant \lceil \tau \cdot v(G) \rceil$ for every $I \in \mathcal{I}(G)$, which immediately implies the claimed bound on $|\mathcal{C}|$. To do so, we will show that when a vertex v is added to S in the algorithm, at least $d/2$ vertices[4] are added to X. This will be sufficient to prove the claimed bound on $|S(I)|$, since after $\tau \cdot v(G)$ vertices have been added to S, we will have $|X| \geqslant \delta v(G)$, as required.

To show that $d/2$ vertices are added to X, we will use the fact that $\Delta(G) \leqslant c \cdot d$. Indeed, if $|S| \leqslant \tau v(G)$, $|X| \leqslant \delta v(G)$ and $|W| \leqslant d/2$, then

$$e\big(G[A \setminus W]\big) \geqslant e(G) - \Big((\tau + \delta)v(G) + d/2\Big) \cdot \Delta(G) \geqslant \Big(1 - 2c(c+1)\delta\Big)e(G) \geqslant \frac{e(G)}{2}$$

if $\delta < 1/4c(c+1)$, since $\tau \leqslant 2c\delta$ and assuming that $d \leqslant 2\delta v(G)$ (since otherwise $|S(I)| = 1$). Since v is a vertex of maximum degree in $G[A \setminus W]$ (by the definition of the max-degree order), it follows that v has at least $d/2$ neighbours in $G[A]$, which are all moved to X, as required. $\square$

[4] More precisely, this should be $\min\{d/2, \delta v(G)\}$ vertices, but for simplicity let us assume that $d \leqslant 2\delta v(G)$.

We remark that the graph container lemma itself already has a large number of interesting applications, see e.g. the surveys of Balogh, Treglown and Wagner [19] and Samotij [53].

We are now ready to prove the container lemma for 3-uniform hypergraphs. The approach is similar, but instead of removing the neighbours of a vertex v when it is placed in S, we will need to store the graph induced by the edges containing it. Once we have acquired enough edges (by adding many vertices to S, and taking the union of the edges containing them), we will be able to use the graph container algorithm to move $\delta v(\mathcal{H})$ vertices into X.

We begin by defining the algorithm we will use to reduce a 3-uniform hypergraph to a graph.

The container algorithm for 3-uniform hypergraphs. Let $\mathcal{H}$ be a 3-uniform hypergraph, and let $I \in \mathcal{I}(\mathcal{H})$ be an independent set. We will maintain a fingerprint S, a 3-uniform hypergraph $\mathcal{A}$ of 'available' edges of $\mathcal{H}$, and a graph G of 'forbidden' pairs on $V(\mathcal{H})$. We start with $\mathcal{A} = \mathcal{H}$ and $S = E(G) = \emptyset$. Now, while $|S| < \lfloor \tau v(\mathcal{H})/2 \rfloor$ and $V(\mathcal{A}) \cap I \neq \emptyset$, repeat the following steps:

1. Let u be the first vertex of I in the max-degree order on $\mathcal{A}$.
2. Move u into S, i.e., set $S := S \cup \{u\}$.
3. Move the edges $N(u) = \big\{vw : uvw \in E(\mathcal{A})\big\}$ into G, i.e., set $E(G) := E(G) \cup N(u)$.
4. Remove u from $V(\mathcal{A})$, and also the vertices which preceded u in the max-degree order on $\mathcal{A}$, i.e., set $\mathcal{A} := \mathcal{A}[V(\mathcal{A}) \backslash W]$, where $W = \big\{w \in V(\mathcal{A}) : w \leqslant u \text{ in the max-degree order on } \mathcal{A}\big\}$.
5. Remove from $V(\mathcal{A})$ every vertex whose degree in the graph G is larger than $c \cdot \sqrt{d}$, i.e., set $\mathcal{A} := \mathcal{A}[V(\mathcal{A}) \setminus Y]$, where $Y = \big\{w \in V(\mathcal{A}) : d_G(w) > c \cdot \sqrt{d}\big\}$.
6. Remove from $E(\mathcal{A})$ every edge which contains an edge of G, i.e., set $E(\mathcal{A}) := E(\mathcal{A}) \setminus Z$, where $Z = \big\{f \in E(\mathcal{A}) : e(G[f]) \neq 0\big\}$.

Finally, set $\mathcal{A}(I) := \mathcal{A}$, $S_2(I) := S$ and $G(I) := G$.

The reason we need to remove the vertices of high degree in G from $V(\mathcal{A})$ is to ensure that the final graph $G(I)$ satisfies the condition $\Delta(G(I)) = O(\sqrt{d})$ required by the graph container lemma. We will show that either we remove many vertices from $\mathcal{A}$, or $e(G(I)) \geqslant \sqrt{d}n$, in which case we may apply the graph container algorithm to $G(I)$.

Proof of the hypergraph container lemma for 3-uniform hypergraphs. As in the proof of the graph container lemma, we begin by observing that if $I, I' \in \mathcal{I}(G)$ are such that $S_2(I) = S_2(I')$, then $G(I) = G(I')$ and $\mathcal{A}(I) = \mathcal{A}(I')$. Indeed, we query the vertices of $\mathcal{H}$ in the same order and always receive the same answer, so the process is identical for the two independent sets. We may therefore define $G(S)$ and $\mathcal{A}(S)$ to be $G(I)$ and $\mathcal{A}(I)$ for any $I \in \mathcal{I}(G)$ such that $S_2(I) = S$.

Fix $I \in \mathcal{I}(\mathcal{H})$, let $S_2 = S_2(I)$, and suppose first that

$$e\big(G(S_2)\big) \geqslant \frac{\sqrt{d} \cdot v(\mathcal{H})}{8c} \qquad \text{and} \qquad \Delta\big(G(S_2)\big) \leqslant 2c\sqrt{d}. \tag{2}$$

Observing that I is an independent set in $G(S_2)$, we may apply the graph container algorithm to $G(S_2)$ and I, and obtain a fingerprint $S_1 = S_1(I)$ with $|S_1| < \tau v(\mathcal{H})/2$, and a container C for I, depending only on $S_1 \cup S_2$, with $|C| \leqslant (1-\delta)v(\mathcal{H})$, as required.

Let us therefore assume that (2) fails to hold; we will show how to define a container C for I, depending only on S_2, with $|C| \leqslant (1-\delta)v(\mathcal{H})$. Note that the bound $\Delta\big(G(S_2)\big) \leqslant 2c\sqrt{d}$ always holds, since the degree of a vertex increases by at most $\Delta\big(N(u)\big) \leqslant \Delta_2(\mathcal{H}) \leqslant c\sqrt{d}$ in each step of the algorithm, and once a vertex of G has degree larger than $c\sqrt{d}$ it enters Y, so no more edges incident to it are added to G. Moreover, the bound on $e\big(G(S_2)\big)$ holds if at least $d/4$ edges are added to G in each step, since in that case[5]

$$e\big(G(S_2)\big) \;\geqslant\; \frac{\tau \cdot v(\mathcal{H})}{2} \cdot \frac{d}{4} \;\geqslant\; \frac{\sqrt{d} \cdot v(\mathcal{H})}{8},$$

and it also holds if at any stage of the algorithm we have $|Y| \geqslant v(\mathcal{H})/16c$, since then

$$e\big(G(S_2)\big) \;\geqslant\; \frac{|Y| \cdot c\sqrt{d}}{2} \;\geqslant\; \frac{\sqrt{d} \cdot v(\mathcal{H})}{32},$$

as required, since we may assume (without loss of generality) that $c \geqslant 4$.

We may therefore assume that there exists a step of the algorithm for which $e(G) < (\sqrt{d}/8c)v(\mathcal{H})$, $|Y| < v(\mathcal{H})/16c$ and at most $d/4$ edges are added to G. If at most $v(\mathcal{H})/16c$ other vertices have been removed from $\mathcal{A}$, then since $\Delta_1(\mathcal{H}) \leqslant c \cdot d$ and $\Delta_2(\mathcal{H}) \leqslant c \cdot \sqrt{d}$, it follows that

$$e(\mathcal{A}) \;\geqslant\; e(\mathcal{H}) - \frac{v(\mathcal{H}) \cdot \Delta_1(\mathcal{H})}{8c} - e(G) \cdot \Delta_2(\mathcal{H}) \;\geqslant\; e(\mathcal{H}) - \frac{d \cdot v(\mathcal{H})}{4} \;=\; \frac{e(\mathcal{H})}{4},$$

and hence, by the definition of the max-degree order, u has degree at least $d/4$ in $\mathcal{A}$. Since we have removed all edges of $\mathcal{A}$ containing an edge of G, these must correspond to at least $d/4$ new edges of G, which is a contradiction. Hence at least $v(\mathcal{H})/16c$ vertices of Y^c must have been removed from $\mathcal{A}$, and these are all either in S_2 or in I^c. Since $|S_2| \leqslant \tau v(\mathcal{H}) \leqslant \delta v(\mathcal{H})$ (by assumption), it follows that we can define a container C for I (depending only on S_2) with $|C| \leqslant (1-\delta)v(\mathcal{H})$, as claimed. This completes the proof of the hypergraph container lemma for 3-uniform hypergraphs. □

A stronger (but slightly more technical) version of the container lemma. We end this lecture by noting that the proof above in fact implies slightly stronger versions of the various theorems above, which in particular allow us to remove the final factor of $\log n$ required for the proof of Theorem 2. The statements of these stronger versions are slightly more technical, but seeing them in this relatively simple (and, by now, familiar) setting will help prepare us for the general statement at the start of Lecture 2, which will be given in this stronger form. We begin with the container lemma for 3-uniform hypergraphs.

The Hypergraph Container Lemma for 3-uniform hypergraphs (stronger version). *For every $c > 0$, there exists $\delta > 0$ such that the following holds. Let $\mathcal{H}$ be a 3-uniform hypergraph with average degree d, set $\tau := 1/\sqrt{d}$, and suppose that $\tau \leqslant \delta$, and that*

$$\Delta_1(\mathcal{H}) \leqslant c \cdot d \qquad \text{and} \qquad \Delta_2(\mathcal{H}) \leqslant c \cdot \sqrt{d}.$$

Then there exists a collection $\mathcal{C}$ of subsets of $V(\mathcal{H})$, and a function $f \colon \mathcal{P}\big(V(\mathcal{H})\big) \to \mathcal{C}$ such that:

(a) *For every $I \in \mathcal{I}(\mathcal{H})$ there exists $S \subset I$ with $|S| \leqslant \tau \cdot v(\mathcal{H})$ and $I \subset f(S)$,*

(b) *$|C| \leqslant (1-\delta)v(\mathcal{H})$ for every $C \in \mathcal{C}$.*

We leave it to the reader to check that the proof given above implies this stronger statement, which implies the following container theorem for triangle-free graphs.

[5]Here we assume that $V(\mathcal{A}) \cap I \neq \emptyset$ at the end of the algorithm; if it is, then our container is $(Y \cup S_2)^c$.

Theorem 7 (Another container theorem for triangle-free graphs)**.** *For every $\delta > 0$, there exists $C > 0$ such that the following holds. For every $n \in \mathbb{N}$, there exists a collection $\mathcal{G}$ of graphs on n vertices, and a function $f \colon \mathcal{P}\big(E(K_n)\big) \to \mathcal{G}$ such that:*

(a) Each $G \in \mathcal{G}$ contains at most δn^3 triangles.

(b) For every triangle-free graph H on n vertices, there exists a subgraph $S \subset H$ with

$$e(S) \leqslant Cn^{3/2} \qquad \text{and} \qquad H \subset f(S).$$

The proof is essentially identical to that of Theorem 3, given earlier, except using the stronger version of the hypergraph container lemma for 3-uniform hypergraphs.[6]

We are finally able to deduce the theorem of Frankl and Rödl.

Proof of Theorem 2. Let $\varepsilon > 0$, and suppose that $G(n,p)$ contains a triangle-free graph H with $m \geqslant (1/4 + 2\varepsilon)pn^2$ edges. By Theorem 7, there exists a subgraph $S \subset H$ with $e(S) \leqslant Cn^{3/2}$ and a corresponding graph $G(S) \in \mathcal{G}$ such that $H \subset G(S)$. Since $G(S)$ contains at most δn^3 triangles, it follows from Theorem 4 that

$$e\big(G(S)\big) \leqslant \left(\frac{1}{4} + \varepsilon\right) n^2.$$

Now, let $\mathcal{S}$ denote the collection of all graphs S obtained in this way, and note that if $S \subset H \subset G(n,p)$, then $S \subset G(n,p)$. Observe also that, by Chernoff's inequality, the probability that $G(S)$ contains more than $m - e(S)$ edges of $G(n,p)$ other than $E(S)$ is at most

$$\mathbb{P}\Big(\mathrm{Bin}\big(e(G(S)), p\big) \geqslant m - e(S)\Big) \leqslant e^{-\Omega(pn^2)}$$

if $p \gg 1/\sqrt{n}$, since then $e(S) \ll m$. Hence, writing Y for the number of sets $S \in \mathcal{S}$ such that $S \subset G(n,p)$ and $\big(E(G(S)) \setminus E(S)\big) \cap E\big(G(n,p)\big) \geqslant m - e(S)$, we obtain

$$\mathbb{P}\big(Y \geqslant 1\big) \leqslant \mathbb{E}[Y] \leqslant \sum_{S \in \mathcal{S}} p^{e(S)} e^{-\Omega(pn^2)} \leqslant \sum_{k=0}^{Cn^{3/2}} p^k \binom{\binom{n}{2}}{k} e^{-\Omega(pn^2)} \to 0$$

if $p \gg 1/\sqrt{n}$, as required. $\square$

Let us finish the lecture by mentioning the following related theorems of DeMarco and Kahn [30] and Osthus, Prömel and Taraz [48], which motivate several of the results stated later.

Theorem 8 (DeMarco and Kahn, 2015)**.** *There exists a constant $C > 0$ such that if*

$$p \geqslant C\sqrt{\frac{\log n}{n}}$$

then with high probability the largest triangle-free subgraph of $G(n,p)$ is bipartite.

Theorem 9 (Osthus, Prömel and Taraz, 2003)**.** *There exists a constant $C > 0$ such that if*

$$m \geqslant Cn^{3/2}\sqrt{\log n}$$

then almost all triangle-free graphs with m edges are bipartite.

[6] One technical point, which we ignore here for simplicity, is that as stated this gives a sequence of fingerprints for each triangle-free graph, rather than a single fingerprint. However, it is not hard to show that one can reconstruct the sequence of sets from their union, see [13].

Exercises

1. (a) Prove that if G is a d-regular graph on n vertices, and $A \subset V(G)$ is a set of vertices of size

$$|A| \geqslant \frac{n}{2} + \frac{\beta n^2}{2d}$$

for some $\beta > 0$, then $e(A) \geqslant \beta\binom{|A|}{2}$.

(b) Use the graph container lemma to deduce the following theorem of Alon [2].[7]

Theorem 10. *If G is a d-regular graph, then G has at most $2^{n/2+o(n)}$ independent sets.*[8]

(c) Prove as strong a bound on the error term as you can. How close to best possible is it?

(d) How many independent sets of size m can a d-regular graph have?

2. (a) Recall that, by Roth's theorem, every subset of $\mathbb{N}$ of positive density contains a 3-term arithmetic progression. Use this result to deduce the following supersaturation theorem.

Theorem 11. *For every $\varepsilon > 0$, there exists a $\delta > 0$ such that the following holds. If $A \subset [n]$ has size $|A| \geqslant \varepsilon n$, then A contains at least δn^2 triples $\{x, y, z\}$ with $x + y = 2z$.*

(b) Let us write $[n]_p$ for a *p-random* subset of $[n]$, that is, a subset which includes each element of $[n]$ independently at random with probability p. Use Theorem 11 and the container lemma for 3-uniform hypergraphs to deduce the following theorem of Kohayakawa, Łuczak and Rödl [42].

Theorem 12. *For every $\varepsilon > 0$ there exists $C > 0$ such that if $p \geqslant C/\sqrt{n}$, then with high probability every subset $A \subset [n]_p$ of size $|A| \geqslant \varepsilon pn$ contains a 3-term arithmetic progression.*

3. (a) Prove the following supersaturation theorem for Schur triples in $[n]$.

Theorem 13. *For every $\varepsilon > 0$, there exists a $\delta > 0$ such that the following holds. If $A \subset [n]$ has size $|A| \geqslant (1/2 + \varepsilon)n$, then A contains at least δn^2 triples $\{x, y, z\}$ with $x + y = z$.*

(b) Use Theorem 13 and the container lemma for 3-uniform hypergraphs to deduce the following theorem, which was first proved by Schacht [59] and by Conlon and Gowers [25].

Theorem 14. *If $p \gg 1/\sqrt{n}$ then with high probability the largest sum-free subset of a p-random subset $A \subset [n]$ has size $\big(1/2 + o(1)\big)pn$.*

(c) State and prove a 'supersaturated stability theorem' for Schur triples in $\mathbb{Z}_{2n}$, which says that a set $A \subset \mathbb{Z}_{2n}$ with $(1 - o(1))n$ elements and $o(n^2)$ Schur triples must contain $o(n)$ even numbers. Use the container lemma for 3-uniform hypergraphs to deduce the following theorem.

Theorem 15. *If $m \gg \sqrt{n}$ then almost all sum-free m-subsets of $\mathbb{Z}_{2n}$ contain $o(m)$ even numbers.*

(d) What should the corresponding theorem in $[n]$ say?

---

[7] This proof is due to Sapozhenko [54], who independently developed the container method for graphs. The details can be found in the survey of Samotij [53], as well as several other beautiful applications of graph containers.

[8] Here $o(1)$ is a function which tends to zero as $d \to \infty$.

4. Say that a graph G is *t-far from being bipartite* if G has no bipartite subgraph $H \subset G$ with $e(H) > e(G) - t$, and say G is *t-close to being bipartite* otherwise. For each $v \in V(G)$, define $B(v) = N(v)$ and $A(v) = B(v)^c$. Observe that

(a) The number of triangles containing v is equal to $e(B(v))$, the number of edges in $B(v)$.

(b) If G is t-far from being bipartite, then $e(A(v)) + e(B(v)) \geqslant t$.

(c) If we sum the degrees of the vertices in $A(v)$, then we count each edge between $A(v)$ and $B(v)$ once, and each edge inside $A(v)$ twice.

Using these simple observations, prove the following strengthening of Theorem 4.[9]

Theorem 16. *For every $n, t \geqslant 1$, the following holds. Every graph G on n vertices which is t-far from being bipartite contains at least*

$$\frac{n}{6}\left(e(G) + t - \frac{n^2}{4}\right)$$

triangles.

5. (a) Use Theorem 16 and the container lemma for 3-uniform hypergraphs to prove the following approximate version of Theorem 9, which was originally proved by Łuczak.

Theorem 17 (Łuczak, 2000)**.** *For each $\varepsilon > 0$, there exists a constant $C > 0$ such that if*

$$m \geqslant C n^{3/2}$$

then almost all triangle-free graphs with m edges are εm-close to bipartite.

(b) Prove that if $m \ll n^{3/2}$ then almost all triangle-free graphs are far from being bipartite.

6. (a) Prove the following slight extension of Ramsey's theorem for triangles.

Theorem 18. *For every $r \in \mathbb{N}$, there exists a $\delta > 0$ such that the following holds for all sufficiently large $n \in \mathbb{N}$. If G is a graph with n vertices and $e(G) > (1/2 - \delta)n^2$ edges, then every colouring $c\colon E(G) \to [r]$ contains at least δn^3 monochromatic triangles.*

(b) Use Theorem 31 and the container lemma for 3-uniform hypergraphs to deduce the following theorem of Rödl and Ruciński [50].[10]

Theorem 19. *For every $r \in \mathbb{N}$, and every function $p \gg 1/\sqrt{n}$, the following holds. With high probability every r-colouring of the edges of $G(n,p)$ contains a monochromatic triangle.*

(c) Use the same idea to prove the following theorem of Graham, Rödl and Rucinski [35].

Theorem 20. *For every $r \in \mathbb{N}$, and every function $p \gg 1/\sqrt{n}$, the following holds. With high probability every r-colouring of $[n]_p$ contains a monochromatic triple $\{x, y, z\}$ with $x + y = z$.*

(d) Can you prove corresponding lower bounds on the thresholds for these events?

[9] We remark that this proof is essentially due to Füredi, see [7] for the details.

[10] This proof is due to Nenadov and Steger [47].

Lecture 2: A general container lemma

In Lecture 1, we stated and proved the container lemmas for graphs, and for 3-uniform hypergraphs. The following generalisation to k-uniform hypergraphs was proved independently by Balogh, Morris and Samotij [13] and by Saxton and Thomason [57].

The Hypergraph Container Lemma. *For every $k \in \mathbb{N}$ and $c > 0$, there exists a $\delta > 0$ such that the following holds. Let $\mathcal{H}$ be a k-uniform hypergraph, and suppose that $\tau \in (0,1)$ satisfies*

$$\Delta_\ell(\mathcal{H}) \leqslant c \cdot \tau^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})} \tag{3}$$

for every $1 \leqslant \ell \leqslant k$. Then there exists a collection $\mathcal{C}$ of subsets of $V(\mathcal{H})$, and a function $f \colon \mathcal{P}\big(V(\mathcal{H})\big) \to \mathcal{C}$ such that:

- *(a) For every $I \in \mathcal{I}(\mathcal{H})$ there exists $S \subset I$ with $|S| \leqslant \tau \cdot v(\mathcal{H})$ and $I \subset f(S)$,*
- *(b) $|C| \leqslant (1-\delta)v(\mathcal{H})$ for every $C \in \mathcal{C}$.*

The proof for general k is similar to that described above, but the details are somewhat more complicated and technical, and so we will omit the proof and refer the interested reader to [13, 57] for the details. Instead, in this lecture we will aim to build up our intuition by focussing on some important (but relatively straightforward) applications.

Sparse sets avoiding k-term arithmetic progressions. Perhaps the most famous theorem in extremal combinatorics is that of Szemerédi, which states that every subset of $\mathbb{N}$ of positive density contains arbitrarily long arithmetic progressions. Our first application of the hypergraph container theorem for k-uniform hypergraphs will be the following strengthening of this theorem.

Theorem 21. *For every $\beta > 0$ and every $k \in \mathbb{N}$, there exist a constant C such that the following holds. For every sufficiently large $n \in \mathbb{N}$, if $m \geqslant Cn^{1-1/(k-1)}$, then there are at most*

$$\binom{\beta n}{m}$$

m-subsets of $\{1, \ldots, n\}$ that contain no k-term arithmetic progression.

As a almost immediate consequence, we will obtain the following theorem of Conlon and Gowers [25] and Schacht [59], which was originally conjectured by Kohayakawa, Łuczak and Rödl [42].

Corollary 22. *For every $\varepsilon > 0$ there exists $C > 0$ such that if $p \geqslant Cn^{-1/(k-1)}$, then with high probability every subset $A \subset [n]_p$ of size $|A| \geqslant \varepsilon pn$ contains a k-term arithmetic progression.*[11]

To prove Theorem 21, we will apply the hypergraph container lemma to the k-uniform hypergraph $\mathcal{H}$ that encodes k-term arithmetic progressions in $[n]$. To spell it out, this hypergraph has vertex set $V(\mathcal{H}) = [n]$ and edge set

$$E(\mathcal{H}) = \left\{ e \in \binom{[n]}{k} \,:\, e = \big\{a, a+d, \ldots, a+(k-1)d\big\} \text{ for some } a, d \in [n] \right\}.$$

[11] Recall that $[n]_p$ denotes a p-random subset of $[n] = \{1, \ldots, n\}$.

Note that

$$e(\mathcal{H}) = \Theta(n^2), \qquad \Delta_1(\mathcal{H}) = O(n) \qquad \text{and} \qquad \Delta_2(\mathcal{H}) = O(1),$$

where the implicit constants are all allowed to depend on k. It is now easy to check that (3) holds for $\mathcal{H}$ (and for some $c = c(k) > 0$) with $\tau = n^{-1/(k-1)}$; indeed, the inequality is tight (up to the value of c) only when $\ell = 1$ and $\ell = k$. The following supersaturated version of Szemerédi's theorem (which follows from the original via a simple averaging argument, originally observed by Varnavides) allows us to show that (3) also holds for $\mathcal{H}[A]$ for any subset $A \subset [n]$ with $|A| \geqslant \varepsilon n$.

Theorem 23. *For every $\varepsilon > 0$ and $k \in [n]$, there exists $\delta > 0$ such that the following holds. Every subset $A \subset [n]$ with $|A| \geqslant \varepsilon n$ contains at least δn^2 k-term arithmetic progressions.*

We are ready to count the k-AP-free sets of size m.

Proof of Theorem 21. Set $\varepsilon = \beta/2$. We claim that there exists a family $\mathcal{A} \subset \mathcal{P}(n)$, and a function $f\colon \mathcal{P}(n) \to \mathcal{A}$ such that:

(a) Each $A \in \mathcal{A}$ has at most εn elements.

(b) For every k-AP-free set $B \subset [n]$, there exists a subset $S \subset B$ with

$$|S| = O\big(n^{1-1/(k-1)}\big) \qquad \text{and} \qquad B \subset f(S).$$

Indeed, let B be a k-AP-free set, and apply the hypergraph container lemma to $\mathcal{H}$, with $\tau = n^{-1/(k-1)}$ and a suitable value of $c = c(k)$. We obtain a sets $S_1 \subset B$ and $C_1 = C_1(S_1) \supset B$ with $|S_1| \leqslant n^{1-1/(k-1)}$ and $|C_1| \leqslant (1-\delta)n$. We now iterate: given sets $S_t \subset B$ and $C_t = C_t(S_t) \supset B$ with $|S_t| \leqslant t \cdot n^{1-1/(k-1)}$ and $|C_t| \leqslant (1-\delta)^t n$, we do the following: if $|C_t| \leqslant \varepsilon n$ then we place C_t in $\mathcal{A}$ and set $f(S_t) = C_t$; otherwise we apply the hypergraph container lemma to $\mathcal{H}[C_t]$, with $\tau = n^{-1/(k-1)}$ and a suitable value of $c = c(k, \varepsilon)$, using Theorem 23 to prove that (3) holds. We obtain a fingerprint $S'_{t+1} \subset B$ and a container $C_{t+1} = C_{t+1}(C_t, S'_{t+1}) \supset B$ with $|C_{t+1}| \leqslant (1-\delta)|C_t| \leqslant (1-\delta)^{t+1} n$, and set $S_{t+1} = S_t \cup S'_{t+1} \subset B$, so $|S_{t+1}| \leqslant (t+1) n^{1-1/(k-1)}$, as required. Finally, since δ depends only on c, ε and k, after a constant number of steps we obtain a container with $|C_t| \leqslant \varepsilon n$.

Let $\mathcal{S}$ denote the collection of fingerprints S obtained in (b), and let C be a sufficiently large constant. Then, for each $m \geqslant C n^{1-1/(k-1)}$, the number of subsets $A \subset [n]$ of size m containing no k-term arithmetic progression is at most

$$\sum_{S \in \mathcal{S}} \binom{|f(S)|}{m - |S|} \leqslant \sum_{s \leqslant \varepsilon m} \binom{n}{s} \binom{\varepsilon n}{m - s} \leqslant \sum_{s \leqslant \varepsilon m} \left(\frac{en}{s}\right)^s \left(\frac{m}{\varepsilon n - m}\right)^s \binom{\varepsilon n}{m}.$$

Observe that, by Szemerédi's theorem, we may assume that $m = o(n)$. Noting that the function $x \mapsto (y/x)^x$ is increasing on $(0, y/e)$, it follows that the right-hand side is at most

$$\sum_{s \leqslant \varepsilon m} \left(\frac{2em}{\varepsilon s}\right)^s \binom{\varepsilon n}{m} \leqslant m \left(\frac{2e}{\varepsilon^2}\right)^{\varepsilon m} \binom{\varepsilon n}{m} \leqslant \binom{\beta n}{m},$$

where the final inequality follows since $\beta = 2\varepsilon$, so $\binom{\varepsilon n}{m} \leqslant 2^{-m} \binom{\beta n}{m}$. This proves Theorem 21. $\square$

We can now easily deduce Corollary 22 using Markov's inequality.

Proof of Corollary 22. For each $m \in \mathbb{N}$, let Y_m denote the number of m-subsets of $[n]_p$ that contain no k-term arithmetic progression. Set $\beta = \varepsilon$, and let $C = C(\beta, k)$ be the constant given by Theorem 21; it follows that if $p \geqslant (C/\varepsilon)n^{-1/(k-1)}$ then the conclusion of Theorem 21 holds for $m = \varepsilon pn$. Thus

$$\mathbb{P}\big(Y_m \geqslant 1\big) \leqslant \mathbb{E}\big[Y_m\big] \leqslant \binom{\beta n}{m} p^m \leqslant \left(\frac{\beta epn}{m}\right)^m = e^{-m},$$

as required. □

Containers for H-free graphs. We will next see how to deduce various results about the family of H-free graphs (these were first proved by Conlon and Gowers [25] and Schacht [59]). The first is a generalisation of the theorem of Frankl and Rödl proved in Lecture 1.

Theorem 24. *For every graph H with $\Delta(H) \geqslant 2$ and every $\varepsilon > 0$, there exists $C > 0$ such that if $p \geqslant Cn^{-1/m_2(H)}$, then*[12]

$$\mathrm{ex}\big(G(n,p), H\big) \leqslant \left(1 - \frac{1}{\chi(H) - 1} + \varepsilon\right) p \binom{n}{2}$$

with high probability.

The second is the following 'stability' version of Theorem 24.

Theorem 25. *For every graph H with $\Delta(H) \geqslant 2$ and every $\varepsilon > 0$, there exist $C > 0$ and $\delta > 0$ such that if $p \geqslant Cn^{-1/m_2(H)}$, then the following holds with high probability. Every H-free subgraph $G \subset G(n,p)$ with*

$$e(G) \geqslant \left(1 - \frac{1}{\chi(H) - 1} - \delta\right) p \binom{n}{2}$$

is εpn^2-close to being $\big(\chi(H) - 1\big)$-partite.

The third is a variant of Theorem 25 for H-free graphs with m edges.

Theorem 26. *For every graph H with $\chi(H) \geqslant 3$, and every $\varepsilon > 0$, there exists $C > 0$ such that the following holds. If $m \geqslant Cn^{2-1/m_2(H)}$, then almost all H-free graphs with n vertices and m edges are εm-close to being $\big(\chi(H) - 1\big)$-partite.*

In each case, we will apply the hypergraph container lemma to the k-uniform hypergraph that encodes copies of H in K_n, where $k = e(H)$, cf. Definition 5. This hypergraph $\mathcal{H}$ has vertex set $V(\mathcal{H}) = E(K_n)$ and edge set

$$E(\mathcal{H}) = \Big\{ \big\{f_1, \ldots, f_{e(H)}\big\} \subset E(K_n) \,:\, \big\{f_1, \ldots, f_{e(H)}\big\} = E(H) \Big\}.$$

We claim that there exists a constant $c = c(H)$ such that (3) holds with $\tau = n^{-1/m_2(H)}$. Indeed, note that

$$v(\mathcal{H}) = \Theta\big(n^2\big), \qquad e(\mathcal{H}) = \Theta\big(n^{v(H)}\big) \qquad \text{and} \qquad \tau^{e(F)-1} n^{v(F)-2} \geqslant 1$$

[12]Recall that $m_2(H) = \max\big\{\frac{e(F)-1}{v(F)-2} : F \subset H,\, v(F) \geqslant 3\big\}$.

for every $F \subseteq H$. Since we have

$$\Delta_\ell(\mathcal{H}) = O(1) \cdot \max\Big\{ n^{v(H)-v(F)} \colon F \subset H \text{ with } e(F) = \ell \Big\}$$

for each $1 \leqslant \ell \leqslant k$, it follows that

$$\Delta_\ell(\mathcal{H}) \cdot \left(\tau^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})} \right)^{-1} = O(1) \cdot \max_{F \subset H \colon e(F)=\ell} \big\{ \tau^{1-e(F)} n^{2-v(F)} \big\} = O(1),$$

as required.

We will also need the following supersaturated version of the Erdős-Simonovits stability theorem, which follows (for example) from the Szemerédi regularity lemma combined with the graph removal lemma. Note that when $H = K_3$ it follows from Exercise 3 of Lecture 1.

Theorem 27. *For every graph H and every $\varepsilon > 0$, there exists $\delta > 0$ such that the following holds for every $n \in \mathbb{N}$. If G is a graph on n vertices with*

$$e(G) \geqslant \left(1 - \frac{1}{\chi(H) - 1} - \delta \right) \binom{n}{2},$$

then either G is εn^2-close to being $\big(\chi(H) - 1\big)$-partite, or G contains at least $\varepsilon n^{v(H)}$ copies of H.

We can now prove the following container theorem for H-free graphs, which implies the three theorems stated above.

Theorem 28 (A container theorem for H-free graphs)**.** *For every graph H and every $\varepsilon > 0$, there exists $\delta > 0$ and $C > 0$ such that the following holds. For every $n \in \mathbb{N}$, there exists a collection $\mathcal{G}$ of graphs on n vertices, and a function $f \colon \mathcal{P}\big(E(K_n)\big) \to \mathcal{G}$ such that:*

(*a*) *Each $G \in \mathcal{G}$ either satisfies*

$$e(G) \leqslant \left(1 - \frac{1}{\chi(H) - 1} - \delta \right) \binom{n}{2},$$

or is εn^2-close to being $\big(\chi(H) - 1\big)$-partite.

(*b*) *For every H-free graph I on n vertices, there exists a subgraph $S \subset I$ with*

$$e(S) \leqslant C n^{2-1/m_2(H)} \qquad \text{and} \qquad I \subset f(S).$$

Proof. Let I be an H-free graph, and apply the hypergraph container lemma to the hypergraph $\mathcal{H}$ that encodes copies of H in K_n, with $\tau = n^{-1/m_2(H)}$ and a suitable value of $c = c(H)$. We obtain graphs $S_1 \subset I$ and $C_1 = C_1(S_1) \supset I$ with $|S_1| \leqslant n^{1-1/m_2(H)}$ and $|C_1| \leqslant (1-\delta)n$. We now iterate, as in the proof of Theorem 21, except placing C_t in $\mathcal{G}$ if either

$$e(C_t) \leqslant \left(1 - \frac{1}{\chi(H) - 1} - \delta \right) \binom{n}{2},$$

or C_t is εn^2-close to being $\big(\chi(H) - 1\big)$-partite. By Theorem 27, if neither of these holds, then C_t contains at least $\varepsilon n^{v(H)}$ copies of H, and so we can apply the hypergraph container lemma to $\mathcal{H}[C_t]$. Since $e(C_t) \leqslant (1-\delta)^t n^2$, we will arrive at a member of $\mathcal{G}$ after a bounded number of steps. □

The deduction of Theorems 24, 25 and 26 is now straightforward; we will give a quick sketch of the proof of Theorem 26, and leave the details to the reader.

Sketch proof of Theorem 26. Let $\mathcal{S}$ denote the collection of fingerprints S obtained in Theorem 28(b), and partition $\mathcal{S}$ into

$$\mathcal{S}_1 = \big\{ S \in \mathcal{S} : f(S) \text{ is } \varepsilon n^2\text{-close to being } \big(\chi(H)-1\big)\text{-partite} \big\}$$

and $\mathcal{S}_2 = \mathcal{S} \setminus \mathcal{S}_1$. Observe that the number of graphs G with n vertices and m edges that are βm-far from being $\big(\chi(H)-1\big)$-partite, and are contained in $f(S)$ for some $S \in \mathcal{S}_1$, is at most

$$\sum_{S \in \mathcal{S}_1} \binom{\varepsilon n^2}{\beta m - |S|} \binom{|f(S)|}{m - \beta m} \leqslant \binom{\beta n^2}{m}$$

if $\varepsilon = \varepsilon(\beta)$ is sufficiently small and $m \geqslant C n^{2-1/m_2(H)}$ for some sufficiently large constant $C > 0$. On the other hand, the total number of graphs G with n vertices and m edges that are contained in $f(S)$ for some $S \in \mathcal{S}_2$ is at most

$$\sum_{S \in \mathcal{S}_2} \binom{|f(S)|}{m - |S|} \leqslant 2^{-\delta^2 m} \binom{\mathrm{ex}(n,H)}{m}$$

if $m \geqslant C n^{2-1/m_2(H)}$, since $|f(S)| \leqslant \big(1 - \frac{1}{\chi(H)-1} - \delta\big)\binom{n}{2}$ for every $S \in \mathcal{S}_2$. Summing these two bounds, it follows that the number of H-free graphs G with n vertices and m edges that are βm-far from being $\big(\chi(H)-1\big)$-partite is at most

$$2^{-\delta^2 m + 1} \binom{\mathrm{ex}(n,H)}{m} \ll \binom{\mathrm{ex}(n,H)}{m}$$

for every $m \geqslant C n^{2-1/m_2(H)}$. Since the right-hand side is a lower bound on the number of H-free graphs with n vertices and m edges, the theorem follows. $\square$

The proof of Theorem 25 is similar, except we replace the counting above with an application of Chernoff's inequality in each container (cf. the proof of Theorem 2); we again leave the details to the reader. Finally, note that Theorem 24 follows from Theorem 25.

Ramsey properties of random graphs. In this section we will show how to use the hypergraph container lemma prove the following celebrated theorem of Rödl and Ruciński [50], using a nice argument of Nenadov and Steger [47].

Theorem 29. *For every graph H and $r \in \mathbb{N}$, if $p \gg n^{-1/m_2(H)}$ then the following holds. With high probability every r-colouring of the edges of $G(n,p)$ contains a monochromatic copy of H.*

We will apply the hypergraph container lemma to the k-uniform hypergraph $\mathcal{H}$ that encodes monochromatic copies of H in r-colourings of $E(K_n)$, where $k = e(H)$. This hypergraph consists of r disjoint copies of the hypergraph that encodes copies of H in K_n, so

$$v(\mathcal{H}) = r\binom{n}{2} \qquad \text{and} \qquad e(\mathcal{H}) = \Theta\big(n^{v(H)}\big).$$

It follows from our earlier calculation that there exists a constant $c = c(H,r)$ such that (3) holds with $\tau = n^{-1/m_2(H)}$. The definition of this hypergraph motivates the following definition.

Definition 30. An $[r]$-*coloured graph* is a graph G together with a function $c\colon E(G) \to \mathcal{P}(r)$ that associates a non-empty subset of $[r]$ to each edge. We say that an $[r]$-coloured graph is H-*free* if it contains no monochromatic copy of H, i.e., if there does not exist a copy of H in G whose edges have the following property: their associated sets have non-empty intersection.

We will need the following straightforward extension of Ramsey's theorem, cf. Exercise 5 of Lecture 1.

Theorem 31. *For every graph H and $r \in \mathbb{N}$, there exists a $\delta > 0$ such that the following holds for all sufficiently large $n \in \mathbb{N}$. If G is a graph with n vertices and $e(G) > (1/2 - \delta)n^2$ edges, then every $[r]$-colouring of G contains at least $\delta n^{v(H)}$ monochromatic copies of H.*

Now, to prove Theorem 29 we simply run the usual proof for each independent set $I \in \mathcal{I}(\mathcal{H})$ (that is, for every H-free $[r]$-colouring of a graph on n vertices), applying the hypergraph container lemma repeatedly to produce subsets $S_t \subset I$ and $C_t = C_t(S_t) \supset I$, and placing C_t in our collection of containers as soon as the graph G (of edges that receive at least one colour in C_t) satisfies $e(G) \leqslant (1/2 - \delta)n^2$. Doing so, we obtain the following container theorem for H-free colourings.

Theorem 32 (A container theorem for H-free colourings)**.** *For every graph H, and every $r \in \mathbb{N}$ and $\varepsilon > 0$, there exists $\delta > 0$ and $C > 0$ such that the following holds. For every $n \in \mathbb{N}$, there exists a collection $\mathcal{G}$ of $[r]$-coloured graphs on n vertices, and a function $f\colon \mathcal{P}\big(E(K_n)\big)^r \to \mathcal{G}$ such that:*

(a) *Each $G \in \mathcal{G}$ has at most $(1-\delta)\binom{n}{2}$ edges.*

(b) *For every H-free $[r]$-coloured graph I on n vertices, there exists a sub-colouring*[13] *$S \subset I$ with*

$$e(S) \leqslant C n^{2-1/m_2(H)} \qquad \text{and} \qquad I \subset f(S).$$

The Rödl–Ruciński theorem now follows easily.

Proof of Theorem 29. Let $\mathcal{S}$ denote the collection of fingerprints S obtained in Theorem 32(b). If there exists an r-colouring I of the edges of $G(n,p)$ containing no monochromatic copy of H, it follows that $E(S) \subset E(G(n,p))$ and $E(G(n,p)) \subset E(f(S))$, and also that $e(f(S)) \leqslant (1-\delta)\binom{n}{2}$, since $f(S) \in \mathcal{G}$. By the union bound, it follows that this has probability at most

$$\begin{aligned}\sum_{S \in \mathcal{S}} p^{e(S)} (1-p)^{\delta\binom{n}{2}} &\leqslant \sum_{s \leqslant Cn^{2-1/m_2(H)}} \binom{n^2}{s} p^s (1-p)^{\delta\binom{n}{2}} \\ &\leqslant \sum_{s \leqslant Cn^{2-1/m_2(H)}} \left(\frac{epn^2}{s}\right)^s \exp\left(-\delta p \binom{n}{2}\right) \to 0\end{aligned}$$

as $n \to \infty$, as required, since $p \gg n^{-1/m_2(H)}$. □

Note that here, and in the other applications considered in this lecture, we obtain bounds on the probability of failure / the size of the exceptional set that are exponential in pn^2 / m.

[13] This means an $[r]$-colouring $c'\colon E(S) \to \mathcal{P}(r)$ satisfying $c'(e) \subset c(e)$ for every $e \in E(S) \subset E(I)$.

Exercises

1. The following supersaturation version of Sperner's theorem was proved by Kleitman in 1968.

Theorem 33. *Let $n, k \in \mathbb{N}$, and let $\mathcal{A} \subset \mathcal{P}(n)$. If $|\mathcal{A}| \geqslant \binom{n}{n/2} + k$, then $\mathcal{A}$ contains at least $kn/2$ pairs $A, B \in \mathcal{A}$ with $A \subset B$.*

(a) By adding one edge at a time, deduce the following 'balanced' version of Theorem 33.

Theorem 34. *For every $\varepsilon > 0$, there exists $\delta > 0$ such that the following holds. If $\mathcal{A} \subset \mathcal{P}(n)$ satisfies $|\mathcal{A}| \geqslant (1+\varepsilon)\binom{n}{n/2}$, then there exist a collection $\mathcal{H}$ of $\delta^2 n\binom{n}{n/2}$ pairs $A, B \in \mathcal{A}$ with $A \subset B$, such that each set $A \in \mathcal{A}$ is contained in at most δn elements of $\mathcal{H}$.*

(b) Use Theorem 34 and the graph container lemma to prove the following theorem, which is also due to Kleitman [38].[14]

Theorem 35. *There are $2^{(1+o(1))\binom{n}{n/2}}$ antichains in $\mathcal{P}(n)$.*

(c) Let $\mathcal{P}(n,p)$ denote the random set-system obtained by including each set $A \in \mathcal{P}(n)$ independently at random with probability p. Prove the following conjecture[15] of Osthus from 2000.

Theorem 36. *If $p \gg 1/n$, then with high probability the largest antichain in $\mathcal{P}(n,p)$ has size*

$$\big(1+o(1)\big)p\binom{n}{n/2}.$$

(d) What do you expect to happen for smaller values of p? Can you prove it?

2. (a) Prove the following supersaturation theorem for pairs with equal sums.

Theorem 37. *There exist constants $C > 0$ and $\delta > 0$ such that the following holds. If $A \subset [n]$ is a set of size $|A| \geqslant C\sqrt{n}$, then A contains at least $\delta m^4/n$ sets $\{x,y,z,w\}$ with $x+y = z+w$.*

(b) A set $A \subset \mathbb{Z}$ is said to be a *Sidon set* if it contains no solutions of the equation $x + y = z + w$ with $\{x,y\} \neq \{z,w\}$. Deduce the following theorem of Kohayakawa, Lee, Rödl and Samotij, and also (independently) Saxton and Thomason.

Theorem 38. *There are $2^{O(\sqrt{n})}$ Sidon sets in $\{1,\ldots,n\}$.*

(c) Use similar ideas to prove the following theorem about C_4-free graphs.[16]

Theorem 39. *There are $2^{O(n^{3/2})}$ C_4-free graphs on n vertices.*

(d) What bounds can you prove on $\mathrm{ex}\big(G(n,p), C_4\big)$?

[14] This proof appears in the survey by Balogh, Treglown and Wagner [19], which contains a number of further applications to problems involving $\mathcal{P}(n)$.

[15] This conjecture was proved by Balogh, Mycroft and Treglown, and independently by Collares and Morris.

[16] This theorem was in fact the original application of the container method, by Kleitman and Winston [40].

3. (a) Observe that a graph whose vertex set can be partitioned into a clique and an independent set[17] contains no induced copy of C_4. The following 'supersaturated stability theorem' for induced C_4s in coloured graphs was proved by Keevash and Lochet.

Theorem 40. *For every $\varepsilon > 0$ there exists $\delta > 0$ such that the following holds. Suppose G_R and G_B are 'red' and 'blue' graphs on n vertices, and suppose that*

$$e(G_R) + e(G_B) \geqslant \left(\frac{3}{4} - \delta\right)n^2.$$

Either there are at least δn^4 copies of C_4 whose edges are in G_R and whose non-edges are in G_B, or $G_R \triangle G_B$ is εn^2-close to being two disjoint copies of $K_{n/2}$, one of each colour.

(b) Use Theorem 40 and the hypergraph containers lemma to prove the following theorem of Prömel and Steger from 1991.

Theorem 41. *Almost all induced-C_4-free graphs on n vertices are εn^2-close to being a split graph.*

(c) For which $p = p(n)$ do you expect $G(n,p)$, conditioned to contain no induced copy of C_4, to be close to a split graph?

4. (a) Kuhn, Osthus, Townsend and Zhao [44] proved the following supersaturated stability theorem for transitive triangles in directed graphs. Given a directed graph G, let us write $e_1(G)$ and $e_2(G)$ for the number of single and double edges respectively.

Theorem 42. *For $\varepsilon > 0$, there exists $\delta > 0$ such that the following holds. If G is an directed graph on n vertices and*

$$e_1(G) + \log_2 3 \cdot e_2(G) \geqslant \left(\frac{\log_2 3}{4} - \delta\right)n^2,$$

then either G contains at least δn^3 transitive triangles, or is εn^2-close to being bipartite.

(b) Use hypergraph containers to deduce the following theorem, proved in [44].

Theorem 43. *Almost all oriented graphs on n vertices with no transitive triangle are $o(n^2)$-close to being bipartite.*

(c) For which functions $m = m(n)$ can you prove that almost all oriented graphs with n vertices, m edges and no transitive triangle are $o(m)$-close to being bipartite?

5. (a) Prove the following classical supersaturation theorem of Erdős and Simonovits.

Theorem 44. *For every $r \in \mathbb{N}$ and $\varepsilon > 0$, and every r-uniform hypergraph H, there exists $\delta > 0$ such that the following holds. If G is an r-uniform hypergraph with at least $\mathrm{ex}(n,H) + \varepsilon n^r$ edges, then G contains at least δn^r copies of H.*

(b) Use hypergraph containers to deduce the following theorem of Nagle, Rödl and Schacht.

Theorem 45. *If H is an r-uniform hypergraph, then the number of H-free r-uniform hypergraphs on n vertices is $2^{\mathrm{ex}(n,H)+o(n^r)}$.*

[17] A graph with this property is usually called a 'split graph'.

Lectures 3 and 4: Some more advanced applications

In the final two lectures, we will see a selection of more complicated applications of the hypergraph container method. We begin with the original application of Kleitman and Winston [40], and its recent generalisations by Balogh and Samotij [18] and Morris and Saxton [46].

$K_{s,t}$-free and C_{2k}-free graphs. The following theorems were proved by Balogh and Samotij (in 2011), and by Morris and Saxton (in 2016), respectively.

Theorem 46. *There are at most $2^{O(n^{2-1/s})}$ $K_{s,t}$-free graphs on n vertices.*

Theorem 47. *There are at most $2^{O(n^{1+1/k})}$ C_{2k}-free graphs on n vertices.*

Both theorems follow from the container lemma (in fact, the proof of Balogh and Samotij was an important step in the development of the hypergraph container method), together with a suitable 'balanced' supersaturation theorem. In order to illustrate the method while avoiding too many technical difficulties, we'll focus of the case of C_4-free graphs.

Theorem 48. *There exist constants $\delta > 0$ and $k_0 \in \mathbb{N}$ such that the following holds for every $k \geqslant k_0$ and every $n \in \mathbb{N}$. Given a graph G with n vertices and $kn^{3/2}$ edges, there exists a collection $\mathcal{H}$ of copies of C_4 in G, satisfying:*

- (a) $|\mathcal{H}| \geqslant \delta^3 k^4 n^2$,
- (b) *Each edge is used in at most $\delta^2 k^3 \sqrt{n}$ members of $\mathcal{H}$,*
- (c) *Each pair of edges is used in at most $\delta k \sqrt{n}$ members of $\mathcal{H}$.*

We will give the proof of Theorem 48, since it provides a template for similar proofs in various other settings, where one would like to 'balance' a supersaturation result.

Proof of balanced supersaturation for C_4s. Observe first that it is easy to find a collection $\mathcal{H}$ satisfying part (a): following Erdős' argument from 1935, we simply count paths of length two and use convexity. To be precise, since G has at least $kn^{3/2}$ edges, it follows that G contains at least

$$\sum_{v \in V(G)} \binom{d(v)}{2} \geqslant n \cdot \binom{2k\sqrt{n}}{2} \geqslant k^2 n^2$$

paths of length two, and hence at least

$$\frac{1}{2}\binom{n}{2}\binom{2k^2}{2} \geqslant \frac{k^4 n^2}{4}$$

C_4s. In order to find a collection which also satisfies (b) and (c), we apply this argument many times, at each step finding one new copy of C_4 that may be added to the current collection without violating the bounds in (b) and (c). In fact, for technical reasons we will need to 'colour' the two sides of each $K_{2,2} = C_4$, and maintain a different bound for P_2s centred on each side.

Indeed, suppose we have already found a collection $\mathcal{H}$ of 'coloured'[18] copies of $C_4 = K_{2,2}$ in G which satisfies condition (b) and the following condition (c'), but not (a):

[18]This just means when we find a copy of $K_{2,2}$, we label two non-adjacent vertices 'red' and the other two 'blue'.

(c') Each path of length two with centre vertex v is used in at most δk^2 members of $\mathcal{H}$ in which v is red, and at most $\delta k\sqrt{n}$ members of $\mathcal{H}$ in which v is blue.

(Note that, for example, the empty collection has this property.) The first step is to remove from G all 'saturated' edges, i.e., all edges that are already used in $\lfloor \delta^2 k^3 \sqrt{n} \rfloor$ members of $\mathcal{H}$. Since we have $|\mathcal{H}| < \delta^3 k^4 n^2$ (as otherwise we are already done), it follows that we remove at most

$$\frac{4\delta^3 k^4 n^2}{\lfloor \delta^2 k^3 \sqrt{n} \rfloor} \leqslant 5\delta k n^{3/2}$$

edges of G, and thus we still have at least $kn^{3/2}/2$ edges. Let us call the remaining edges G'.

Claim: There exist in G' at least $2\delta k^4 n^2$ coloured copies of $K_{2,2} = C_4$, such that if any one of them is added to $\mathcal{H}$, the resulting hypergraph still satisfies conditions (b) and (c').

Note that if we can prove the claim then we will be done, since at most $\delta k^4 n^2$ of the C_4s are already members of $\mathcal{H}$, so at least one of them must be new, and by the claim we can add it to $\mathcal{H}$ without violating (b) or (c').

Proof of Claim. To find such a collection of C_4s, we count paths of length two, as before, removing those that are already 'saturated'. More precisely, let us say that a path of length two whose centre vertex v is red (respectively blue) is *saturated* if it is contained in exactly $\lfloor \delta k^2 \rfloor$ (resp. $\lfloor \delta k \sqrt{n} \rfloor$) members of $\mathcal{H}$ in which v is red (resp. blue). We first find $k^4 n^2/4$ paths of length two (whose centre points we colour red) in G', none of which is saturated. To do so note that each edge is in at most

$$\frac{2\delta^2 k^3 \sqrt{n}}{\lfloor \delta k^2 \rfloor} \leqslant 3\delta k \sqrt{n}$$

saturated paths of length two with centre vertex red, since each edge $e \in E(G')$ is used in at most $\delta^2 k^3 \sqrt{n}$ members of $\mathcal{H}$, each such path containing e is in at least $\lfloor \delta k^2 \rfloor$ of them, and each C_4 contains at most two of these paths. By convexity, it follows that G' contains at least

$$\frac{1}{2} \sum_{v \in V(G')} d(v)\big(d(v) - 3\delta k \sqrt{n}\big) \geqslant \frac{1}{3n} \Bigg(\sum_{v \in V(G')} d(v) \Bigg)^2 \geqslant \frac{k^2 n^2}{4}$$

non-saturated paths of length two with centre vertex red, as claimed.

Now, fix two vertices u and w (which we will colour blue), and consider the collection $P(u, w)$ of such paths with leaves u and w. Let $\{uv, vw\} \in P(u, w)$ and consider the set of v' such that $\{uv', v'w\} \in P(u, w)$, for which the C_4 made by joining these two paths cannot be included in our collection. This implies that either $\{vu, uv'\}$ or $\{vw, wv'\}$ is saturated (with centre vertex blue), and hence each is used in at least $\lfloor \delta k \sqrt{n} \rfloor$ members of $\mathcal{H}$. However, each edge of G' is used in at most $\delta^2 k^3 \sqrt{n}$ members of $\mathcal{H}$, so there are at most

$$\frac{2\delta^2 k^3 \sqrt{n}}{\lfloor \delta k \sqrt{n} \rfloor} \leqslant 3\delta k^2$$

such vertices v'. By convexity, it follows that G' contains at least

$$\frac{1}{2} \binom{n}{2} \frac{k^4}{4} \left(\frac{k^4}{4} - 3\delta k^2 \right) \geqslant 2\delta k^4 n^2$$

C_4s in G' containing no saturated path of length two, as required. $\square$

As noted above, the theorem follows from the claim. $\square$

We will next use Theorem 48 and the hypergraph container lemma to prove the following container theorem for C_4-free graphs, which immediately implies the claimed bound on the number of C_4-free graphs, and which will moreover allow us to bound $\mathrm{ex}\big(G(n,p), C_4\big)$.

Theorem 49. *There exist constants $k_0 > 0$ and $C > 0$ such that the following holds for all $n \in \mathbb{N}$ and $k_0 \leqslant k \leqslant n^{1/6}/\log n$. There exists a collection $\mathcal{G}(n,k)$ of at most*

$$\exp\left(\frac{C\log k}{k} \cdot n^{3/2}\right)$$

graphs on n vertices such that

$$e(G) \leqslant kn^{3/2}$$

for every $G \in \mathcal{G}(n,k)$, and every C_4-free graph on n vertices is a subgraph of some $G \in \mathcal{G}(n,k)$.

Proof. As usual, we will apply the hypergraph container lemma repeatedly, each time refining the set of containers obtained at the previous step. However, it will be important that the value of τ varies, depending on the current size of the container, so that the later steps (when the container is small) matter much more than the early steps (when it is still big).

More precisely, suppose that after t steps we have constructed a family $\mathcal{C}_t$ such that

$$|\mathcal{C}_t| \leqslant \exp\left(\frac{n^{3/2}}{\varepsilon}\sum_{i=1}^{t}\max\left\{\frac{\log k(i)}{k(i)}, \frac{\log n}{n^{1/6}}\right\}\right),$$

$e(G) \leqslant k(t)n^{1+1/\ell}$ for every $G \in \mathcal{C}_t$, and every C_4-free graph is a subgraph of some $G \in \mathcal{C}_t$, where

$$k(i) = \max\left\{(1-\varepsilon)^i\sqrt{n}, k_0\right\}$$

and k_0 and ε are sufficiently large and small constants respectively. (Note that $\mathcal{C}_0 = \{K_n\}$ satisfies these conditions.) We will construct a family $\mathcal{C}_{t+1}$ by applying Theorem 48 and the hypergraph container lemma to each graph $G \in \mathcal{C}_t$ with more than $k(t+1)n^{1+1/\ell}$ edges.

Let $G \in \mathcal{C}_t$ be a graph with $kn^{3/2}$ edges, where $k \geqslant k_0$. We claim that there exists a collection $\mathcal{C}(G)$ of at most

$$\exp\left(\frac{n^{3/2}}{\varepsilon} \cdot \max\left\{\frac{\log k}{k}, \frac{\log n}{n^{1/6}}\right\}\right)$$

subgraphs of G such that:

(*a*) Every C_4-free subgraph of G is a subgraph of some $C \in \mathcal{C}(G)$, and

(*b*) $e(C) \leqslant (1-\varepsilon)e(G)$ for every $C \in \mathcal{C}(G)$.

Indeed, this follows easily by applying the container lemma to the 4-uniform hypergraph $\mathcal{H}$ with vertex set $E(G)$ and edge set given by the collection of copies of C_4 given by Theorem 48. Observe that

$$v(\mathcal{H}) = kn^{3/2}, \quad e(\mathcal{H}) \geqslant \delta k^4 n^2, \quad \Delta_1(\mathcal{H}) \leqslant k^3\sqrt{n} \quad \text{and} \quad \Delta_2(\mathcal{H}) \leqslant \delta^{-1}k\sqrt{n}.$$

Since $\Delta_3(\mathcal{H}) = \Delta_4(\mathcal{H}) = 1$, it follows that (3) holds with $c = \delta^{-1}$ and

$$\tau = \max\left\{\frac{\delta}{k^2}, \frac{1}{kn^{1/6}}\right\}.$$

Hence, by the hypergraph container lemma, there exists a collection of at most

$$\binom{e(G)}{\tau e(G)} \leqslant \left(\frac{O(1)}{\tau}\right)^{kn^{3/2}} \leqslant \exp\left(\frac{n^{3/2}}{\varepsilon} \cdot \max\left\{\frac{\log k}{k}, \frac{\log n}{n^{1/6}}\right\}\right),$$

as required. Now simply set $\mathcal{C}_{t+1} = \bigcup_{G \in \mathcal{C}_t} \mathcal{C}(G)$, and observe that $\mathcal{C}_{t+1}$ satisfies the required conditions.

Finally, let us show that if $k \leqslant n^{1/6}/\log n$ and m is chosen to be minimal so that $k(m) \leqslant k$, then

$$|\mathcal{C}_m| \leqslant \exp\left(O(1) \cdot \frac{\log k}{k} \cdot n^{3/2}\right)$$

as required. To see this, note first that $m = O(\log n)$, and that

$$\frac{(\log n)^2}{n^{1/6}} = O\left(\frac{\log k}{k}\right),$$

by our upper bound on k. Since $k(i)$ decreases exponentially in i, it follows that

$$\sum_{i=1}^{m} \max\left\{\frac{\log k(i)}{k(i)}, \frac{\log n}{n^{1/6}}\right\} = O\left(\frac{\log k}{k}\right),$$

as claimed, and so the theorem follows. □

As noted above, we can now bound the number of C_4-free graphs simply by choosing k to be a suitably large constant. We also obtain the following bound on $\mathrm{ex}\big(G(n,p), C_4\big)$, which is not far from best possible.

Theorem 50. *For every $\ell \geqslant 2$, and every function $p = p(n) \gg n^{-1/3}(\log n)^3$,*

$$\mathrm{ex}\big(G(n,p), C_4\big) \leqslant p^{1/2} n^{3/2} \log n$$

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

Proof. Choose $k \in \mathbb{N}$ so that

$$\frac{\log k}{k^2} = p \gg \frac{(\log n)^3}{n^{1/3}},$$

and note that this implies that $k \leqslant n^{1/6}/\log n$. Let $\mathcal{G}(n,k)$ be the collection of graphs given by Theorem 49. Observe that, if there exists a C_4-free subgraph of $G(n,p)$ with m edges, then some graph in $\mathcal{G}(n,k)$ must contain at least m edges of $G(n,p)$. If , then we may apply Theorem 49 and deduce that the expected number of such graphs is at most

$$\exp\left(\frac{C \log k}{k} \cdot n^{3/2}\right) \cdot \binom{kn^{3/2}}{m} \cdot p^m \leqslant \left(\frac{O\big(pkn^{3/2}\big)}{m}\right)^m \to 0$$

as $n \to \infty$, if

$$m \gg pkn^{3/2} = \frac{\log k}{k} \cdot n^{3/2}.$$

Since this inequality holds if $m \geqslant p^{1/2} n^{3/2} \log n$, the result follows. □

To see that Theorem 50 is close to best possible (apart for the log factors), consider the following two constructions:

1. Choose a random subgraph $G \subset G(n,p)$ by retaining each edge independently with probability
$$q := \min\big\{\varepsilon p^{-1} n^{-2/3},\, 1\big\},$$
and remove one edge from each copy of C_4 in G.
2. Let G be a C_4-free bipartite graph on $N = n/a$ vertices, where $a = \varepsilon/p$, with at least $\varepsilon N^{3/2}$ edges. Now blow up each vertex of G into a set of size a, and retain from $G(n,p)$ a maximal matching between each pair that corresponds to an edge of G.

We leave it to the reader to check that these constructions prove almost lower bounds in the following theorem. The upper bounds follow by using the full power of the container lemma, see [46].

Theorem 51. *There exists a constant $C > 0$ such that, with high probability,*
$$n^{4/3} \leqslant \mathrm{ex}\big(G(n,p), C_4\big) \leqslant n^{4/3} (\log n)^2$$
if $n^{-2/3} \ll p \leqslant n^{-1/3} (\log n)^4$, and
$$\mathrm{ex}\big(G(n,p), C_4\big) = \Theta\big(p^{1/2} n^{3/2}\big)$$
if $p \geqslant n^{-1/3} (\log n)^4$.

We remark that the lower bound of $n^{4/3}$ is not sharp, as Kohayakawa, Kreuter and Steger showed that, with high probability,
$$\mathrm{ex}\big(G(n,p), C_4\big) \;=\; \Theta\Big(n^{4/3} (\log \alpha)^{1/3}\Big) \tag{4}$$
if $p = \alpha n^{-4/3}$ and $2 \leqslant \alpha \leqslant n^{1/9}$.

We remark that the proof above can be adapted to count $K_{s,t}$-free graphs without great difficulty. For longer even cycles, on the other hand, the proof in [46] of the corresponding supersaturation theorem is significantly more complicated, and for general bipartite graphs we have no idea how to prove a suitable supersaturation theorem. In [46] we made the following general conjecture, and showed that (for each H) it implies that the number of H-free graphs on n vertices is $2^{O(\mathrm{ex}(n,H))}$.

Conjecture 52 (Balanced supersaturation conjecture for general bipartite H)**.** *Given a bipartite graph H, there exist constants $C > 0$, $\varepsilon > 0$ and $k_0 \in \mathbb{N}$ such that the following holds. Let $k \geqslant k_0$, and suppose that G is a graph on n vertices with $k \cdot \mathrm{ex}(n,H)$ edges. Then there exists a (non-empty) collection $\mathcal{H}$ of copies of H in G, satisfying*
$$d_{\mathcal{H}}(\sigma) \leqslant \frac{C \cdot |\mathcal{H}|}{k^{(1+\varepsilon)(|\sigma|-1)} e(G)} \quad \text{for every } \sigma \subset E(G) \text{ with } 1 \leqslant |\sigma| \leqslant e(H), \tag{5}$$
where $d_{\mathcal{H}}(\sigma) = |\{A \in \mathcal{H} : \sigma \subset A\}|$ denotes the 'degree' of the set σ in $\mathcal{H}$.

Finally, we remark that similar results were proved in the closely-related setting of "B_h-sets" by Dellamonica, Kohayakawa, Lee, Rödl and Samotij [27, 28, 29] using only graph containers (together with a number of new ideas). It would be very interesting to apply their techniques to graphs.

The KŁR conjecture. In this section we will show how to use the container method to prove the following famous conjecture of Kohayakawa, Łuczak, and Rödl [43]. Let us write $\mathcal{G}(H, n, m, p, \varepsilon)$ for the collection of all graphs G constructed as follows:

(a) $V(G) = V_1 \cup \cdots V_{v(H)}$, where the V_i are disjoint sets of size n.

(b) For each $ij \in E(H)$, choose an (ε, p)-regular[19] bipartite graph with m edges between V_i and V_j. The edge set of G is the union of these graphs.

The following theorem can be though of as a probabilistic embedding lemma for sparse graphs.

Theorem 53 (The KŁR conjecture). *For every graph H and every $\beta > 0$, there exist constants $C > 0$, $n_0 \in \mathbb{N}$ and $\varepsilon > 0$ such that the following holds. For every $n \in \mathbb{N}$ with $n \geqslant n_0$ and $m \in \mathbb{N}$ with $m \geqslant Cn^{2-1/m_2(H)}$, there are at most*

$$\beta^m \binom{n^2}{m}^{e(H)}$$

H-free graphs in $\mathcal{G}(H, n, m, p, \varepsilon)$.

We remark that Łuczak observed that, somewhat surprisingly, for any graph H that contains a cycle and any function p satisfying $p = o(1)$, there are graphs in $\mathcal{G}(H, n, pn^2, p, \varepsilon)$ with no canonical copy of H. To prove the KŁR conjecture, we will (as usual) combine the hypergraph container lemma with a suitable supersaturation lemma, which in this case is essentially just the usual embedding lemma for dense graphs. To state the version we need, let us write $\mathcal{G}(H; n_1, \ldots, n_{v(H)})$ the collection of all graphs G with vertex set $V_1 \cup \ldots \cup V_{v(H)}$, where the V_i are disjoint and $|V_i| = n_i$, and all edges of G lie between those pairs of sets (V_i, V_j) such that $\{i, j\}$ is an edge of H. The proof of the following lemma follows from the usual proof of the embedding lemma; we leave the details to the reader.

Lemma 54. *Let H be a graph and let $\delta\colon (0,1] \to (0,1)$ be an arbitrary function. There exist constants $\alpha_0, \xi, n_0 > 0$ such that for every $n_1, \ldots, n_{v(H)} \geqslant n_0$ and every graph $G \in \mathcal{G}(H; n_1, \ldots, n_{v(H)})$, one of the following holds:*

(a) G contains at least $\xi n_1 \ldots n_{v(H)}$ canonical copies of H.

(b) There exist $\alpha \geqslant \alpha_0$, an edge $ij \in E(H)$, and sets $A_i \subset V_i$, $A_j \subset V_j$ such that

$$|A_i| \geqslant \alpha n_i, \qquad |A_j| \geqslant \alpha n_j \qquad \text{and} \qquad d_G(A_i, A_j) < \delta(\alpha).$$

We next show that, by choosing δ sufficiently small, we can easily count the (ε, p)-regular subgraphs of a graph with a 'hole' as in Lemma 54(b). To be precise, for each $\beta \in (0, 1)$ let us define

$$\delta(x) = \frac{1}{4e} \left(\frac{\beta}{2} \right)^{2/x^2} \tag{6}$$

for each $x \in (0, 1]$. The following lemma says that a graph G that has a hole of size αn and density at most $\delta(\alpha)$ has very few subgraphs in $\mathcal{G}(K_2, n, m, m/n^2, \varepsilon)$.

[19]This means that for every $X \subset V_i$ and $Y \subset V_j$ with $|X| \geqslant \varepsilon |V_i|$ and $|Y| \geqslant \varepsilon |V_j|$, the density $d(X, Y)$ of edges between X and X satisfies $\big| d(X, Y) - d(V_i, V_j) \big| \leqslant \varepsilon p$.

Lemma 55. *For every $\alpha_0 > 0$ and $\beta > 0$, there exists $\varepsilon > 0$ such that the following holds. Let $G \subseteq K_{n,n}$ be such that there exist subsets $A \subseteq V_1(G)$ and $B \subseteq V_2(G)$ with*

$$\min\{|A|, |B|\} \geqslant \alpha n \qquad \text{and} \qquad d_G(A, B) < \delta(\alpha)$$

for some $\alpha \in [\alpha_0, 1]$, and let $S \subseteq G$. Then, for every m with $|S|/\varepsilon \leqslant m \leqslant n^2$, there are at most

$$\beta^m \binom{n^2}{m - |S|}$$

subgraphs of $\tilde{\mathcal{G}}$ that belong to $\mathcal{G}(K_2, n, m, m/n^2, \varepsilon)$ and contain S.

The proof of Lemma 55 is just some straightforward counting, so we again leave the details to the reader. We can now easily deduce the following "KŁR container theorem" by applying the hypergraph container lemma to the $e(H)$-uniform hypergraph whose edges are the $n^{v(H)}$ canonical copies of H in the (unique) graph $G_n^* \in \mathcal{G}(H, n, n^2, 1, 1)$.

Theorem 56 (The KŁR container theorem)**.** *For every graph H and every $\beta > 0$, there exist $\alpha_0 > 0$ and $C > 0$ such that the following holds. For every $n \in \mathbb{N}$, there exists a collection $\mathcal{G} \subset \mathcal{G}(H, n, n^2, 1, 1)$, and a function $f \colon \mathcal{P}\big(E(G_n^*)\big) \to \mathcal{G}$ such that:*

(*a*) *For each $G \in \mathcal{G}$, there exist $ij \in E(H)$ and sets $A_i \subset V_i$, $A_j \subset V_j$ such that*

$$|A_i| \geqslant \alpha n_i, \qquad |A_j| \geqslant \alpha n_j \qquad \text{and} \qquad d_G(A_i, A_j) < \delta(\alpha)$$

for some $\alpha \geqslant \alpha_0$, where δ is as defined in (6)*.*

(*b*) *For every H-free graph I on n vertices, there exists a subgraph $S \subset I$ with*

$$e(S) \leqslant C n^{2 - 1/m_2(H)} \qquad \text{and} \qquad I \subset f(S).$$

We leave the proof of Theorem 56, and the deduction of the KŁR conjecture, to the reader.

References

[1] D. Achlioptas and E. Friedgut, A Sharp Threshold for k-Colorability, *Random Structures Algorithms*, **14** (1999), 63–70.

[2] N. Alon, Independent sets in regular graphs and sum-free subsets of finite groups, *Israel J. Math.*, **73** (1991), 247–256.

[3] N. Alon, J. Balogh, R. Morris and W. Samotij, Counting sum-free subsets in abelian groups, *Israel J. Math.*, **199** (2014), 309–344.

[4] N. Alon, J. Balogh, R. Morris and W. Samotij, A refinement of the Cameron-Erdős conjecture, *Proc. London Math. Soc.*, **108** (2014), 44–72.

[5] J. Balogh, B. Bollobás and M. Simonovits, On the number of graphs without forbidden subgraph, *J. Combin. Theory Ser. B*, **91** (2004), 1–24.

[6] J. Balogh, B. Bollobás and M. Simonovits, The typical structure of graphs without given excluded subgraphs, *Random Structures Algorithms*, **34** (2009), 305–318.

[7] J. Balogh, N. Bushaw, M. Collares, H. Liu, R. Morris and M. Sharifzadeh, The typical structure of graphs with no large cliques, *Combinatorica*, to appear.

[8] J. Balogh, S. Das, M. Delcourt, H. Liu and M. Sharifzadeh, The typical structure of intersecting families of discrete structures, *J. Combin. Theory, Ser. A*, **132** (2015), 224–245.

[9] J. Balogh, H. Liu, S. Petrickova and M. Sharifzadeh, The typical structure of maximal triangle-free graphs, *Forum Math., Sigma*, **3** (2015), 19pp.

[10] J. Balogh, H. Liu and M. Sharifzadeh, The number of subsets of integers with no k-term arithmetic progression, *Int. Math. Res. Not.*, to appear.

[11] J. Balogh, H. Liu, M. Sharifzadeh and A. Treglown, The number of maximal sum-free subsets of integers, *Proc. Amer. Math. Soc.*, **143** (2015), 4713–4721.

[12] J. Balogh, R. Morris and W. Samotij, Random sum-free subsets of abelian groups, *Israel J. Math.*, **199** (2014), 651–685.

[13] J. Balogh, R. Morris and W. Samotij, Independent sets in hypergraphs, *J. Amer. Math. Soc.*, **28** (2015), 669–709.

[14] J. Balogh, R. Morris, W. Samotij and L. Warnke, The typical structure of sparse K_{r+1}-free graphs, *Trans. Amer. Math. Soc.*, to appear.

[15] J. Balogh, R. Mycroft and A. Treglown, A random version of Sperner's theorem, *J. Combin. Theory, Ser. A*, **128** (2014), 104–110.

[16] J. Balogh and S. Petrickova, The number of the maximal triangle-free graphs, *Bull. London Math. Soc.*, **46** (2014), 1003–1006.

[17] J. Balogh and W. Samotij, The number of $K_{m,m}$-free graphs, *Combinatorica*, **31** (2011), 131–150.

[18] J. Balogh and W. Samotij, The number of $K_{s,t}$-free graphs, *J. London Math. Soc.*, **83** (2011), 368–388.

[19] J. Balogh, A. Treglown and A. Zs. Wagner, Applications of graph containers in the Boolean lattice, *Random Structures Algorithms*, to appear.

[20] J. Balogh and A. Zs. Wagner, Further applications of the Container Method, book chapter: Recent Trends in Combinatorics, Editors: A. Beveridge, J. R. Griggs, L. Hogben, G. Musiker, P. Tetali; Springer 2016.

[21] J. Balogh and A. Zs. Wagner, On the number of union-free families, *Israel J. Math.*, to appear.

[22] N. Bushaw, M. Collares Neto, R. Morris and P. Smith, Sharp thresholds for sum-free subsets of abelian groups of even order, *Combin. Probab. Computing*, **24** (2015), 609–640.

[23] M. Collares and R. Morris, Maximum-size antichains in random set-systems, *Random Structures Algorithms*, to appear.

[24] D. Conlon, Combinatorial theorems relative to a random set, *Proc. Int. Cong. Math.*, **4** (2014), 303–328.

[25] D. Conlon and W.T. Gowers, Combinatorial theorems in sparse random sets, *Ann. Math.*, to appear.

[26] D. Conlon, W.T. Gowers, W. Samotij and M. Schacht, On the KŁR conjecture in random graphs, *Israel J. Math.*, **203** (2014), 535–580.

[27] D. Dellamonica, Y. Kohayakawa, S. Lee, V. Rödl and W. Samotij, On the number of B_h-sets, *Combin. Prob. Computing*, **25** (2016), 108–129.

[28] D. Dellamonica, Y. Kohayakawa, S. Lee, V. Rödl and W. Samotij, The number of B_3-sets of a given cardinality, *J. Combin. Theory, Ser. A*, **142** (2016), 44–76.

[29] D. Dellamonica, Y. Kohayakawa, S. Lee, V. Rödl and W. Samotij, The number of B_h-sets of a given cardinality, submitted.

[30] B. DeMarco and J. Kahn, Mantel's theorem for random graphs, *Random Structures Algorithms*, **47** (2015), 59–72.

[31] B. DeMarco and J. Kahn, Turan's theorem for random graphs, submitted.

[32] P. Frankl and V. Rödl, Large triangle-free subgraphs in graphs without K_4, *Graphs Combin.*, **2** (1986) 135–144.

[33] E. Friedgut, V. Rödl, A. Ruciński and P. Tetali, A sharp threshold for random graphs with a monochromatic triangle in every edge coloring, *Mem. Amer. Math. Soc.*, **179** (2006), 66pp.

[34] E. Friedgut, V. Rödl and M. Schacht, Ramsey properties of random discrete structures, *Random Structures Algorithms*, **37** (2010), 407–436.

[35] R. Graham, V. Rödl and A. Ruciński, On Schur properties of random subsets of integers, *J. Number Theory*, **61** (1996), 388–408.

[36] B. Green, The Cameron-Erdős conjecture, *Bull. London Math. Soc.*, **36** (2004), 769–778.

[37] B. Green and R. Morris, Counting sets with small sumset and applications, *Combinatorica*, to appear, arXiv:1305.3079v2.

[38] D. Kleitman, On Dedekind's problem: the number of monotone boolean functions, *Proc. Amer. Math. Soc.*, **21** (1969), 677–682.

[39] D. Kleitman and D. Wilson, On the number of graphs which lack small cycles, manuscript, 1996.

[40] D. Kleitman and K. Winston, On the number of graphs without 4-cycles, *Discrete Math.*, **41** (1982), 167–172.

[41] Y. Kohayakawa, B. Kreuter and A. Steger, An extremal problem for random graphs and the number of graphs with large even-girth, *Combinatorica*, **18** (1998), 101–120.

[42] Y. Kohayakawa, T. Łuczak and V. Rödl, Arithmetic progressions of length three in subsets of a random set, *Acta Arith.*, **75** (1996), 133–163.

[43] Y. Kohayakawa, T. Łuczak and V. Rödl, On K_4-free subgraphs of random graphs, *Combinatorica*, **17** (1997), 173–213.

[44] D. Kuhn, D. Osthus, T. Townsend and Y. Zhao, On the structure of oriented graphs and digraphs with forbidden tournaments or cycles, submitted, arXiv:1404.6178.

[45] T. Łuczak, On triangle-free random graphs, *Random Structures Algorithms*, **16** (2000), 260–276.

[46] R. Morris and D. Saxton, The number of $C_{2\ell}$-free graphs, *Adv. Math.*, **298** (2016), 534–580.

[47] R. Nenadov and A. Steger, A short proof of the Random Ramsey Theorem, *Combin. Probab. Computing*, to appear.

[48] D. Osthus, H.J. Prömel and A. Taraz, For which densities are random triangle-free graphs almost surely bipartite?, *Combinatorica*, **23** (2003), 201–250.

[49] F.P. Ramsey, On a Problem of Formal Logic, *Proc. London Math. Soc.*, **30** (1930), 264–286.

[50] V. Rödl and A. Ruciński, Threshold functions for Ramsey properties, *J. Amer. Math. Soc.*, **8** (1995), 917–942.

[51] V. Rödl and A. Ruciński, Rado partition theorem for random subsets of integers, *Proc. London Math. Soc.*, **74** (1997), 481–502.

[52] V. Rödl and M. Schacht, Extremal results in random graphs, In: Erdős Centennial, *Bolyai Soc. Math. Stud.*, **25** (2013), 535–583.

[53] W. Samotij, Counting independent sets in graphs, *European J. Combin.*, **48** (2015), 5–18.

[54] A.A. Sapozhenko, On the number of independent sets in extenders, Diskret. Mat. 13 (2001), 56–62.

[55] A.A. Sapozhenko, The Cameron-Erdős Conjecture, (Russian) *Dokl. Akad. Nauk.*, **393** (2003), no. 6, 749–752.

[56] D. Saxton and A. Thomason, List colourings of regular hypergraphs, *Combin. Probab. Computing*, **21** (2011), 315–322.

[57] D. Saxton and A. Thomason, Hypergraph containers, *Inventiones Math.*, **201** (2015), 1–68.

[58] D. Saxton and A. Thomason, Simple Containers for Simple Hypergraphs, *Combin. Probab. Computing*, **25** (2016), 448–459.

[59] M. Schacht, Extremal results for random discrete structures, *Ann. Math.*, **184** (2016), 331–363.