Combinatorics, Probability and Computing (2016) 25, 172–212.
doi:10.1017/S0963548315000061

On the Chromatic Thresholds of Hypergraphs

JÓZSEF BALOGH[1†], JANE BUTTERFIELD[2‡], PING HU[1], JOHN LENZ[3§] and DHRUV MUBAYI[3¶]

[1]Department of Mathematics, University of Illinois, Urbana-Champaign, IL 61801, USA
(e-mail: jobal@math.uiuc.edu, pinghu1@math.uiuc.edu)

[2]University of Victoria, Victoria BC V8P 5C2, Canada
(e-mail: jvbutter@uvic.ca)

[3]Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL 60607, USA
(e-mail: lenz@math.uic.edu, mubayi@math.uic.edu)

Received 15 September 2011; revised 29 December 2014; first published online 6 March 2015

Let $\mathcal{F}$ be a family of r-uniform hypergraphs. The *chromatic threshold* of $\mathcal{F}$ is the infimum of all non-negative reals c such that the subfamily of $\mathcal{F}$ comprising hypergraphs H with minimum degree at least $c\binom{|V(H)|}{r-1}$ has bounded chromatic number. This parameter has a long history for graphs ($r = 2$), and in this paper we begin its systematic study for hypergraphs.

Łuczak and Thomassé recently proved that the chromatic threshold of the so-called near bipartite graphs is zero, and our main contribution is to generalize this result to r-uniform hypergraphs. For this class of hypergraphs, we also show that the exact Turán number is achieved uniquely by the complete $(r + 1)$-partite hypergraph with nearly equal part sizes. This is one of very few infinite families of non-degenerate hypergraphs whose Turán number is determined exactly. In an attempt to generalize Thomassen's result that the chromatic threshold of triangle-free graphs is $1/3$, we prove bounds for the chromatic threshold of the family of 3-uniform hypergraphs not containing $\{abc, abd, cde\}$, the so-called generalized triangle.

[†] This material is based on work partly done at University of California, San Diego, and at SZTE, Bolyai Institute, Szeged, Hungary. Research supported by NSF CAREER grant DMS-0745185, UIUC Campus Research Board Grant 11067, OTKA Grant K76099, and the Arnold O. Beckman Research Award (UIUC Campus Research Board 13039) grant. Also supported by the European Union and co-funded by the European Social Fund under the project 'Telemedicine-focused research activities on the field of Mathematics, Informatics and Medical sciences' of project number 'TÁMOP-4.2.2.A-11/1/KONV-2012-0073'.

[‡] Research partly supported by the Dr Lois M. Lackner Mathematics Fellowship and NSF grant DMS 08-38434, 'EMSW21-MCTP: Research Experience for Graduate Students'.

[§] Research partly supported by NSA grant H98230-13-1-0224.

[¶] Research supported in part by NSF grant 0969092.

In order to prove upper bounds we introduce the concept of *fibre bundles*, which can be thought of as a hypergraph analogue of directed graphs. This leads to the notion of *fibre bundle dimension*, a structural property of fibre bundles that is based on the idea of Vapnik–Chervonenkis dimension in hypergraphs. Our lower bounds follow from explicit constructions, many of which use a hypergraph analogue of the Kneser graph. Using methods from extremal set theory, we prove that these Kneser hypergraphs have unbounded chromatic number. This generalizes a result of Szemerédi for graphs and might be of independent interest. Many open problems remain.

2010 *Mathematics subject classification*: Primary 05C35
Secondary 05C65, 05C15, 05D40

1. Introduction

An *r-uniform hypergraph on n vertices* is a collection of r-subsets of V, where V is a set of n elements. If $r = 2$ then we call it a graph. The r-sets in a hypergraph are called *edges*, and the n elements of V are called *vertices*. For a hypergraph H, let $V(H)$ denote the set of vertices. We denote the set of edges by either $E(H)$ or simply H. The *chromatic number* of a hypergraph H, denoted $\chi(H)$, is the least integer k for which there exists a map $f : V(H) \to [k]$ such that if E is an edge in the hypergraph then there exist $v, u \in E$ for which $f(v) \neq f(u)$. For a vertex v in a hypergraph H we let $d(v)$ denote the number of edges in H that contain v. We let $\delta(H) = \min\{d(v) : v \in V(H)\}$, called the *minimum degree* of H.

Definition. Let $\mathcal{F}$ be a family of r-uniform hypergraphs. The *chromatic threshold* of $\mathcal{F}$ is the infimum of the values $c \geqslant 0$ such that the subfamily of $\mathcal{F}$ consisting of hypergraphs H with minimum degree at least $c\binom{|V(H)|}{r-1}$ has bounded chromatic number.

We say that F is a subhypergraph of H if there is an injection from $V(F)$ to $V(H)$ such that every edge in F gets mapped to an edge of H. Notice that if H is r-uniform for some r then this is only possible if F is also r-uniform. If F is an r-uniform hypergraph, then the family of *F-free* hypergraphs is the family of r-uniform hypergraphs that do not contain F as a (not necessarily induced) subhypergraph.

The study of the chromatic thresholds of graphs was motivated by a question of Erdős and Simonovits [7]: 'If G is non-bipartite, what bound on $\delta(G)$ forces G to contain a triangle?' This question was answered by Andrásfai, Erdős and Sós [3], who showed that the answer is $2/5|V(G)|$, and sharpness is shown by the graph obtained from C_5 by replacing each edge with a copy of $K_{n/5,n/5}$. Andrásfai, Erdős and Sós's [3] idea, that is, blowing up a small triangle-free graph to create a new graph with the same chromatic number and large minimum degree, can be generalized to show that for every k and ϵ there exists a triangle-free graph G with $\chi(G) \geqslant k$ and $\delta(G) \geqslant (1/3 - \epsilon)|V(G)|$. This led to the following conjecture: if $\delta(G) > (1/3 + \epsilon)|V(G)|$ and G is triangle-free, then $\chi(G) \leqslant k_\epsilon$, where k_ϵ is a constant depending only on ϵ.

Note that the conjecture is equivalent to the statement that the family of triangle-free graphs has chromatic threshold 1/3. The conjecture was proved by Thomassen [37]. Subsequently, there have been three more proofs of the conjecture: one by Łuczak [24]

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

using the Regularity Lemma, a result of Brandt and Thomassé [5] proving that one can take $k_\epsilon = 4$, and a recent proof by Łuczak and Thomassé [25] using the concept of Vapnik–Chervonenkis dimension (which is defined later in this paper).

For other graphs, Goddard and Lyle [15] proved that the chromatic threshold of the family of K_r-free graphs is $(2r-5)/(2r-3)$, while Thomassen [38] showed that the chromatic threshold of the family of C_{2k+1}-free graphs is zero for $k \geqslant 2$. Recently, Łuczak and Thomassé [25] gave another proof that the class of C_{2k+1}-free graphs has chromatic threshold zero for $k \geqslant 2$, as well as several other results about related families, such as Petersen graph-free graphs. The main result of Allen, Böttcher, Griffiths, Kohayakawa and Morris [1] is to determine the chromatic threshold of the family of H-free graphs for all H.

We finish this section with some definitions.

Definition. For an r-uniform hypergraph H and a set of vertices $S \subseteq V(H)$, let $H[S]$ denote the r-uniform hypergraph consisting of exactly those edges of H that are completely contained in S. We call this the hypergraph *induced by* S. A set of vertices $S \subseteq V(H)$ is called *independent* if $H[S]$ contains no edges and *strongly independent* if there is no edge of H containing at least two vertices of S. A hypergraph is s-partite if its vertex set can be partitioned into s parts, each of which is strongly independent.

If $\mathcal{H}$ is a family of r-uniform hypergraphs, then the family of *$\mathcal{H}$-free* hypergraphs is the family of r-uniform hypergraphs that contain no member of $\mathcal{H}$ as a (not necessarily induced) subgraph. For an r-uniform hypergraph H and an integer n, let $\mathrm{ex}(n,H)$ be the maximum number of edges an r-uniform hypergraph on n vertices can have while being H-free, and let

$$\pi(H) = \lim_{n\to\infty} \frac{\mathrm{ex}(n,H)}{\binom{n}{r}}.$$

We call $\pi(H)$ the *Turán density* of H.

Let $T_{r,s}(n)$ be the complete n-vertex, r-uniform, s-partite hypergraph with part sizes as equal as possible. When $s = r$, we write $T_r(n)$ for $T_{r,r}(n)$. Let $t_r(n)$ be the number of edges in $T_r(n)$; notice that

$$t_r(n) \approx \frac{r!}{r^r}\binom{n}{r}.$$

We say that an r-uniform hypergraph H is *stable* with respect to $T_r(n)$ if $\pi(H) = r!/r^r$ and for any $\epsilon > 0$ there exists some positive δ depending only on ϵ such that if G is an n-vertex, H-free, r-uniform hypergraph with at least $(1-\delta)t_r(n)$ edges, then there is a partition of $V(G)$ into $U_1, U_2, \ldots, U_r$ such that all but at most ϵn^r edges of G have exactly one vertex in each part.

Let $\mathrm{TK}^r(s)$ be the r-uniform hypergraph obtained from the complete graph K_s by enlarging each edge with $r-2$ new vertices. The *core vertices* of $\mathrm{TK}^r(s)$ are the s vertices of degree larger than one. For $s > r$, let $\mathcal{TK}^r(s)$ be the family of r-uniform hypergraphs such that there exists a set S of s vertices where each pair of vertices from S are contained together in some edge. The set S is called the set of *core vertices* of the hypergraph.

For $s \leqslant r$, let $\mathcal{TK}^r(s)$ be the family of r-uniform hypergraphs such that there exists a set S of s vertices where, for each pair of vertices $x \neq y \in S$, there exists an edge E with $E \cap S = \{x, y\}$ (the definition is different when $s \leqslant r$ so that a hypergraph consisting of a single edge is not in $\mathcal{TK}^r(s)$). It is obvious that $\mathrm{TK}^r(s) \in \mathcal{TK}^r(s)$.

2. Results

Motivated by the above results, we investigate the chromatic thresholds of the families of A-free hypergraphs for some r-uniform hypergraphs A. One of our main results concerns a generalization of cycles to hypergraphs. A *partial matching* is a hypergraph whose edges are pairwise disjoint (note that it can contain vertices that lie in no edge).

Definition. Let H be an r-uniform hypergraph. We say that H is *near r-partite* if H is not r-partite and there exists a partition $V_1 \cup \cdots \cup V_r$ of $V(H)$ such that all edges of H either cross the partition (have one vertex in each V_i) or are contained entirely in V_1, and in addition $H[V_1]$ is a partial matching. We call such a partition a *near r-partition* if it witnesses a smallest $H[V_1]$. The edges in $H[V_1]$ of a near r-partition are called the *special edges*. Say that H is *mono near r-partite* if in addition in a near r-partition $H[V_1]$ contains exactly one edge.

A hypergraph H is *connected* if, for every $x, y \in V(H)$, there exists a sequence of hyperedges $E_1, \ldots, E_t$ such that $x \in E_1$, $y \in E_t$, and $E_i \cap E_{i+1} \neq \emptyset$ for $1 \leqslant i \leqslant t-1$. Let H be an r-uniform hypergraph and let X, Y be two disjoint sets of vertices of H.

Let $C_1, \ldots, C_t$ be the components of $H|_Y$, where $H|_Y$ is the (potentially non-uniform) hypergraph $\{A \cap Y : A \in E(H)\}$ and the *components of* $H|_Y$ are the maximal connected induced subhypergraphs of $H|_Y$. The vertex set X is *partite-extendible* to Y if there exists a partition of X into r strong independent sets $X_1, \ldots, X_r$ so that, for every $1 \leqslant i \leqslant t$, there do not exist $x_1 \in X_j$ and $x_2 \in X_\ell$ for $j \neq \ell$ and two edges $E_1, E_2 \in E(C_i)$ such that $E_1 \cup \{x_1\} \in E(H)$ and $E_2 \cup \{x_2\} \in E(H)$. Informally, each component extends to at most one part of the partition of X.

Our main theorem claims that for an infinite family of hypergraphs H the chromatic threshold of the family of H-free hypergraphs is zero. We will demonstrate that this family of hypergraphs is infinite below, applying this Theorem 2.1 to a type of hypergraph cycle (see Corollary 2.5).

Theorem 2.1. *Let H be an r-uniform, near r-partite hypergraph for which there exists near r-partition $V_1, \ldots, V_r$. If every component, which may be a single vertex, of $H[V_1]$ is partite-extendible to $V_2 \cup \cdots \cup V_r$, then the chromatic threshold of the family of H-free hypergraphs is zero.*

One interesting aspect of the chromatic threshold of graphs, proved by Łuczak and Thomassé [25], is that there exist graphs G for which the chromatic threshold of the family of G-free graphs is zero while the Turán density of G is non-zero. We show that a similar phenomenon occurs in hypergraphs; for a subfamily of the hypergraphs considered in

Theorem 2.1 we in fact determine the exact extremal hypergraph (see Theorem 2.2). We prove that if a mono near r-partite hypergraph H has Turán density $r!/r^r$ and is stable with respect to $T_r(n)$ (an example of such a graph is given in Theorem 2.4), then its unique extremal hypergraph is the complete r-partite hypergraph. Similar results occur for graphs; see Simonovits [35], where for critical graphs the Erdős–Stone theorem [9] was sharpened.

Definition. Let H be an r-uniform hypergraph. We say that H is *critical* if:

- H is mono near r-partite,
- there exists a near r-partition of H whose special edge has at least $r-2$ vertices of degree one,
- H is stable with respect to $T_r(n)$.

Recall that the stability of H implies that $\pi(H) = r!/r^r$.

Theorem 2.2. *Let H be an r-uniform critical hypergraph. Then there exists some n_0 such that for $n > n_0$, $T_r(n)$ is the unique H-free hypergraph with the most edges.*

A particularly interesting critical family is one that generalizes cycles to hypergraphs.

Definition. Fix $m \geqslant 4$ and let

$$n = \begin{cases} r\left\lfloor \dfrac{m}{2} \right\rfloor + r - 1 & \text{if } m \text{ is odd,} \\ r\dfrac{m}{2} & \text{if } m \text{ is even.} \end{cases}$$

Then C_m^r is the r-uniform hypergraph with vertices $v_1, \ldots, v_n$ and edges $E_1, \ldots, E_m$ such that:

(1) each edge contains r consecutively labelled vertices, modulo m, and in particular $E_1 = \{v_1, \ldots, v_r\}$,
(2) edges E_i and E_j intersect if and only if i and j are consecutive modulo m,
(3) if i is odd and $1 < i < m$ then $|E_{i-1} \cap E_i| = r - 1$ and $|E_i \cap E_{i+1}| = 1$.
(4) if m is odd then $|E_1 \cap E_m| = 1$; if m is even then $|E_1 \cap E_m| = r - 1$.

We say that C_m^r is *odd* if m is odd, and *even* otherwise.

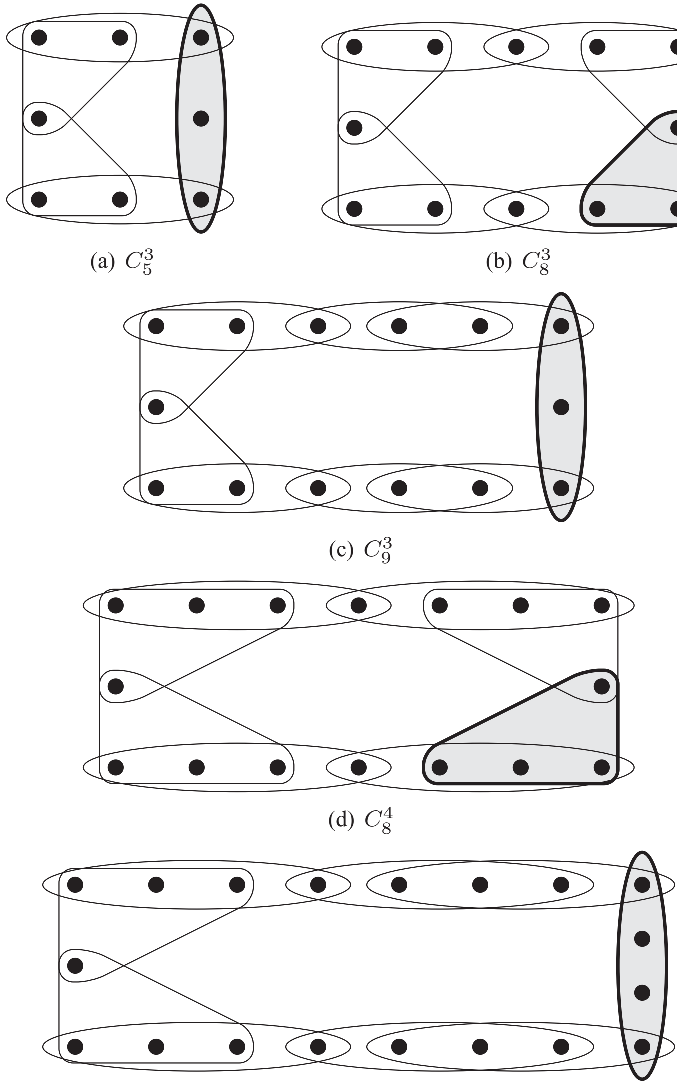
Lemma 2.3. *If $m = 2k+1 \geqslant 5$ is odd then C_m^r is not r-partite but is mono near-r-partite with partition*

$$V_1 = E_1 \cup \{v_{ir} : 1 \leqslant i \leqslant k\} \quad \textit{and} \quad V_j = \{v_{ir+j-1} : 1 \leqslant i \leqslant k+1\} \quad \textit{for } 2 \leqslant j \leqslant r.$$

Also, every component of $C_m^r[V_1]$ is partite-extendible to $V_2 \cup \cdots \cup V_r$.

Proof. Let $m = 2k+1$ for some integer k. Notice that because m is odd, we have $|E_{2k+1} \cap E_1| = 1$. Because each edge contains consecutively indexed vertices (modulo m), it follows

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press



(a) C_5^3 (b) C_8^3

(c) C_9^3

(d) C_8^4

(e) C_9^4

Figure 1. Hypergraph cycles; E_1 is indicated in each.

that v_1 is the common vertex. Then E_{2k+1} consists of the vertices $v_{rk+1}, v_{rk+2}, \ldots, v_{rk+r-1}, v_1$. Suppose $f : V \to \{0, \ldots, r-1\}$ is an r-colouring of the vertices of C^r_{2k+1} such that each colour class induces a strongly independent set. Now, $|E_1 \cap E_2| = 1$ and $|E_2 \cap E_3| = r-1$ (see Figure 2). It therefore follows that v_r is the only vertex in $E_2 \setminus E_3$ and that v_{2r} is the only vertex in $E_3 \setminus E_2$. Therefore, $f(v_r) = f(v_{2r})$. Similarly, vertices $v_r, v_{2r}, v_{3r}, \ldots, v_{kr}$ all have the same colour. Finally, $v_1 = E_m \setminus E_{m-1}$ and $v_{kr} = E_{m-1} \setminus E_m$, and so $f(v_1) = f(v_{kr})$. This shows that C^r_m is not r-partite, because $f(v_{kr}) = f(v_r)$ and v_1, v_r are in E_1. The hypergraph $C^r_m - E_1$ is r-partite via the colouring $f(v_i) = i \pmod r$. Also, all vertices of E_1 can be coloured by zero to obtain a colouring where the colour classes form a near r-partition of C^r_m.

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

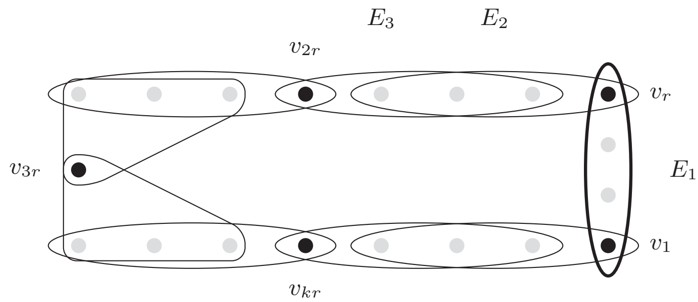


Figure 2. Odd cycles are not r-partite.

Let V_i be the vertices coloured $i-1$ for $1 \leqslant i \leqslant r$. The components of $C_m^r[V_1]$ are the edge E_1 plus the single vertex components $\{v_{ir}\}$ for $2 \leqslant i \leqslant k$. The components of $C_m^r|_{V_2 \cup \cdots \cup V_r}$ (which is a $(r-1)$-uniform hypergraph) consists of a matching. One $(r-1)$-edge of this matching is $E_2 \cap E_3$, one is $E_4 \cap E_5$, and so forth (see Figure 2). First, E_1 is partite-extendible to $V_2 \cup \cdots \cup V_r$. Indeed, only E_2 and E_{2k+1} use vertices of E_1 and they use vertices from different components of $C_m^r|_{V_2 \cup \cdots \cup V_r}$. Also, trivially each single vertex component $\{v_{ir}\}$ is partite-extendible to $V_2 \cup \cdots \cup V_r$, finishing the proof. □

A theorem of Keevash and the last author [20], combined with a theorem of Pikhurko [29], the supersaturation result of Erdős and Simonovits [8], and the hypergraph removal lemma of Gowers, Nagle, Rödl and Skokan [17, 28, 31, 32, 36] prove that C_{2k+1}^3 and C_{2k+1}^4 are critical; see Theorem 2.4.

For r larger than four, however, C_{2k+1}^r is not critical. A result of Frankl and Füredi [11] can easily be extended to prove that if $r \geqslant 5$ then

$$\pi(C_{2k+1}^r) \geqslant \frac{1}{\binom{r}{2} e^{1+1/(r-1)}} > \frac{r!}{r^r}.$$

Using techniques similar to those in Section 6, it can in fact be shown that

$$\pi(C_{2k+1}^5) = \frac{6!}{11^4} > \frac{5!}{5^5} \quad \text{and} \quad \pi(C_{2k+1}^6) = \frac{11 \cdot 6!}{12^5} > \frac{6!}{6^6}.$$

Theorem 2.4. *The cycles C_{2k+1}^3 and C_{2k+1}^4 are critical for every $k \geqslant 2$.*

Theorems 2.1, 2.2, and 2.4 together with Lemma 2.3 prove the following corollary, which extends the results in [38] and [25] that the chromatic threshold of the family of C_{2k+1}-free graphs is zero.

Corollary 2.5. *For $r = 3$ or $r = 4$ and every $k \geqslant 2$, there exists some n_0 such that for $n > n_0$, the unique n-vertex, r-uniform, C_{2k+1}^r-free hypergraph with the largest number of edges is $T_r(n)$. For all $r, k \geqslant 2$, the chromatic threshold of the family of C_{2k+1}^r-free hypergraphs is zero.*

Note that Łuczak and Thomassé [25] proved Theorem 2.1 for graphs, and they conjectured that the family of H-free graphs has chromatic threshold zero if and only if H is near acyclic and triangle free. (A graph G is *near acyclic* if there exists an independent set S in G such that $G - S$ is a forest and every odd cycle has at least two vertices in S.) This conjecture was verified by Allen, Böttcher, Griffiths, Kohayakawa and Morris [1]. We pose a similar question for hypergraphs.

Problem 2.6. *Characterize the r-uniform hypergraphs H for which the chromatic threshold of the family of H-free hypergraphs is zero.*

Note that two of the authors made some progress [4] toward solving Problem 2.6. Another way to generalize the triangle to 3-uniform hypergraphs is the hypergraph F_5, which is the hypergraph with vertex set $\{a,b,c,d,e\}$ and edges $\{a,b,c\}$, $\{a,b,d\}$, and $\{c,d,e\}$. Frankl and Füredi [10] proved that $\mathrm{ex}(n,F_5)$ is achieved by $T_3(n)$ for $n > 3000$ (recently Goldwasser [16] has determined $\mathrm{ex}(n,F_5)$ for all n). We prove the following bounds on the chromatic threshold of the family of F_5-free 3-uniform hypergraphs.

Theorem 2.7. *The chromatic threshold of the family of F_5-free 3-uniform hypergraphs is between $6/49$ and $(\sqrt{41}-5)/8 \approx 7/40$.*

The rest of the paper is organized as follows. First, in Section 3 we define and motivate fibre bundles and fibre bundle dimension, the main tools in the proofs of Theorem 2.1 and 2.7. Next, in Section 4 we show the power of fibre bundle dimension by giving a relatively short proof of Theorem 2.1. We prove our key theorem about fibre bundle dimension, Theorem 3.1, in Section 5. In Section 6, we prove that C^3_{2k+1} and C^4_{2k+1} are critical (Theorem 2.4), and then prove Theorem 2.2. The proof of Theorem 2.7 is given in Section 7. Section 9 gives lower bounds for several other families of hypergraphs, along with conjectures and open problems. The lower bounds all follow from specific constructions, some of which use a generalized Kneser hypergraph; this graph is defined and discussed in Section 8. We also make a conjecture about the chromatic number of generalized Kneser hypergraphs; see Conjecture 8.1.

Throughout this paper, we occasionally omit the floor and ceiling signs for the sake of clarity.

3. Fibre bundles and fibre bundle dimension

The proofs of Theorems 2.1 and 2.7 are based on a method by Łuczak and Thomassé [25] to colour graphs, which itself was based on the Vapnik–Chervonenkis dimension. Let H be a hypergraph. A subset X of $V(H)$ is *shattered* by H if, for every $Y \subseteq X$, there exists an $E \in H$ such that $E \cap X = Y$. Introduced in [34] and [39], the *Vapnik–Chervonenkis dimension of H* (or VC-dimension) is the maximum size of a vertex subset shattered by H.

Definition. A *fibre bundle* is a tuple (B,γ,F) such that B is a hypergraph, F is a finite set, and $\gamma : V(B) \to 2^{2^F}$. That is, γ maps vertices of B to collections of subsets of F, which we can think of as hypergraphs on vertex set F. The hypergraph B is called the *base hypergraph* of the bundle and F is the *fibre* of the bundle. For a vertex $b \in V(B)$, the hypergraph $\gamma(b)$ is called the *fibre over b*.

We should think about a fibre bundle as taking a base hypergraph and putting a hypergraph 'on top' of each base vertex. There is one canonical example of a fibre bundle. Given a hypergraph B, define the *neighbourhood bundle of* B to be the bundle (B, γ, F), where $F = V(B)$ and γ maps $b \in V(B)$ to $\{A \subseteq F : A \cup \{b\} \in E(B)\}$.

Why define and use the language of fibre bundles? We can consider that in some sense fibre bundles are a generalization of directed graphs to hypergraphs, where we think of $\gamma(x)$ as the 'out-neighbourhood' of x. In the neighbourhood bundle, $\gamma(x)$ is related to the neighbours of x, so we can consider the neighbourhood bundle as some sort of directed analogue of the undirected hypergraph B, where each edge is directed 'both ways'. By thinking of the 'out-neighbourhood' of x as $\gamma(x)$ and not requiring any dependency between $\gamma(x)$ and $\gamma(y)$ for $x \neq y$, we have no dependency between the neighbourhood of x and the neighbourhood of y, which is one of the defining differences between directed and undirected graphs. Note that the definition of a fibre bundle differs from the usual definition of *directed hypergraph* used in the literature, which is the reason we use the term 'fibre bundle' instead of 'directed hypergraph.'

A fibre bundle (B, γ, F) is (r_B, r_γ)*-uniform* if B is an r_B-uniform hypergraph and $\gamma(b)$ is an r_γ-uniform hypergraph for each $b \in V(B)$. Given $X \subseteq V(B)$, the *section of* X is the hypergraph with vertex set F and edges $\cap_{x \in X} \gamma(x)$. In other words, the section of X is the collection of subsets of F that appear in the fibre over x for every $x \in X$. Motivated by a definition of Łuczak and Thomassé [25], we define the *H-dimension of a fibre bundle*. Let H be a hypergraph and define $\dim_H(B, \gamma, F)$ to be the maximum integer d such that there exist d disjoint edges $E_1, \dots, E_d$ of B (*i.e.*, a matching) such that, for every $x_1 \in E_1, \dots, x_d \in E_d$, the section of $\{x_1, \dots, x_d\}$ contains a copy of H. Our definition of dimension coincides with the definition of paired VC-dimension in [25] when (B, γ, F) is $(2, 1)$-uniform and $H = \{\{x\}\}$, the complete 1-uniform, 1-vertex hypergraph.

Let A be an r-uniform hypergraph. Our method for proving an upper bound on the chromatic threshold of the family of A-free hypergraphs, used in Theorems 2.1 and 2.7, is the following. Let G be an A-free r-uniform hypergraph with minimum degree at least $c\binom{|V(G)|}{r-1}$. We now need to show that G has bounded chromatic number, which we do in two steps. Let (G, γ, F) be the neighbourhood bundle of G. First, we show that the dimension of (G, γ, F) is bounded, by showing that if the dimension is large then we can find A as a subhypergraph. Then, given that $\dim_H(G, \gamma, F)$ is bounded, we use the following theorem to bound the chromatic number of G. In most applications, we will let H be an $(r-1)$-uniform, $(r-1)$-partite hypergraph.

Theorem 3.1. *Let $r_B \geqslant 2$, $r_\gamma \geqslant 1$, $d \in \mathbb{Z}^+$, $0 < \epsilon < 1$, and H be an r_γ-uniform hypergraph with zero Turán density. Then there exist constants*

$$K_1 = K_1(r_B, r_\gamma, d, \epsilon, H), \quad K_2 = K_2(r_B, r_\gamma, d, \epsilon, H)$$

such that the following holds. Let (B,γ,F) *be any* (r_B,r_γ)*-uniform fibre bundle where*

$$\dim_H(B,\gamma,F) < d$$

and, for all $b \in V(B)$*,*

$$|\gamma(b)| \geqslant \epsilon\binom{|F|}{r_\gamma}.$$

If $|F| \geqslant K_1$*, then* $\chi(B) \leqslant K_2$*.*

The above theorem is sufficient for our purposes, but our proof of Theorem 3.1 proves something slightly stronger. The conclusion of the above theorem can be reworded to say that either F is small, the chromatic number of B is bounded, or $\dim_H(B,\gamma,F)$ is large, which means that we can find d hyperedges $E_1,\dots,E_d$ such that every section of $x_1 \in E_1,\dots,x_d \in E_d$ contains a copy of H. In fact, the proof shows that if F is large and the chromatic number of B is large, we can guarantee not only one copy of H but at least $\Omega(|F|^h)$ copies of H in each section, where h is the number of vertices in H.

We conjecture a similar statement for all r_γ-uniform hypergraphs H, instead of just those hypergraphs with a Turán density of zero.

Conjecture 3.2. *Let* $r_B \geqslant 2$*,* $r_\gamma \geqslant 1$*,* $d \in \mathbb{Z}^+$*,* $0 < \epsilon < 1$*, and* H *be an* r_γ*-uniform hypergraph. Then there exist constants* $K_1 = K_1(r_B,r_\gamma,d,\epsilon,H)$ *and* $K_2 = K_2(r_B,r_\gamma,d,\epsilon,H)$ *such that the following holds. Let* (B,γ,F) *be any* (r_B,r_γ)*-uniform fibre bundle where* $\dim_H(B,\gamma,F) < d$ *and, for all* $b \in V(B)$*,*

$$|\gamma(b)| \geqslant (\pi(H)+\epsilon)\binom{|F|}{r_\gamma}.$$

If $|F| \geqslant K_1$*, then* $\chi(B) \leqslant K_2$*.*

The motivation behind defining and using the language of fibre bundles rather than using the language of hypergraphs is that in the course of the proof of Theorem 3.1 we will modify B and γ and apply induction. As mentioned above, fibre bundles can be thought of as a directed version of a hypergraph. When applying Theorem 3.1 in Sections 4 and 7, we start with the neighbourhood bundle, which carries no 'extra' information beyond just the hypergraph B. But if we tried to prove Theorem 3.1 in the language of hypergraphs, we would run into trouble when we needed to modify γ. In the neighbourhood bundle, γ is related to the neighbourhood of a vertex, and if we restricted ourselves to neighbourhood bundles or just used the language of hypergraphs, modifying $\gamma(x)$ would imply that some $\gamma(y)$ would change at the same time. The notion of a fibre bundle allows us to change the 'out-neighbourhood' of x independently of changing the 'out-neighbourhood' of $y \neq x$, and this power is critical in the proof of Theorem 3.1.

4. Chromatic threshold for near r-partite hypergraphs

In this section we show an application of Theorem 3.1 by proving Theorem 2.1. Recall that H is an r-uniform, near r-partite hypergraph with near r-partition $V_1,\dots,V_r$ such that every component of $H[V_1]$ is partite-extendible to $V_2 \cup \cdots \cup V_r$. Fix $\epsilon > 0$ and let G be an n-vertex, r-uniform, H-free hypergraph with $\delta(G) \geqslant \epsilon\binom{n}{r-1}$. We would like to use Theorem 3.1 to bound the chromatic number of G, so we need to choose an appropriate bundle. We will not use the neighbourhood bundle of G, but a closely related bundle. Once we have defined this bundle, we show it has bounded dimension by proving that if the dimension is large then we can find a copy of H in G.

Proof of Theorem 2.1. Let H be an r-uniform, near r-partite, h-vertex hypergraph and let $\epsilon > 0$ be fixed. Let $V_1, \ldots, V_r$ be a near r-partition of H and assume every component of $H[V_1]$ is partite-extendible to $V_2 \cup \cdots \cup V_r$. Let

$$d = |V_1|.$$

Let G be an n-vertex, H-free hypergraph with $\delta(G) \geqslant \epsilon\binom{n}{r-1}$. We need to show that the chromatic number of G is bounded by a constant depending only on ϵ and H.

First, choose a partition $X_1, \ldots, X_r$ of $V(G)$ such that the sizes of $X_1, \ldots, X_r$ are as equal as possible, and for every $x \in V(G)$ the number of edges containing x and one vertex from each X_i is at least $(1/2r^r)\epsilon\binom{n}{r-1}$. (Almost every nearly equitable partition has this property.) We will show how to bound the chromatic number of $G[X_1]$; the same argument can be used to bound the chromatic number of each $G[X_i]$ and thus the chromatic number of G.

Define the $(r, r-1)$-uniform fibre bundle (B, γ, F) as follows. Let $B = G[X_1]$, let $F = X_2 \cup \cdots \cup X_r$, and for $x \in X_1$ define

$$\gamma(x) = \{\{x_2, \ldots, x_r\} \subseteq F : x_2 \in X_2, \ldots, x_r \in X_r, \{x, x_2, \ldots, x_r\} \in G\}.$$

Then $\gamma(x)$ has size at least $(1/2r^r)\epsilon\binom{n}{r-1}$. Let L be the complete $(r-1)$-uniform, $(r-1)$-partite hypergraph on $(rh(r-1))^{h(r-1)}$ vertices with colour classes of (nearly) equal sizes. Using that the Turán density of L is zero, we apply Theorem 3.1 to show that there exist constants $K_1 = K_1(r, \epsilon, H)$ and $K_2 = K_2(r, \epsilon, H)$ such that one of the following holds: either $|F| \leqslant K_1$, $\chi(B) \leqslant K_2$, or $\dim_L(B, \gamma, F) \geqslant d$. Since $|F| = (1 - 1/r)|V(G)|$, if $|F| \leqslant K_1$ then

$$|V(G)| < K_1\left(\frac{r}{r-1}\right).$$

Therefore, if either of the first two possibilities occur then the chromatic number of $G[X_1]$ is bounded. We may therefore assume that $\dim_L(B, \gamma, F) \geqslant d$.

We now show that if $\dim_L(B, \gamma, F) \geqslant d$ then G contains a copy of H, which follows from the definition of near r-partite and partite-extendible. Since $\dim_L(B, \gamma, F) \geqslant d$, there are d edges $E_1, \ldots, E_d$ such that, for each $x_1 \in E_1, \ldots, x_d \in E_d$, we have that $\gamma(x_1) \cap \cdots \cap \gamma(x_d)$ contains a copy of L; see Figure 3. Since $h = |V(H)|$, from each $\gamma(x_1) \cap \cdots \cap \gamma(x_d)$ we can pick a copy of the complete $(r-1)$-uniform, $(r-1)$-partite hypergraph on h vertices in each part, whose colour classes are of nearly equal size so that all these copies are vertex-disjoint. Assume $V_1 = A_1 \cup \cdots \cup A_\ell \cup \{a_{\ell+1}\} \cup \cdots \cup \{a_{\ell'}\}$, where $A_1, \ldots, A_\ell$ are the special edges of H. Because $\ell \leqslant \ell' \leqslant d$, we can embed a copy of H in G by mapping A_i to E_i for $1 \leqslant i \leqslant \ell$, mapping a_i to any vertex in E_i for $\ell + 1 \leqslant i \leqslant \ell'$, and mapping the

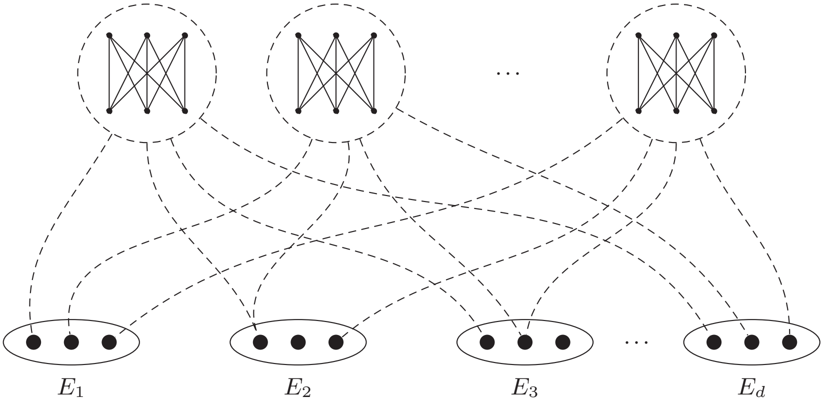


Figure 3. The structure guaranteed by dimension d.

components of $H|_{V_2\cup\cdots\cup V_r}$ to the complete $(r-1)$-uniform, $(r-1)$-partite hypergraphs as follows.

Consider some component C in $H|_{V_2\cup\cdots\cup V_r}$. Any such C is an $(r-1)$-uniform, $(r-1)$-partite hypergraph on at most h vertices. Let $D_1,\dots,D_{\ell'}$ be the components of $H[V_1]$; D_i is either one of the special edges $A_1,\dots,A_\ell$ or D_i consists only of the vertex a_i for some $\ell+1\leqslant i\leqslant \ell'$. Since $V(D_i)$ is partite-extendible to $V_2\cup\cdots\cup V_r$, edges in C, extend to at most one vertex $z_i\in D_i$. Since vertices in V_1 are embedded to vertices in $E_1,\dots,E_d$, this means that C must be embedded in $\gamma(x_1)\cap\cdots\cap\gamma(x_d)$ for some $x_i\in E_i$. It is crucial that C does not need to be embedded in $\gamma(x)\cap\gamma(y)$ for $x\neq y\in E_i$; this is what is guaranteed by the definition of partite-extendible. Embedding C is possible since $\gamma(x_1)\cap\cdots\cap\gamma(x_d)$ contains a complete $(r-1)$-uniform, $(r-1)$-partite hypergraph h vertices in each part, and $h=|V(H)|$ (so even if more than one component is embedded in the same $\gamma(x_1)\cap\cdots\cap\gamma(x_d)$, there is enough room for both of them.) □

5. Colouring hypergraphs with bounded dimension

In this section we will prove Theorem 3.1. To prove Theorem 3.1, given a fibre bundle (B,γ,F) satisfying the conditions of the theorem, we must show how to produce a proper colouring of B with a bounded number of colours. We do this via a partition refinement strategy. Below, we give an algorithm to refine a partition of (B,γ,F) (a partition is formally defined below). The algorithm will increase a density measure (also defined below) by a constant amount and add a constant number of new parts, so the refinement will halt after a constant number of iterations. Each part of the resulting partition will either correspond to an independent set in B or to a vertex set X where $B[X]$ has a maximal matching of bounded size (so $B[X]$ has bounded chromatic number), therefore producing a proper colouring of B with a bounded number of colours.

Throughout this section, fix $r_B\geqslant 2$, $r_\gamma\geqslant 1$, $d\in\mathbb{Z}^+$, $0<\epsilon<\frac{1}{4}r_B^{-d}$, and H an r_γ-uniform hypergraph with zero Turán density.

Condition 5.1. *Let (B, γ, F) be an (r_B, r_γ)-uniform fibre bundle for which* $\dim_H(B, \gamma, F) < d$, *and if $b \in V(B)$, then $|\gamma(b)| \geqslant \epsilon\binom{|F|}{r_\gamma}$.*

Define the following constants:

$$\alpha = \frac{1}{1000}\left(\frac{\epsilon}{4^{r_B^d+1}}\right)^{d+1}, \qquad \eta = \frac{1}{4}\epsilon^2\alpha, \qquad \beta = \alpha^{1/\eta}, \qquad K_2 = \lceil r_B d(r_B^d + 2)^{1/\eta}\rceil.$$

Next, pick K_1 sufficiently large that if $|F| \geqslant K_1$ and $S \subseteq \binom{F}{r_\gamma}$ with $|S| \geqslant \alpha\beta\epsilon\binom{|F|}{r_\gamma}$, then S contains a copy of H.

If (B, γ, F) is a fibre bundle, a *partition* P of (B, γ, F) is a family

$$P = \{(X_1, S_1), \ldots, (X_p, S_p)\}$$

such that $X_1, \ldots, X_p$ is a partition of $V(B)$ and $S_1, \ldots, S_p$ is a partition of $\binom{F}{r_\gamma}$, where we allow $X_i = \emptyset$ or $S_i = \emptyset$. A partition Q is a refinement of a partition P if, for each $(X, S) \in P$, there exist $(Y_1, T_1), \ldots, (Y_q, T_q) \in Q$ such that $X = \cup Y_i$ and $S = \cup T_i$. For $X \subseteq V(B)$ and $S \subseteq 2^F$, the *density of* (X, S) is

$$d(X,S) = \begin{cases} 1 & S = \emptyset \text{ or } X = \emptyset, \\ \min\left\{\dfrac{|\gamma(x) \cap S|}{|S|} : x \in X\right\} & \text{otherwise,} \end{cases}$$

and define

$$d(P) = \min\{d(X,S) : (X,S) \in P\}.$$

A partition P is a *partial colouring* if, for every $(X, \emptyset) \in P$, we have that $B[X]$ is independent. The *rank* of a partition P is the minimum of $|S|$ over all $(X, S) \in P$ with $S \neq \emptyset$.

The key lemma in this section is the following.

Lemma 5.2. *Let (B, γ, F) be a fibre bundle satisfying Condition 5.1 and $|F| \geqslant K_1$. Let $X \subseteq V(B)$ and $S \subseteq \binom{F}{r_\gamma}$ with $X \neq \emptyset$, $d(X,S) \geqslant \epsilon$, and $|S| \geqslant \beta\binom{|F|}{r_\gamma}$. Then there exists a partition $Y_1, \ldots, Y_q, Z$ of X and a partition $T_1, \ldots, T_q$ of S such that $q \leqslant r_B^d + 1$ and*

- *$|T_i| \geqslant \alpha|S|$,*
- *$d(Y_i, T_i) \geqslant \min\{1, \eta + d(X,S)\}$,*
- *$B[Z]$ is independent.*

This lemma has an easy corollary.

Corollary 5.3. *Let (B, γ, F) be a fibre bundle satisfying Condition 5.1 and $|F| \geqslant K_1$. Let P be a partial colouring of (B, γ, F) where P has rank at least $\alpha^k\binom{|F|}{r_\gamma}$ with $k \leqslant 1/\eta$. Then there exists a refinement Q of P such that*

- *$|Q| \leqslant (r_B^d + 2)|P|$,*
- *Q is also a partial colouring,*

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

- *the rank of Q is at least* $\alpha^{k+1}\binom{|F|}{r_\gamma}$,
- $d(Q) \geqslant \min\{1, \eta + d(P)\}$.

Proof. For each pair $(X,S) \in P$ with $X \neq \emptyset$ and $S \neq \emptyset$, apply Lemma 5.2. Since $k \leqslant 1/\eta$,

$$|S| \geqslant \alpha^k \binom{|F|}{r_\gamma} \geqslant \alpha^{1/\eta}\binom{|F|}{r_\gamma} \geqslant \beta\binom{|F|}{r_\gamma}.$$

Lemma 5.2 produces $Y_1,\ldots,Y_q,Z$ and $T_1,\ldots,T_q$ with $q \leqslant r_B^d + 1$. We replace the pair (X,S) with the pairs $(Y_1,T_1),\ldots,(Y_q,T_q),(Z,\emptyset)$. The resulting partition satisfies all the required properties. □

We can now easily prove Theorem 3.1.

Proof of Theorem 3.1. By assumption, (B,γ,F) satisfies Condition 5.1. Start with the partition

$$P = \left\{\left(V(B), \binom{F}{r_\gamma}\right)\right\}$$

and apply Corollary 5.3 repeatedly until the partition satisfies $d(P) = 1$. Since the value of $d(P)$ increases by η at each step, the partition is refined at most $1/\eta$ times, and so the resulting partition P has at most $(r_B^d + 2)^{1/\eta}$ parts. Consider a part $(X,S) \in P$. If $S = \emptyset$, then since P is a partial colouring $B[X]$ must be independent, so $\chi(B[X]) = 1$. If $S \neq \emptyset$, then because the partition was refined at most $1/\eta$ times we know that $|S| \geqslant \beta\binom{|F|}{r_\gamma}$, which by the choice of β and K_1 forces a copy of H in S. Since $d(X,S) = 1$ we must have $S \subseteq \gamma(x)$ for every $x \in X$, so that a matching of size d in $B[X]$ witnesses that $\dim_H(B,\gamma,F) \geqslant d$. Therefore, the maximum size of a matching in $B[X]$ is $d-1$. Since the size of a maximal matching in $B[X]$ is $d-1$, it follows that $\chi(B[X]) \leqslant r_B(d-1)+1$. This implies that the chromatic number of B is at most $r_B d(r_B^d+2)^{1/\eta}$. □

All that remains is to prove Lemma 5.2. Before proving this lemma, we make some definitions. If $E_1,\ldots,E_t \in B$ and $S \subseteq \binom{F}{r_\gamma}$, then the *minimum section density of* $E_1,\ldots,E_t$ *with respect to* S is

$$\delta(E_1,\ldots,E_t,S) = \min\left\{\frac{|\gamma(x_1)\cap\cdots\cap\gamma(x_t)\cap S|}{|S|} : x_1 \in E_1,\ldots,x_t \in E_t\right\}.$$

Notice that if $E_1,\ldots,E_d$ are disjoint, $\delta(E_1,\ldots,E_d,S) > 0$, S contains a constant fraction of $\binom{F}{r_\gamma}$, and F is large, then $E_1,\ldots,E_d$ witness that $\dim_H(B,\gamma,F) \geqslant d$. Define constants $\psi_1,\ldots,\psi_d$ recursively by $\psi_1 = 1$ and $\psi_{i+1} = \frac{1}{2}4^{-r_B^d}\epsilon\psi_i$ for $1 \leqslant i \leqslant d-1$.

Proof of Lemma 5.2. Start by greedily selecting disjoint edges $E_1,\ldots,E_i$ of $B[X]$ such that $\delta(E_1,\ldots,E_i,S) \geqslant \epsilon\psi_i$. Since for every $x \in X$

$$\frac{|\gamma(x)\cap S|}{|S|} \geqslant d(X,S) \geqslant \epsilon\psi_1,$$

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

the greedy algorithm can start with any edge E_1 in $B[X]$. Assume the greedy algorithm has selected $E_1, \ldots, E_m$ with $\delta(E_1, \ldots, E_m, S) \geqslant \epsilon\psi_m$ but for every other edge E in $B[X]$ disjoint from $E_1, \ldots, E_m$, we have $\delta(E_1, \ldots, E_m, E, S) < \epsilon\psi_{m+1}$.

First, we prove that $\dim_H(B, \gamma, F) \geqslant m$. Let $m' = \min\{m, d\}$. Since $\delta(E_1, \ldots, E_{m'}, S) \geqslant \epsilon\psi_{m'} \geqslant \epsilon\psi_d$, we have that every section of $x_1 \in E_1, \ldots, x_{m'} \in E_{m'}$ has size at least $\epsilon\psi_d|S| \geqslant \epsilon\alpha|S| \geqslant \alpha\epsilon\beta\binom{|F|}{r_\gamma}$. By the choice of K_1, the section of $x_1, \ldots, x_{m'}$ contains a copy of H, and so $m' < d$ and $m' = m$. Then $E_1, \ldots, E_m$ witness that $\dim_H(B, \gamma, F) \geqslant m$.

We make the following definitions.

- Let $R_1, \ldots, R_t$ be all r_B^m sections of $v_1 \in E_1, \ldots, v_m \in E_m$ intersected with S.
- Now remove elements from each R_i to form T_i via the following steps.
  - Start with $T_i = R_i$ for all $1 \leqslant i \leqslant t$.
  - If there exists some $i \neq j$ with $T_i \cap T_j \neq \emptyset$, divide $T_i \cap T_j$ into two sets A and B with size as equal as possible and remove A from T_i and B from T_j. Repeat until $T_1, \ldots, T_t$ are pairwise disjoint.
  - Remove elements of T_i arbitrarily until $|T_i| < 2\epsilon|S|$. (If T_i is already smaller than $2\epsilon|S|$, nothing needs to be removed.)
- Let $T_{t+1} = S \setminus T_1 \setminus \cdots \setminus T_t$.
- For $1 \leqslant i \leqslant t+1$, define
$$Y_i = \left\{ x \in X : \frac{|\gamma(x) \cap T_i|}{|T_i|} \geqslant \min\{1, \eta + d(X, S)\} \right\}.$$
  If some x appears in more than one Y_i, remove it from all but the least-indexed Y_i.
- Let $Z = X \setminus Y_1 \setminus \cdots \setminus Y_{t+1}$.

By the definition of Y_i, $d(Y_i, T_i) \geqslant \min\{1, \eta + d(X, S)\}$. Therefore, to finish the proof we need to check that $|T_i| \geqslant \alpha|S|$ and $B[Z]$ is independent.

Claim 1. $|T_i| \geqslant 2\psi_{m+1}|S| \geqslant \alpha|S|$ for all $1 \leqslant i \leqslant t+1$.

Proof. Since $\delta(E_1, \ldots, E_m, S) \geqslant \epsilon\psi_m$, each R_i has size at least $\epsilon\psi_m|S|$, so initially each T_i has size at least $\epsilon\psi_m|S|$. Now consider how many elements are removed from T_i for some fixed i. For each $j \neq i$, half of $T_i \cap T_j$ will be removed from T_i, so even if T_i is contained inside T_j, at most half of T_i will be removed. To deal with the case when $T_i \cap T_j$ is odd, certainly the size of T_i is cut down to at most one-fourth. There are $t - 1 = r_B^m - 1 \leqslant r_B^d$ of these potential removals, so after making $T_1, \ldots, T_t$ disjoint,
$$|T_i| \geqslant \frac{1}{4^{r_B^d}}|R_i| \geqslant \frac{\epsilon\psi_m}{4^{r_B^d}}|S| = 2\psi_{m+1}|S|.$$
Finally, since $\psi_1 = 1$ and $m \geqslant 1$, $\psi_{m+1} < \epsilon/4$, we have that $2\psi_{m+1}|S| < 2\epsilon|S|$, so if after making $T_1, \ldots, T_t$ disjoint, T_i is still larger than $2\epsilon|S|$, cutting T_i down to size $2\epsilon|S|$ still preserves that $|T_i| \geqslant 2\psi_{m+1}|S|$. By the choice of constants, $2\psi_{m+1} \geqslant \alpha$ so $|T_i| \geqslant \alpha|S|$.

Now consider the size of T_{t+1}. Since each T_i with $i \leqslant t$ has size at most $2\epsilon|S|$ and we assumed that $\epsilon < \frac{1}{4}t^{-1}$, the set T_{t+1} has at least $\frac{1}{2}|S| \geqslant 2\psi_{m+1}|S| \geqslant \alpha|S|$ elements. □

Claim 2. $B[Z]$ is independent.

Proof. Assume E is an edge in $B[Z]$. We would like to show that there exists some $x \in E$ and some T_j such that

$$\frac{|\gamma(x) \cap T_j|}{|T_j|} \geqslant \min\{1, \eta + d(X,S)\}, \tag{5.1}$$

since this would show that $x \in Y_j$, contradicting that $x \in Z$. Assume E intersects some E_i for some $1 \leqslant i \leqslant m$, with $x \in E \cap E_i$. Since $x \in E_i$ there is a section R_j that *selects* x, by which we mean that R_j was formed by choosing x from E_i. Fix some such section R_j that selects x, in which case $R_j \subseteq \gamma(x)$. Then $T_j \subseteq R_j \subseteq \gamma(x)$ and $|\gamma(x) \cap T_j|/|T_j| = 1$, so (5.1) is satisfied.

Now assume E is disjoint from $E_1, \dots, E_m$. Since the greedy algorithm could not continue, $\delta(E_1, \dots, E_m, E, S) < \epsilon\psi_{m+1}$, which implies that there exists some $v_1 \in E_1, \dots, v_m \in E_m, x \in E$ such that

$$|\gamma(v_1) \cap \cdots \cap \gamma(v_m) \cap \gamma(x) \cap S| < \epsilon\psi_{m+1}|S|.$$

By the definition of T_i, there exists some T_i such that $T_i \subseteq \gamma(v_1) \cap \cdots \cap \gamma(v_m) \cap S$. Therefore,

$$|\gamma(x) \cap T_i| < \epsilon\psi_{m+1}|S| \leqslant \frac{\epsilon}{2}|T_i|,$$

where the last inequality uses

$$|S| \leqslant \frac{1}{2\psi_{m+1}}|T_i|$$

from Claim 1. Assume that for every $j \neq i$, (5.1) fails. Then

$$|\gamma(x) \cap S| = |\gamma(x) \cap T_i| + \sum_{j \neq i} |\gamma(x) \cap T_j| \leqslant \frac{\epsilon}{2}|T_i| + \sum_{j \neq i} (\eta + d(X,S))|T_j|.$$

Dividing through by $|\gamma(x) \cap S|$, we obtain

$$1 \leqslant \frac{\epsilon}{2}\frac{|T_i|}{|S|}\frac{|S|}{|\gamma(x) \cap S|} + (\eta + d(X,S))\left(1 - \frac{|T_i|}{|S|}\right)\frac{|S|}{|\gamma(x) \cap S|}.$$

Because

$$|S|/|\gamma(x) \cap S| \leqslant \frac{1}{d(X,S)} \leqslant \frac{1}{\epsilon},$$

we have

$$1 \leqslant \frac{1}{2}\frac{|T_i|}{|S|} + \left(\frac{\eta}{\epsilon} + 1\right)\left(1 - \frac{|T_i|}{|S|}\right). \tag{5.2}$$

Let $w = |T_i|/|S|$. The right-hand side of the above inequality is a weighted average of $1/2$ and $(1 + \eta/\epsilon)$:

$$\frac{1}{2}w + \left(1 + \frac{\eta}{\epsilon}\right)(1 - w).$$

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

Since $1/2 < 1 + \eta/\epsilon$, this will be maximized when w is as small as possible. By Claim 1, $w \geqslant \alpha$, and we have

$$\frac{1}{2}\alpha + \left(1 + \frac{\eta}{\epsilon}\right)(1 - \alpha) < \frac{1}{2}\alpha + 1 + \frac{\eta}{\epsilon} - \alpha \leqslant 1 + \frac{\eta}{\epsilon} - \frac{1}{2}\alpha < 1.$$

This implies that for any $w \geqslant \alpha$ the inequality in (5.2) is false. This contradiction shows that there must be some $j \neq i$ such that $|\gamma(x) \cap T_j|/|T_j|$ is at least $\eta + d(X,S)$, which contradicts that E is contained in $B[Z]$. □

Thus $B[Z]$ is independent and the proof is complete. □

6. Extremal results for critical hypergraphs

In this section we prove Theorems 2.2 and 2.4. First, by Lemma 2.3, C^r_{2k+1} is mono near r-partite. Thus to complete the proof of Theorem 2.4 we need only prove that C^3_{2k+1} and C^4_{2k+1} are stable with respect to $T_3(n)$ and $T_4(n)$. One tool we will use is the hypergraph removal lemma of Gowers, Nagle, Rödl and Skokan [17, 28, 31, 32, 36].

Theorem 6.1. *For every integer $r \geqslant 2$, $\epsilon > 0$, and r-uniform hypergraph H, there exists a $\delta > 0$ such that any r-uniform hypergraph with at most $\delta n^{|V(H)|}$ copies of H can be made H-free by removing at most ϵn^r edges.* □

The second tool we will use is supersaturation, proved by Erdős and Simonovits [8]. There are several equivalent formulations of supersaturation; the one we will use is the following.

Theorem 6.2 (Corollary 2 of [8]). *Let $K^r_{t_1,\ldots,t_r}$ be the complete r-uniform, r-partite hypergraph with part sizes $t_1, \ldots, t_r$. Let $t = \sum t_i$. For every $\epsilon > 0$, there exists a $\delta = \delta(r,t,\epsilon)$ such that any r-uniform hypergraph with at least ϵn^r edges contains at least δn^t copies of $K^r_{t_1,\ldots,t_r}$.* □

For any hypergraph H, let $H(t)$ denote the hypergraph obtained from H by blowing up each vertex into an independent set of size t. An easy extension of supersaturation is the following (see Theorem 2.2 in the survey by Keevash [19]).

Corollary 6.3. *For every $r, t \geqslant 2$, $\epsilon > 0$, and r-uniform hypergraph H, there exists an n_0 such that if $n \geqslant n_0$ and G is an n-vertex, r-uniform hypergraph that contains at least $\epsilon n^{|V(H)|}$ copies of H, then G contains a copy of $H(t)$.* □

Next, we will need stability results for F_5 and the book $B_{4,2}$, proved by Keevash and the last author [20] and Pikhurko [29] respectively. Let the *book* $B_{r,m}$ be the r-uniform hypergraph with vertices $x_1, \ldots, x_{r-1}, y_1, \ldots, y_r$ and hyperedges $\{x_1, \ldots, x_{r-1}, y_i\}$ for $1 \leqslant i \leqslant m$ and $\{y_1, \ldots, y_r\}$. Note that $F_5 = B_{3,2}$.

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

Theorem 6.4 ([20]). *F_5 is stable with respect to $T_3(n)$.* □

Theorem 6.5 ([29]). *$B_{4,2}$ is stable with respect to $T_4(n)$.* □

The last piece of the proof of Theorem 2.4 is the following lemma.

Lemma 6.6. *If H is an r-uniform hypergraph that is stable with respect to $T_r(n)$ and F is a non-r-partite subhypergraph of $H(t)$ for some t, then F is also stable with respect to $T_r(n)$.*

Proof. First, $\pi(F) \geqslant r!/r^r$. Indeed, since F is non-r-partite, $T_r(n)$ is an F-free hypergraph. To complete the proof that F is stable with respect to $T_r(n)$, it is therefore enough to prove that, given $\epsilon > 0$, there exists a $\delta > 0$ such that if G is an F-free hypergraph with at least $t_r(n) - \delta n^r$ edges, then G differs from $T_r(n)$ in at most ϵn^r edges. This is enough since this implies that $\pi(F) \leqslant r!/r^r$ so $\pi(F) = r!/r^r$.

Let h denote the number of vertices in H and let $\epsilon > 0$ be fixed. We now show how to define δ. Since H is stable with respect to $T_r(n)$, there exists an $\alpha \leqslant \epsilon/2$ such that if G' has at least $t_r(n) - 2\alpha n^r$ edges and contains no copy of H, then G' differs from $T_r(n)$ in at most $\epsilon n^r/2$ edges. By Theorem 6.1, there exists $\beta = \beta(\alpha)$ such that if there are at most βn^h copies of H in G, then by deleting at most αn^r edges of G we can remove all copies of H. Lastly, choose $\delta \ll \beta$.

Now, fix some G that contains no copy of F and has at least $t_r(n) - \delta n^r$ edges. Because G contains no copy of F it contains no copy of $H(t)$. Therefore, by Corollary 6.3 there are at most βn^h copies of H in G. By Theorem 6.1, we may therefore delete αn^r edges in order to find a subhypergraph G' of G that contains no copy of H. Notice that G' has at least $t_r(n) - (\delta + \alpha)n^r$ edges, and $(\delta + \alpha) < 2\alpha$, so G' differs from $T_r(n)$ in at most $\epsilon n^r/2$ edges. Therefore, G differs from $T_r(n)$ in at most $(\alpha + \epsilon/2)n^r$ edges, and $\alpha + \epsilon/2 < \epsilon$. □

It is easy to see that C^r_{2k+1} is a non-r-partite subhypergraph of $B_{r,2}(k)$. Thus Theorem 6.4 combined with Lemma 6.6 shows that C^3_{2k+1} is stable with respect to $T_3(n)$, and similarly Theorem 6.5 combined with Lemma 6.6 shows that C^4_{2k+1} is stable with respect to $T_4(n)$, which completes the proof of Theorem 2.4.

For $r \geqslant 5$, a result of Frankl and Füredi [11] can be used to show that C^r_{2k+1} is not critical.

Lemma 6.7. *For $r \geqslant 5$ and every $k \geqslant 1$,*

$$\pi(C^r_{2k+1}) > \frac{r!}{r^r}.$$

Proof. Let $\mathcal{H}_n$ be the family of r-uniform hypergraphs H on n vertices that satisfy $|E_1 \cap E_2| \leqslant r-2$ whenever E_1 and E_2 are distinct edges of H. It is easy to check that for any $t > 0$ the blow-up $H(t)$ of H is C^r_{2k+1}-free. Therefore, $\mathrm{ex}(n, C^r_{2k+1}) \geqslant \max_{H \in \mathcal{H}_{n/t}} \{|H(t)|\}$. Frankl and Füredi [11] showed that, for $r \geqslant 7$,

$$\max_{H \in \mathcal{H}_{n/t}} \{|H(t)|\} > \frac{n^r}{r!} \frac{1}{\binom{r}{2} e^{1+1/(r-1)}}.$$

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

Thus, for $r \geqslant 7$,

$$\pi(C^r_{2k+1}) > \frac{r!}{r^r}.$$

All that remains is the case $r = 5$ or 6. Let F be an n-vertex, r-uniform hypergraph where no three edges E_1, E_2, E_3 satisfy $|E_1 \cap E_2| = r-1$ and $E_1 \Delta E_2 \subseteq E_3$. Frankl and Füredi [11] proved that if $r = 5$ then, for all such F, we have that $|E(F)| \leqslant \frac{6}{11^4} n^5$. In addition, if 11 divides n there exists a hypergraph F achieving equality. They also proved that if $r = 6$ then for all such F we have that $|E(F)| \leqslant \frac{11}{12^5} n^6$; again, if 12 divides n then there exists a hypergraph F achieving equality.

Notice that if H is the r-uniform hypergraph consisting of three hyperedges E_1, E_2, and E_3 such that $|E_1 \cap E_2| = r-1$ and $E_1 \Delta E_2 \subseteq E_3$, then C^r_{2k+1} is a subhypergraph of a blow-up of H. Using supersaturation and an argument similar to that used in the proof of Lemma 6.6, it follows that

$$\pi(C^5_{2k+1}) = \frac{6!}{11^4} > \frac{5!}{5^5} \text{ and } \pi(C^6_{2k+1}) = \frac{11 \cdot 6!}{12^5} > \frac{6!}{6^6},$$

as claimed. □

Proof of Theorem 2.2. Let H be a critical n-vertex, r-uniform hypergraph. Suppose H has h vertices and assume that E is the special edge of a near r-partition that exhibits the fact that H is critical, that is, E has at least $r-2$ vertices of degree one. Suppose G is an H-free, r-uniform, n-vertex hypergraph with $|G| \geqslant t_r(n)$. We would like to show that $G = T_r(n)$. Partition the vertices of G into parts $X_1, \ldots, X_r$ such that the number of edges with one vertex in each X_i is maximized. Let $\epsilon_1 = (2r)^{-h}$, let $\epsilon_2 = \epsilon_1/8r^3$, let $\delta = \delta(r, h, \epsilon_2)$ from Theorem 6.2, and let $\epsilon < 2^{-2r}\epsilon_1\epsilon_2\delta$. Organize r-sets of vertices into the following sets.

- Let M be the set of r-sets with one vertex in each of $X_1, \ldots, X_r$ that are not edges of G (the missing cross-edges).
- Let B be the collection of edges of G that have at least two vertices in some X_i (the bad edges).
- Let $G' = G - B + M$, so that G' is a complete r-partite hypergraph.
- Let $B_i = \{W \in B : |W \cap X_i| \geqslant 2\}$.

Since $B = \cup_i B_i$, there is some B_i that has size at least $\frac{1}{r}|B|$. Assume without loss of generality that $|B_1| \geqslant \frac{1}{r}|B|$. For $a \in X_1$, make the following definitions.

- $B_a = \{W \in B_1 : a \in W\}$.
- Let $C_{a,i}$ be the edges in B_a that have exactly two vertices in X_1 and exactly one vertex in each X_j with $j \geqslant 2$ and $j \neq i$.
- Let $D_a = B_a \setminus C_{a,2} \setminus \cdots \setminus C_{a,r}$.

First, $|B| < \epsilon n^r$ because G is stable with respect to $T_r(n)$. Also, since $|G| \geqslant t_r(n)$, the number of r-sets in M is at most the number of edges in B, so $|M| \leqslant |B| < \epsilon n^r$.

In the rest of the proof, we will assume that B is non-empty and then count the r-sets in M in several different ways. Our counting will imply that $|M| \geqslant \epsilon n^r$, and this contradiction will force $B = \emptyset$ and so $G = T_r(n)$. We will count r-sets in M by counting embeddings of

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

$H - E$ into G' that also map E to some element of B. Since G is H-free, each embedding must use at least one edge in M. Let Φ be the collection of embeddings $\phi : V(H) \to V(G')$ of $H - E$ into G', by which we mean that ϕ is an injection and, for all $F \in H \setminus E$,

$$\phi(F) = \{\phi(x) : x \in F\} \in G'.$$

We say that $\phi \in \Phi$ is W-*special* if $\phi(E) = W$ and a-*avoiding* if $a \in V(G)$ and some degree one vertex in E is mapped to a. If $W \in B$ and ϕ is W-special, then ϕ must use at least one edge of M. Call one of these edges the *missing edge of* ϕ.

Claim 1. For $\phi \in \Phi$ and $v \in V(H)$, there are at least $\frac{1}{2r}n$ embeddings $\phi' \in \Phi$, where $\phi(x) = \phi'(x)$ for $x \neq v$ and $\phi(v) \neq \phi'(v)$.

Proof. This follows easily because G' is a complete r-partite hypergraph for which each class has size about n/r, and $\phi(v)$ can be replaced by any unused vertex in the X_i that contains $\phi(v)$. □

Fix some $W \in B$, and consider when there exists a W-special embedding of $H - E$. Since $W \in B_i$ for some i, let $w_1 \neq w_2 \in W \cap X_i$. Then there exists an embedding of $H - E$ where w_1 and w_2 are used for the non-degree one vertices in the special edge of H. Since the other vertices in the special edge have degree zero in $H - E$, the vertices in the special edge can then be embedded to W. Thus for any $W \in B$, by Claim 1 there are at least $\epsilon_1 n^{h-r}$ W-special embeddings of $H - E$, since we can vary any vertex of H not in W. The situation with a-avoiding is more complicated. If $W \in C_{a,i}$, then the only choice of w_1 and w_2 that we are guaranteed to have are the two vertices in $W \cap X_1$, one of which is a. Thus in a W-special embedding, the only way we can guarantee an embedding is by mapping a vertex whose degree is not one to a. Therefore, only when $W \in D_a$ can we guarantee that there exists at least $\epsilon_1 n^{h-r}$ W-special, a-avoiding embeddings of $H - E$.

Claim 2. For every $a \in X_1$, $|D_a| \leqslant \epsilon_2 n^{r-1}$.

Proof. Assume there exists $a \in X_1$ with $|D_a| \geqslant \epsilon_2 n^{r-1}$. We count a-avoiding, W-special embeddings of $H - E$ into G' where $W \in D_a$. For each $W \in D_a$, we argued above that there are at least $\epsilon_1 n^{h-r}$ embeddings. Since $|D_a| \geqslant \epsilon_2 n^{r-1}$, the number of a-avoiding embeddings that are W-special for some $W \in D_a$ is at least $\epsilon_1\epsilon_2 n^{r-1} \cdot n^{h-r} = \epsilon_1\epsilon_2 n^{h-1}$.

Fix some $L \in M$. We want to count the number of a-avoiding embeddings that are W-special for some $W \in D_a$ and have missing edge L. An upper bound on the number of such embeddings will be the number of choices for W times the number of choices for the $h - |W \cup L|$ vertices of H mapped outside $W \cup L$. Since all these embeddings are a-avoiding, L cannot contain a. For each $0 \leqslant \ell \leqslant r$, there exists at most $\binom{r}{\ell}$ choices for the intersection between L and W, at most $n^{r-\ell-1}$ choices of $W \in D_a$ with $|W \cap L| = \ell$ (here it is crucial that $a \in W$ and $a \notin L$), and at most $n^{h-2r+\ell}$ choices for the vertices of H not in $W \cup L$. Thus each $L \in M$ is in at most $2^r n^{h-r-1}$ potential embeddings. Since there are at least $\epsilon_1\epsilon_2 n^{h-1}$ embeddings, M must have size at least $2^{-r}\epsilon_1\epsilon_2 n^r$, contradicting the choice of ϵ. □

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

Claim 3. For every $a \in X_1$ and every $2 \leqslant i \leqslant r$, $|C_{a,i}| \leqslant \epsilon_2 n^{r-1}$.

Proof. Assume there exist some a and i with $|C_{a,i}| \geqslant \epsilon_2 n^{r-1}$. The proof is similar to the proof of Claim 2, except now we cannot count a-avoiding embeddings. In the previous claim, we used the a-avoiding property to imply that the missing edge does not contain a. In this proof, we will instead guarantee that the missing edge cannot contain a by only counting embeddings that map all edges incident to $\phi^{-1}(a)$ into G.

Let v be one of the non-degree one vertices in the special edge of H, and define $H_v = \{F \in H : v \in F, F \neq E\}$, that is, all edges of H containing v that are not the special edge. Let $Z_a = \{F \in G \setminus B : a \in F\}$, that is, all cross-edges of G that contain a. We now count embeddings $\phi \in \Phi$ that are W-special for some $W \in C_{a,i}$, map v to a, and all edges of H_v are mapped to edges in Z_a. For these embeddings, since edges in H_v are mapped to edges in $Z_a \subseteq G$, the missing edge cannot contain a.

First, $|Z_a| \geqslant |C_{a,i}|$, because otherwise we could move a to X_i and increase the number of edges across the partition, and we chose the partition $X_1, \ldots, X_r$ to maximize the number of cross-edges. Let $H' = \{F - v : F \in H_v\}$ and $Z' = \{F - a : F \in Z_a\}$. Then H' and Z' are $(r-1)$-uniform, $(r-1)$-partite hypergraphs, and Z' has at least $|C_{a,i}| \geqslant \epsilon_2 n^{r-1}$ edges. Let $t = |V(H')|$. Then Theorem 6.2 shows that Z' contains at least δn^t copies of H', so there are at least $\epsilon_2 n^{r-1} \cdot \delta n^t \cdot \epsilon_1 n^{h-r-t} = \epsilon_1 \epsilon_2 \delta n^{h-1}$ embeddings of $H - E$ that are W-special for some $W \in C_{a,i}$, map v to a, and the edges in H_v are embedded into Z_a.

Now fix $L \in M$, and consider how many of these embeddings have L as their missing edge. The computation is almost the same as in the previous claim. For each ℓ_1, ℓ_2, there are $\binom{r}{\ell_1}$ choices for $L \cap W$, there are $\binom{r}{\ell_2}$ choices for $L \cap \phi(H_v)$, there are $n^{r-1-\ell_1}$ choices for W (here we use that L does not contain a), $n^{t-\ell_2}$ choices for $\phi(H_v)$, and $n^{h-2r-t+\ell_1+\ell_2}$ choices for the other vertices of H. Thus each L is in at most $2^{2r} n^{h-r-1}$ potential embeddings. Since there are at least $\epsilon_1 \epsilon_2 \delta n^{h-1}$ embeddings, M must have size at least $2^{-2r} \epsilon_1 \epsilon_2 \delta n^r$, contradicting the choice of ϵ. □

Claims 2 and 3 imply that $|B_a| < 2r\epsilon_2 n^{r-1}$ for each a. Define

$$A = \{a \in X_1 : d_M(a) \geqslant 2r^2 \epsilon_2 n^{r-1}\}.$$

As in the proofs of the previous two claims, we would like to count embeddings of $H - E$ to obtain a lower bound on $|M|$. Once again, the main difficulty is controlling how the missing edge can intersect W. If there were some W with $W \cap A = \emptyset$, then there would be few missing edges intersecting this W, which is how we will overcome this difficulty in this part of the proof.

Claim 4. There exists some $W \in B_1$ with $W \cap A = \emptyset$.

Proof. Assume that every $W \in B_1$ contains an element of A. Then $\sum_{a \in A} |B_a| \geqslant |B_1|$. Since $|B_a| < 2r\epsilon_2 n^{r-1}$ for every a, we have the following contradiction:

$$2r\epsilon_2 n^{r-1}|A| > \sum_{a \in A} |B_a| \geqslant |B_1| \geqslant \frac{1}{r}|B| \geqslant \frac{1}{r}|M| \geqslant \frac{1}{r} \sum_{a \in A} d_M(a) \geqslant \frac{2r^2 \epsilon_2}{r} n^{r-1}|A|. \qquad \square$$

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

We now complete the proof by counting the W-special embeddings whose missing edge does not intersect W. There are at least $\epsilon_1 n^{h-r}$ embeddings that are W-special by Claim 1. If at least half of these have missing edge intersecting W, then W would contain a vertex in A. Thus there are at least $(\epsilon_1/2)n^{h-r}$ W-special embeddings where the missing edge does not intersect W. Each $L \in M$ is in at most n^{h-2r} such potential embeddings, so M has at least $(\epsilon_1/2)n^r$ elements, contradicting the choice of ϵ. □

7. Chromatic threshold of F_5-free hypergraphs

7.1. An upper bound on the chromatic threshold of F_5-free graphs

In this section we prove the upper bound in Theorem 2.7. As in Section 4, we will give an upper bound on the chromatic threshold by first proving that large dimension forces a copy of F_5, and then by applying Theorem 3.1. Let (B, γ, F) be an (r_B, r_γ)-uniform fibre bundle, and make the following definition. A *cut* in (B, γ, F) is a pair (X, S) such that $X \subseteq V(B)$, $S \subseteq \binom{F}{r_\gamma}$, and if $\gamma(x) \cap S \neq \emptyset$, then $x \in X$. In other words, the fibres that intersect S come exclusively from X. A *k-cut* is a cut (X, S) with $|X| \leqslant k$. The size of a k-cut is the size of $|S|$.

We now sketch the proof of the upper bound in Theorem 2.7. Let G be an n-vertex, 3-uniform, F_5-free hypergraph with minimum degree at least $c\binom{n}{2}$. Let (G, γ, F) be the neighbourhood bundle of G, let $H = K_{q,q}$ for some large constant q (see the definition of q in the first line of the proof of Lemma 7.2), and assume $\dim_H(G, \gamma, F)$ is large. We would like to find a copy of F_5 in G. We first use the fact that $\dim_H(G, \gamma, F)$ is large to find a set U of vertices of G such that $G[U]$ has small strong independence number. We then argue that because the minimum degree is large, there must be some vertices x, y such that $N(x, y) = \{z : xyz \in G\}$ has large intersection with U. Next, we show that since $N(x, y)$ has large intersection with U and $G[U]$ has small strong independence number, there must be an edge E with at least two vertices in $N(x, y) \cap U$, which gives a copy of F_5.

The best upper bound on the chromatic threshold will come from the lowest required minimum degree needed in the above proof. The minimum degree is used above to prove that there exists some x, y with $N(x, y) \cap U$ large. If we can find a large cut (X, S) in (G, γ, F) and we make U large enough, we could remove X from U while still maintaining all the useful properties of U. Then for all $\{x, y\} \in S$, we know that $N(x, y) \cap (U - X) = \emptyset$. Since there are now fewer pairs $\{x, y\}$ in $\binom{F}{2}$ with $N(x, y) \cap (U - X) \neq \emptyset$, we can require a weaker lower bound on the minimum degree of G to find $\{x, y\}$ with $N(x, y) \cap U$ large. In other words, the larger the cut of (G, γ, S) we can find, the better upper bound on the chromatic threshold we can prove. This is encoded in the following theorem, which computes the relationship between the minimum degree and the maximum size of a k-cut.

Theorem 7.1. *Let $0 \leqslant c \leqslant 1/5$, and fix an integer k and a constant $c' > c$. Then there exists a constant $L = L(c, c', k)$ such that the following holds. Let G be an n-vertex, F_5-free hypergraph with $\delta(G) \geqslant c'\binom{n}{2}$ and let (G, γ, F) be the neighbourhood bundle of G. Assume (G, γ, F) contains a k-cut of size at least $(1 - 5c)\binom{n}{2}$. Then $\chi(G) \leqslant L$.*

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

Note that if $c = 1/5$, then $1 - 5c = 0$ and so this theorem directly proves an upper bound of $1/5$ on the chromatic threshold of F_5-free hypergraphs. The first part of the proof of Theorem 7.1 is to find a set U with small strong independence number.

Lemma 7.2. *Let $\epsilon > 0$ be fixed. Then there exists constants $d = d(\epsilon)$ and $q = q(\epsilon)$ such that the following holds. Let G be an n-vertex, 3-uniform hypergraph and let (G, γ, F) be the neighbourhood bundle of G. Let $H = K_{q,q}$ and assume $\dim_H(G, \gamma, F) \geqslant d$. Then there exists a vertex set $U \subseteq V(G)$ such that $|U| = 5d$ and any $W \subseteq U$ of size at least $(1+\epsilon)d + 100$ has the following property. There exist three vertices a_1, a_2, a_3 in U such that a_i forms an edge of $G[U]$ with two vertices of $W \setminus \{a_1, a_2, a_3\}$.*

Proof. Let $d = 100 + 100/\epsilon^2$ and $q = 3d + 2 \cdot 3^d$. Since $\dim_H(G, \gamma, F) \geqslant d$, there exists a matching $E_1, \ldots, E_d$ such that for each $w_1 \in E_1, \ldots, w_d \in E_d$ the section of $\{w_1, \ldots, w_d\}$ contains a copy of $K_{q,q}$. (See Figure 3 in Section 4 for a picture of this structure.) Since $q = 3d + 2 \cdot 3^d$, from each of these 3^d copies of $K_{q,q}$ we can pick a copy of K_2 such that each K_2 is vertex-disjoint from $E_1 \cup \cdots \cup E_d$ and all these 3^d copies of K_2 are vertex-disjoint. Now for $1 \leqslant i \leqslant d$, let $y_i z_i$ be a randomly chosen copy of K_2 (with replacement), where each of the 3^d copies of K_2 are equally likely. Let $Z = \{y_1, \ldots, y_d, z_1, \ldots, z_d\}$ and $U = Z \cup E_1 \cup \cdots \cup E_d$. With probability at most

$$\binom{d}{2} \frac{1}{3^d} < \frac{1}{4},$$

some copy of K_2 is selected more than once. To finish the proof, we just need to show that with probability at most $1/4$, for any sufficiently large subset W of U there are three edges that each contain at least two vertices of W and each have at least one vertex that is not shared by either of the other two. Indeed, in this case the union bound shows that with probability at least $1/2$, $|U| = 5d$ and any subset of size at least $(1+\epsilon)d + 100$ has this property.

Let us call a set with the above property a 'good' set, and any set not having this property a 'bad' set. Notice that any bad subset W of U contains at most $d+5$ vertices from $E_1 \cup \cdots \cup E_d$. Otherwise, there are at least three edges, say without loss of generality E_1, E_2, and E_3, from the matching that contain at least two vertices of the subset. In this case, let $a_i \in E_i \setminus W$ (or, if $E_i \setminus W = \emptyset$, let a_i be any vertex in E_i) to see that W is a good set. Similarly, any bad subset W of U contains at most $d+2$ vertices from Z. Therefore, any bad subset of U with at least $(1+\epsilon)d + 100$ vertices must have at least $\epsilon d + 90$ vertices in $E_1 \cup \cdots \cup E_d$ and at least $\epsilon d + 90$ vertices in Z. We need to prove that this occurs with small probability.

Let $x \in E_1 \cup \cdots \cup E_d$ and $1 \leqslant i \leqslant d$. We say that $\{y_i, z_i\}$ is *built from* x if $\{y_i, z_i\}$ is the copy of K_2 assigned to a section of W where $x \in W$. That is, say $x \in E_j$. Each section picks one of the three vertices of E_j and if the section picks x and $\{y_i, z_i\}$ is the edge chosen from the copy of $K_{q,q}$ chosen from this section, then we say that $\{y_i, z_i\}$ is *built from* x. For $x \in E_1 \cup \cdots \cup E_d$ and $1 \leqslant i \leqslant d$, let $A_{x,i}$ be the following event:

$$A_{x,i} : \{y_i, z_i\} \text{ is built from } x.$$

First, $\mathbb{P}[A_{x,i}] = 1/3$. Indeed, say $x \in E_j$ and note that there are 3^d copies of K_2 in total (there are three choices from each of $E_1, \ldots, E_d$ for the section) and there are 3^{d-1} copies of K_2 built from x. Therefore, when randomly picking copies of K_2, the probability that $\{y_i, z_i\}$ is built from x is exactly $1/3$.

Let

$$\mathcal{S} = \{S \subseteq E_1 \cup \cdots \cup E_d : |S| = \epsilon d \text{ and } S \text{ has at most one vertex in each } E_i\}$$

We claim that the events $A_{x,i}$ for $x \in S$ are mutually independent for every $S \in \mathcal{S}$. Indeed, fix some $Q \subseteq S$. Then

$$\mathbb{P}\left[\bigwedge_{x\in Q} A_{x,i}\right] = \frac{3^{d-|Q|}}{3^d} = \left(\frac{1}{3}\right)^{|Q|}$$

since there are $3^{d-|Q|}$ of the copies of K_2 built from x for $x \in Q$ and built on any of three vertices in the edges E_j that do not contain a vertex of Q (recall that S has at most one vertex in each E_j). Thus

$$\mathbb{P}[\wedge_{x\in Q} A_{x,i}] = \prod_{x\in Q} \mathbb{P}[A_{x,i}],$$

so that for every $S \in \mathcal{S}$ the events $A_{x,i}$ for $x \in S$ are mutually independent. Therefore,

$$\mathbb{P}\left[\bigwedge_{x\in S} \overline{A_{x,i}}\right] = \left(\frac{2}{3}\right)^{|S|}.$$

Let $B_{S,i}$ be the event

$$B_{S,i}: \text{ no edge of } G \text{ contains a vertex of } S \text{ and both } y_i \text{ and } z_i.$$

If $B_{S,i}$ holds, then for every $x \in S$ it is the case that the event $A_{x,i}$ fails since if $A_{x,i}$ holds then $\{y_i, z_i, x\} \in E(G)$. Thus

$$\mathbb{P}[B_{S,i}] \leqslant \mathbb{P}\left[\bigwedge_{x\in S} \overline{A_{x,i}}\right] = \left(\frac{2}{3}\right)^{|S|}.$$

Let $X_{S,i}$ be the indicator random variable for the event $\overline{B_{S,i}}$. For each $T \subseteq [d]$ with $|T| = \epsilon d$, let $B_{S,T}$ be the event $\sum_{i\in T} X_{S,i} \leqslant 2$. The events $B_{S,i}$ are mutually independent for $i \in T$ since the copies of K_2 were selected with replacement, so that

$$\mathbb{P}[B_{S,T}] \leqslant \left(\frac{2}{3}\right)^{|S|(|T|-2)} \binom{|T|}{2}.$$

Let $X_{S,T}$ be the indicator random variable for the event $B_{S,T}$ and let X be the sum of all indicator random variables over all $S \in \mathcal{S}$ and all $T \subseteq [d]$ with $|T| = \epsilon d$. We now have $\binom{d}{\epsilon d}$ choices for T and $3^{\epsilon d}\binom{d}{\epsilon d}$ choices for S so that

$$\mathbb{E}[X] = \sum X_{S,T} \leqslant 3^{\epsilon d}\binom{d}{\epsilon d}^2 \left(\frac{2}{3}\right)^{\epsilon d(\epsilon d-2)} \binom{\epsilon d}{2} \leqslant \left(3\left(\frac{e}{\epsilon}\right)^2 \left(\frac{2}{3}\right)^{\epsilon d-2}\right)^{\epsilon d} \frac{\epsilon^2 d^2}{2} < \frac{1}{4}.$$

By Markov's inequality, the probability that $X \geqslant 1$ is at most $1/4$, so with probability at most $1/4$, some $B_{S,T}$ holds. If W is a bad subset of U with $|W| \geqslant (1+\epsilon)d + 100$, then

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

$|W \cap Z| \geqslant \epsilon d + 90$ and $|W \cap (E_1 \cup \cdots \cup E_d)| \geqslant \epsilon d + 90$. Also, W uses at most one vertex from all but at most two pairs in Z, so there exists $T \subseteq [d]$ of size ϵn such that for $i \in T$ exactly one of y_i and z_i is in W. Since W uses at most one vertex from all but at most two edges E_i, there exists $S \subseteq W \cap (E_1 \cup \cdots \cup E_d)$ with $|S| = \epsilon d$ and $S \in \mathcal{S}$. For such S and T, the event $B_{S,T}$ holds since if not, then without loss of generality assume $1, 2, 3 \in T$ and $X_{S,1}, X_{S,2}, X_{S,3} = 1$. Then let $a_i \in \{y_i, z_i\} \setminus W$ to see that W is a good set.

Therefore, the probability that some $B_{S,T}$ holds is an upper bound for the probability that no subset of U of size at least $(1+\epsilon)d + 100$ is bad. Since the probability that some $B_{S,T}$ holds is at most $1/4$, the proof is complete. □

We can now prove Theorem 7.1.

Proof of Theorem 7.1. Pick ϵ so that $c' = (1+2\epsilon)c$ and let $d = d(\epsilon)$ and $q = q(\epsilon)$ be given by Lemma 7.2, and also assume that d is large enough that $5d\epsilon > k(1+2\epsilon)$. Suppose that if $H = K_{q,q}$ then $\dim_H(G, \gamma, F) \leqslant d$. Then by Theorem 3.1, there exist constants $K_1 = K_1(\epsilon, d, H)$ and $K_2 = K_2(\epsilon, d, H)$ (note that K_1 and K_2 depend only on c, c', k) such that either $|F| < K_1$ or $\chi(G) < K_2$. Since $|F| = |V(G)|$, this implies that $\chi(G) < \max\{K_1, K_2\}$.

We can therefore assume that $\dim_H(G, \gamma, F) \geqslant d$. Let U be the set given by Lemma 7.2. Let (X, S) be a k-cut of size at least $(1-5c)\binom{n}{2}$. Let G' be the bipartite graph with partite sets $A = U \setminus X$ and $B = \binom{V(G)}{2} \setminus S$, where $\{u, \{v, w\}\}$ is an edge in G' if and only if $\{u, v, w\}$ is an edge in G. $|A| \geqslant 5d - |X|$, so G' contains at least $(5d - |X|)\delta(G)$ edges. $|B| = \binom{n}{2} - |S|$, so there is some $x \neq y$ such that $d_{G'}(\{x, y\})$ is at least

$$\frac{(5d - |X|)\delta(G)}{\binom{n}{2} - |S|} \geqslant \frac{(5d-k)(1+2\epsilon)c\binom{n}{2}}{5c\binom{n}{2}} = \frac{(5d-k)(1+2\epsilon)}{5} > (1+\epsilon)d + 100.$$

This implies that there is some x, y with $|N(x, y) \cap U| > (1+\epsilon)d + 100$. Then by Lemma 7.2, there exist three distinct vertices a_1, a_2, a_3 in U such that a_i forms an edge E_i of $G[U]$ with two vertices of $(N(x, y) \cap U) \setminus \{a_1, a_2, a_3\}$. There exists $i \in \{1, 2, 3\}$ such that $a_i \neq x, y$. Then x, y together with E_i form a copy of F_5 in G. This contradiction completes the proof. □

7.2. Finding a large cut in an F_5-free hypergraph

In order to use Theorem 7.1 to prove the upper bound in Theorem 2.7, we now need to show the existence of a large cut. Note that in Theorem 7.1 the bound on the chromatic number depends on k but there are no other restrictions on k. Thus to prove an upper bound on the chromatic threshold of a F_5-free graph G, one can pick any fixed integer k and ask what is the size of the largest k-cut. In the following lemma, we set $k = 5$ and prove that if $\delta(G) \geqslant c'\binom{n}{2}$ with $c' > c$, then there exist a 5-cut of G of size approximately $4c^2\binom{n}{2}$. Solving $4c^2 = 1 - 5c$ gives $c = (\sqrt{41} - 5)/8$, the bound in Theorem 2.7.

We suspect that the bound on the chromatic threshold of F_5-free hypergraphs can be improved by finding a larger cut, perhaps by increasing k. In order to achieve a bound of $c = 6/49$, we would need to find a cut of size $s\binom{n}{2}$ with $s = 1 - 5c = 539/36c^2 \approx 15c^2$.

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

Lemma 7.3. *Let $0 < c < c'$ be fixed. There exists a constant $n_0 = n_0(c, c')$ such that for all $n > n_0$ the following holds. Let G be an n-vertex, 3-uniform, F_5-free hypergraph with $\delta(G) \geqslant c'\binom{n}{2}$. Let (G, γ, F) be the neighbourhood bundle of G. Then (G, γ, F) has a 5-cut of size at least $4\binom{c(n-1)}{2}$.*

Combining Theorem 7.1 with Lemma 7.3, we can prove Theorem 2.7.

Proof of Theorem 2.7. Let $c = (\sqrt{41} - 5)/8$, let $c' > c$ be fixed, and let G be any n-vertex, 3-uniform, F_5-free graph with minimum degree at least $c'\binom{n}{2}$. Let (G, γ, F) be the neighbourhood bundle of G. Let $b = (c' + c)/2$ so that $c' > b > c$. Then by Lemma 7.3, either $|V(G)|$ is bounded or (G, γ, F) contains a 5-cut of size at least $4\binom{b(n-1)}{2}$. Since $b > c$, if n is large enough this is at least $4c^2\binom{n}{2}$. Notice that $4c^2 = 1 - 5c$, so Theorem 7.1 implies that the chromatic number of G is bounded. □

The first step in the proof of Lemma 7.3 is the following lemma.

Lemma 7.4. *In a graph G, we call a non-edge $uv \notin E(G)$* good *if $N(u) \cap N(v) \neq \emptyset$. If G is a triangle-free graph with n vertices and m edges, then G has at least $m - n/2$ good non-edges.*

Proof. We prove this by induction on n. It is obviously true for $n = 1$ and $n = 2$. Now assume $n > 2$. If some component of G is not regular, then there exist vertices u, v in that component such that $u \in N(v)$ and $d(u) < d(v)$. Then $G - u$ has $n - 1$ vertices and $m - d(u)$ edges. By induction, $G - u$ has at least $m - d(u) - (n-1)/2$ good non-edges. For any vertex $w \in N(v) - u$, uw is a good non-edge, so G has at least $m - d(u) - (n-1)/2 + d(v) - 1 \geqslant m - n/2$ good non-edges. If all components of G are regular, then pick one component K. Assume K is r-regular, choose a vertex v in K, and let $N_2(v) = \{u :$ there exists a P_3 connecting u and $v\}$. If $|N_2(v)| \geqslant r$, then by the induction hypothesis $G - v$ has at least $m - r - (n-1)/2$ good non-edges, and since for any vertex $u \in N_2(v)$ it is the case that uv is a good non-edge, G has at least $m - r - (n-1)/2 + |N_2(v)| \geqslant m - n/2$ good non-edges. If $|N_2(v)| < r$, then since K is triangle-free and r-regular, K is the complete bipartite graph $K_{r,r}$, which has r^2 edges and $r^2 - r$ good non-edges. Now $G - K$ has $n - 2r$ vertices and $m - r^2$ edges, so by induction it has $m - r^2 - (n-2r)/2$ good non-edges. Then G has $m - r^2 - (n-2r)/2 + r^2 - r = m - n/2$ good non-edges. □

Proof of Lemma 7.3. We examine the copies of F_4 in G, where F_4 is the hypergraph with vertex set $\{1,2,3,4\}$ and edges $\{1,2,3\}$, $\{1,2,4\}$, and $\{2,3,4\}$.

Case 1. There exists a vertex v of G such that v is not contained in any copy of F_4. Consider $L = \gamma(v)[V(G) - v]$, which is a triangle-free graph with $n - 1$ vertices and at least $c\binom{n}{2}$ edges. By Lemma 7.4, L has at least $c\binom{n}{2} - (n-1)/2$ good non-edges. Let $X = \emptyset$ and S be the set of these good non-edges. We claim that (X, S) is a cut in (G, γ, F). Suppose for contradiction that there exists some $x \in V(G)$ and $\{u, w\} \in S$ such that $\{u, w, x\} \in G$. Pick a vertex y from $N_L(u) \cap N_L(w)$. Then u, v, w, x, y form a copy of F_5 in G, which is a contradiction.

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

Case 2. Every vertex of G is contained in some copy of F_4. Pick some $U \subseteq V(G)$ such that $G[U] = F_4$, let $U = \{u_1, u_2, u_3, u_4\}$, and let $G' = \bigcup_{i=1}^{4} \gamma(u_i)$. Consider $\gamma(u_i) \cap \gamma(u_j)$ for $i \neq j$. If $\gamma(u_i) \cap \gamma(u_j)$ contains a matching of size two, then G contains a copy of F_5. Say $ab, cd \in \gamma(u_i) \cap \gamma(u_j)$ with a, b, c, d distinct. Then since $G[U] = F_4$, there is an edge $E = \{u_i, u_j, w\} \in G$. If $w \neq a$ and $w \neq b$, then a, b, u_i, u_j, w form a copy of F_5 and if $w = a$ or $w = b$, then c, d, u_i, u_j, w form a copy of F_5. Thus $\gamma(u_i) \cap \gamma(u_j)$ is a star so has at most n elements. Since each $\gamma(x)$ has size at least $c'\binom{n}{2}$, G' has at least

$$4c'\binom{n}{2} - \binom{4}{2}n > 4c\binom{n}{2}$$

edges if n is large enough.

Then G' has n vertices and at least $4c\binom{n}{2}$ edges, so there exists a vertex v whose degree in G' is at least $4c(n-1)$. Let N denote the neighbourhood of v in G' and let $N_1, \ldots, N_4$ be a partition of N such that for every $1 \leqslant i \leqslant 4$ and every vertex $w \in N_i, vw \in \gamma(u_i)$. Let $X = U \cup \{v\}$ and $S = \bigcup_{i=1}^{4} \binom{N_i}{2}$, so that $|X| = 5$ and

$$|S| \geqslant 4\binom{|N|/4}{2} = 4\binom{c(n-1)}{2}.$$

We claim that (X, S) is a cut in (G, γ, F). Suppose for contradiction that there exists some $z \notin X$ such that $\gamma(z) \cap S \neq \emptyset$. Pick $\{x, y\} \in \gamma(z) \cap S$, then $\{x, y\} \subseteq N_i$ for some $1 \leqslant i \leqslant 4$. Now v, u_i, x, y, z form a copy of F_5, which is a contradiction.

From these two cases we can see that (G, γ, F) has a 5-cut of size at least

$$\min\left\{c\binom{n}{2} - \frac{n-1}{2}, 4\binom{c(n-1)}{2}\right\}.$$

Because G is F_5-free, it follows that $c \leqslant 2/9$ and therefore

$$\min\left\{c\binom{n}{2} - \frac{n-1}{2}, 4\binom{c(n-1)}{2}\right\} = 4\binom{c(n-1)}{2}.$$

□

7.3. A construction for the lower bound

To prove a lower bound on the chromatic threshold of the family of F_5-free hypergraphs, we need to construct an infinite sequence of F_5-free hypergraphs with large chromatic number and large minimum degree. Our construction is inspired by a construction by Hajnal [7] of a dense triangle-free graph with high chromatic number. Hajnal's key idea was to use the Kneser graph to obtain large chromatic number. The Kneser graph $\mathrm{KN}(n, k)$ has vertex set $\binom{[n]}{k}$, and two vertices F_1, F_2 form an edge if and only if $F_1 \cap F_2 = \emptyset$. We use an extension of Kneser graphs to hypergraphs. Alon, Frankl and Lovász [2] considered the Kneser hypergraph $\mathrm{KN}^r(n, k)$, which is the r-uniform hypergraph with vertex set $\binom{[n]}{k}$, and r vertices $F_1, \ldots, F_r$ form an edge if and only if $F_i \cap F_j = \emptyset$ for $i \neq j$. They gave a lower bound on the chromatic number of $\mathrm{KN}^r(n, k)$ as follows.

Theorem 7.5. *If $n \geqslant (t-1)(r-1) + rk$, then $\chi(\mathrm{KN}^r(n, k)) \geqslant t$.*

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

We first show that $\mathrm{KN}^r(n,k)$ is F_5-free for $n < 4k$.

Lemma 7.6. *If $n < 4k$, then $\mathrm{KN}^3(n,k)$ is F_5-free.*

Proof. Say $\{a,b,c\}$, $\{a,b,d\}$ and $\{c,d,e\}$ are edges in $\mathrm{KN}^3(n,k)$. Then by definition a, b, c, and d are four disjoint k-sets in $[n]$, which is impossible because $n < 4k$. □

Proof of the lower bound in Theorem 2.7. Fix $t \geqslant 2$ and $\epsilon > 0$. Pick $k \geqslant 2t$ and $n = 3k + 2(t-1)$ and note that $n < 4k$. By Theorem 7.5, $\mathrm{KN}^3(n,k)$ has chromatic number at least t and by Lemma 7.6 is F_5-free. For integers u, v, and w where n divides u, let U, V and W be disjoint vertex sets of size u, v, and w respectively. Partition U into $U_1,\dots,U_n$ such that $|U_i| = u/n$ for each i. Let H be the hypergraph with vertex set $V(\mathrm{KN}^3(n,k)) \cup U \cup V \cup W$ and the following edges.

- For $\{S_1,S_2,S_3\} \in \mathrm{KN}^3(n,k)$, make $\{S_1,S_2,S_3\}$ an edge of H.
- For $S \in V(\mathrm{KN}^3(n,k))$, $x \in U_i$ with $1 \leqslant i \leqslant n$, and $y \in V$, make $\{S,x,y\}$ an edge of H if $i \in S$.
- For $x \in U$, $y \in V$, and $z \in W$, make $\{x,y,z\}$ an edge of H.

Notice that H has chromatic number at least t because $\mathrm{KN}^3(n,k)$ is a subhypergraph.

Claim 1. H contains no subgraph isomorphic to F_5.

Proof. Suppose $\{a,b,c\},\{a,b,d\}$ and $\{c,d,e\}$ are the hyperedges of a copy of F_5 in H. Notice that the hypergraph induced by $U, V, V(\mathrm{KN}^3(n,k)) \cup W$ is 3-partite, apart from those edges within $\mathrm{KN}^3(n,k)$. Note that a 3-uniform, 3-partite hypergraph is F_5-free, therefore any copy of F_5 must contain an edge from $\mathrm{KN}^3(n,k)$. If that edge is $\{a,b,c\}$ then d must also be contained in $V(\mathrm{KN}^3(n,k))$. But then c and d are both in $V(\mathrm{KN}^3(n,k))$, which means e must be as well. Because $\mathrm{KN}^3(n,k)$ is F_5-free, this is a contradiction. Similarly, $\{a,b,d\} \subsetneq V(\mathrm{KN}^3(n,k))$. Therefore, $\{c,d,e\} \subseteq V(\mathrm{KN}^3(n,k))$, and without loss of generality $b \in U$ and $a \in V$. Because $\{a,b,c\}$ and $\{a,b,d\}$ are edges, b must be in both c and d, which contradicts the fact that $\{c,d,e\}$ is an edge of $\mathrm{KN}^3(n,k)$. □

Claim 2. The minimum degree of H is at least

$$(1-\epsilon)\frac{6}{49}\binom{|V(H)|}{2}.$$

if $|V(H)|$ is large enough.

Proof. Vertices in $\mathrm{KN}^3(n,k)$ have degree at least

$$k\frac{u}{n}v = \frac{kuv}{3k+2(t-1)}.$$

Since t is fixed, we can choose k large enough that vertices in $\mathrm{KN}^3(n,k)$ have degree at least $(1-\epsilon/2)uv/3$. Vertices in A have degree at least vw, vertices in B have degree at

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

least uw, and vertices in C have degree at least uv. Thus the minimum degree of H is at least

$$\min\left\{(1-\epsilon/2)\frac{uv}{3}, uw, vw\right\}.$$

Choose u, v, and w so that $(uv)/3 = uw = vw$; we obtain that $u = v$ and $w = v/3$ and the minimum degree is at least $(1-\epsilon/2)u^2/3$. The number of vertices is

$$u + v + w + \binom{n}{k} = \frac{7}{3}u + \binom{n}{k}.$$

Since $u^2/3 \approx 6/49\binom{7u/3}{2}$, we can choose u sufficiently large that the minimum degree of H is at least

$$(1-\epsilon)\frac{6}{49}\binom{|V(H)|}{2}. \qquad \square$$

We have proved that for every fixed $t \geqslant 2$ and every $\epsilon > 0$, there is a constant N_0 such that for $N > N_0$ there exists an N-vertex, 3-uniform, F_5-free hypergraph with chromatic number at least t and minimum degree at least

$$(1-\epsilon)\frac{6}{49}\binom{|V(H)|}{2}.$$

By the definition of chromatic threshold, this implies that the chromatic threshold of the family of F_5-free hypergraphs is at least $\frac{6}{49}$. $\square$

8. Generalized Kneser hypergraphs

In Section 7.3, we used a generalization of the Kneser graph to hypergraphs to give a lower bound on the chromatic threshold of the family of F_5-free hypergraphs. In Section 9, we will use similar constructions to give lower bounds on the chromatic threshold of the family of A-free hypergraphs, for several other hypergraphs A. For some of these constructions we will need a more general variant of the Kneser hypergraph, which we explore in this section.

Sarkaria [33] considered the generalized Kneser hypergraph $\mathrm{KN}_s^r(n,k)$, that is, the r-uniform hypergraph with vertex set $\binom{[n]}{k}$, in which r vertices $F_1,\dots,F_r$ form an edge if and only if no element of $[n]$ is contained in more than s of them. Note that the Kneser hypergraph $\mathrm{KN}^r(n,k)$ is $\mathrm{KN}_1^r(n,k)$. Sarkaria [33] and Ziegler [40] gave lower bounds on the chromatic number of $\mathrm{KN}_s^r(n,k)$, but Lange and Ziegler [23] showed that the lower bounds obtained by Sarkaria and Ziegler apply only if one allows the edges of $\mathrm{KN}_s^r(n,k)$ to have repeated vertices. We conjecture that for $\mathrm{KN}_s^r(n,k)$, a statement similar to Theorem 7.5 is true.

Conjecture 8.1. *There exists $T(r,s,t)$ such that if $n \geqslant T(r,s,t) + rk/s$, then*

$$\chi(\mathrm{KN}_s^r(n,k)) \geqslant t. \qquad \square$$

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

The following much weaker statement is sufficient for our purposes. The proof is similar to an argument of Szemerédi which appears in a paper of Erdős and Simonovits [7], and the proof of Claim 1 is motivated by an argument of Kleitman [22].

Theorem 8.2. *Let $c > 0$; then for any integers r, t, there exists $K_0 = K_0(c, r, t)$ such that if $k \geqslant K_0$, $s = r - 1$, and $n = (r/s + c)k$, then $\chi(\mathrm{KN}_s^r(n,k)) > t$.*

Before we prove this theorem, we need two definitions. A family $\mathcal{F}$ of subsets of $[n]$ is *monotone decreasing* if $F \in \mathcal{F}$ and $F' \subseteq F$ imply $F' \in \mathcal{F}$. Similarly, it is *monotone increasing* if $F \in \mathcal{F}$ and $F \subseteq F'$ imply $F' \in \mathcal{F}$.

Proof Proof of Theorem 8.2. Fix an integer t. We would like to prove that if k is large enough then it is impossible to t-colour $\mathrm{KN}_s^r(n,k)$. So let k be some integer and assume $\mathrm{KN}_s^r(n,k)$ can be t-coloured. Then the k-subsets of $[n]$ can be divided into t families, $\mathcal{F}_1, \ldots, \mathcal{F}_t$, such that $F_1 \cap \cdots \cap F_r \neq \emptyset$ for all distinct $F_1, \ldots, F_r \in \mathcal{F}_i, 1 \leqslant i \leqslant t$. For $1 \leqslant i \leqslant t$, let

$$\mathcal{F}_i^* = \{A : A \subseteq [n], \exists F \in \mathcal{F}_i \text{ such that } F \subseteq A\}.$$

Then $\mathcal{F}_1^*, \ldots, \mathcal{F}_t^*$ are monotone increasing families of subsets of $[n]$. Let $w = s/r$; since $s = r - 1$, $w = 1 - 1/r$. For a family $\mathcal{F}$ of subsets of $[n]$, define the weighted size $W[\mathcal{F}]$ of $\mathcal{F}$ by

$$W[\mathcal{F}] = \sum_{F \in \mathcal{F}} w^{|F|}(1-w)^{n-|F|}.$$

Claim 1. For $1 \leqslant \ell \leqslant t$,

$$W\left[\bigcup_{i=1}^{\ell} \mathcal{F}_i^*\right] \leqslant 1 - 1/r^{\ell}.$$

Proof. We prove this by induction on ℓ. For $\ell = 1$, Frankl and Tokushige [12] showed that for a family $\mathcal{F}$ of subsets of $[n]$, if $F_1 \cap \cdots \cap F_r \neq \emptyset$ for all distinct $F_1, \ldots, F_r \in \mathcal{F}$, then $W[\mathcal{F}] \leqslant w = 1 - 1/r$. Now assume that the statement is true for ℓ. Let $U = \bigcup_{i=1}^{\ell} \mathcal{F}_i^*$ and $L = \overline{\mathcal{F}_{\ell+1}^*}$. Then $W[U] \leqslant 1 - 1/r^{\ell}$, U is a monotone increasing family of subsets of $[n]$, and L is a monotone decreasing family of subsets of $[n]$. By the FKG inequality,

$$W[U \cap L] \leqslant W[U]W[L].$$

Then

$$\begin{aligned} W\left[\bigcup_{i=1}^{\ell+1} \mathcal{F}_i^*\right] &= W[U \cap L] + W[\mathcal{F}_{\ell+1}^*] \leqslant W[U]W[L] + W[\mathcal{F}_{\ell+1}^*] \\ &\leqslant (1 - 1/r^{\ell})W[L] + W[\mathcal{F}_{\ell+1}^*] = 1 - (1 - W[\mathcal{F}_{\ell+1}^*])/r^{\ell}. \end{aligned}$$

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

Since $W[\mathcal{F}^*_{\ell+1}] \leqslant w = 1 - 1/r$, we have

$$1 - (1 - W[\mathcal{F}^*_{\ell+1}])/r^{\ell} \leqslant 1 - 1/r^{\ell+1},$$

so

$$W\left[\bigcup_{i=1}^{\ell+1} \mathcal{F}^*_i\right] \leqslant 1 - 1/r^{\ell+1}. \qquad \square$$

Now we know that $W[\bigcup_{i=1}^{t} \mathcal{F}^*_i] \leqslant 1 - 1/r^t$, so

$$W\left[\overline{\bigcup_{i=1}^{t} \mathcal{F}^*_i}\right] \geqslant 1/r^t.$$

We also know that $\overline{\bigcup_{i=1}^{t} \mathcal{F}^*_i}$ is the family of subsets of $[n]$ whose size is less than $k = n/(r/s + c)$, so

$$W\left[\overline{\bigcup_{i=1}^{t} \mathcal{F}^*_i}\right] = \sum_{i<n/(r/s+c)} \binom{n}{i} w^i (1-w)^{n-i}.$$

Since

$$wn = \frac{n}{r/s} > \frac{n}{r/s + c},$$

by Chernoff's inequality we have

$$\sum_{i<\frac{n}{r/s+c}} \binom{n}{i} w^i (1-w)^{n-i} \leqslant e^{-\left(\frac{c}{r/s+c}\right)^2 \frac{sn}{2r}} = e^{-\frac{c^2 s}{2(r/s+c)r} k}.$$

Then if k is large and t is fixed,

$$W\left[\overline{\bigcup_{i=1}^{t} \mathcal{F}^*_i}\right] \leqslant e^{-\frac{c^2 s}{2(r/s+c)r} k} < 1/r^t,$$

which contradicts Claim 1. This contradiction implies that for any fixed t, there is no choice of K_0 such that for all $k > K_0$ it is possible to t-colour $\mathrm{KN}^r_s(n,k)$. This completes the proof. $\square$

For an r-uniform hypergraph A, we want to construct an infinite sequence of A-free hypergraphs with $\mathrm{KN}^r(n,k)$ or $\mathrm{KN}^r_{r-1}(n,k)$ as a subhypergraph. This will imply that these A-free hypergraphs have large chromatic number, but we must first show that for any integer k and for some choice of $n = n(k)$, one of $\mathrm{KN}^r(n,k)$, $\mathrm{KN}^r_{r-1}(n,k)$ is A-free. We now show that $\mathrm{KN}^3_2(n,k)$ is T_5-free and $S(7)$-free under some conditions on n and k. Here T_5 is a 3-uniform hypergraph with vertices v_1, v_2, v_3, v_4, v_5 and edges

$$\{v_1, v_2, v_3\}, \quad \{v_1, v_4, v_5\}, \quad \{v_2, v_4, v_5\}, \quad \{v_3, v_4, v_5\},$$

and $S(7)$ denotes the Fano plane (the S stands for Steiner Triple System.)

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

Lemma 8.3. *If* $n < (3/2 + 1/4)k$*, then* $\mathrm{KN}_2^3(n,k)$ *is* T_5*-free.*

Proof. If $n < 3k/2$, then $\mathrm{KN}_2^3(n,k)$ has no edge and of course is T_5-free. Assume $n = (3/2+\epsilon)k$ with $0 \leqslant \epsilon < 1/4$, and suppose T_5 is a subhypergraph of $\mathrm{KN}_2^3(n,k)$. Since $\{v_1, v_4, v_5\}$, $\{v_2, v_4, v_5\}$, $\{v_3, v_4, v_5\}$ are edges of T_5, the vertices v_1, v_2, and v_3 all lie in $\overline{v_4 \cap v_5}$. Because

$$|\overline{v_4 \cap v_5}| \leqslant 2n - 2k = (1+2\epsilon)k < 3k/2,$$

by the pigeonhole principle, $v_1 \cap v_2 \cap v_3 \neq \emptyset$, which means $\{v_1, v_2, v_3\}$ is not an edge, a contradiction. □

Lemma 8.4. *If* $n < (3/2 + 1/10)k$*, then* $\mathrm{KN}_2^3(n,k)$ *is* $S(7)$*-free.*

Proof. Just as in the proof of Lemma 8.3, assume $n = (3/2+\epsilon)k$ with $0 \leqslant \epsilon < 1/10$ and suppose $S(7)$ is a subhypergraph of $\mathrm{KN}_2^3(n,k)$. Let A be a vertex in a copy of $S(7)$ in $\mathrm{KN}_2^3(n,k)$ and let $\{A,B,C\}, \{A,D,E\}, \{A,F,G\}$ be its incident edges in the copy of $S(7)$. Then $B \cap C, D \cap E, F \cap G \subseteq \overline{A}$. Since $|\overline{A}| = (1/2+\epsilon)k$,

$$|B \cap C|, |D \cap E|, |F \cap G| \geqslant (1/2 - \epsilon)k.$$

Then since $3(1/2-\epsilon) > 2(1/2+\epsilon)$, the pigeonhole principle implies that $B \cap C \cap D \cap E \cap F \cap G \neq \emptyset$. Now the copy of $S(7)$ cannot have an edge not containing A, a contradiction. □

We will use Lemma 8.4 in Section 9.2 to provide a lower bound on the chromatic threshold of the family of $S(7)$-free hypergraphs. Similarly, we will use Lemma 8.3 in Section 9.3 to provide a lower bound on the chromatic threshold of the family of T_5-free hypergraphs.

9. Open problems and partial results

Many open problems remain; for most 3-uniform hypergraphs A the chromatic threshold for the family of A-free hypergraphs is unknown. Interesting hypergraphs to study are those for which we know the extremal number, $\mathrm{ex}(n,A)$, and we will examine a few of those here along with partial results and conjectures. We conjecture that most of the lower bounds given by the constructions in this section are tight.

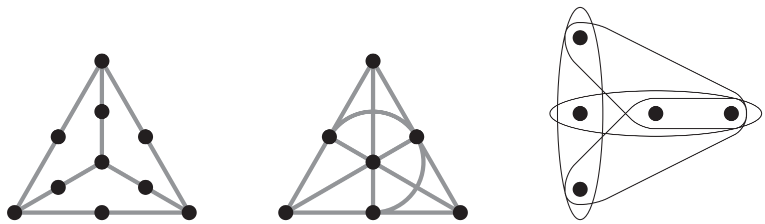
9.1. $\mathcal{TK}^r(s)$-free hypergraphs

For $s > r$, recall that $\mathcal{TK}^r(s)$ is the family of r-uniform hypergraphs such that there exists a set S of s vertices where each pair of vertices from S are contained together in some edge. The set S is called the set of *core vertices* of the hypergraph. Recall also that $T_{r,s}(n)$ is the complete n-vertex, r-uniform, s-partite hypergraph with part sizes as equal as possible.

The last author [27] showed that if $s > r$ then

$$\mathrm{ex}(n, \mathcal{TK}^r(s)) = |T_{r,s-1}(n)| \quad \text{and} \quad \mathrm{ex}(n, \mathrm{TK}^r(s)) = (1+o(1))|T_{r,s}(n)|.$$

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press



(a) $\mathrm{TK}^3(4)$ (b) $S(7)$ (c) T_5

Figure 4. Assorted hypergraphs.

Recently, Pikhurko [30] has shown that for large n and $s > r$, $\mathrm{ex}(n, \mathrm{TK}^r(s)) = |T_{r,s-1}(n)|$ and that $T_{r,s-1}(n)$ is the unique extremal example. Because F_5 is a member of $\mathcal{TK}^3(4)$ it follows that the chromatic threshold of $\mathcal{TK}^3(4)$-free hypergraphs is at most $(\sqrt{41}-5)/8$. The following simple variation on the construction from Section 7.3 provides a lower bound of $18/361$ for both $\mathrm{TK}^3(4)$-free and $\mathcal{TK}^3(4)$-free hypergraphs.

Proposition 9.1. *The chromatic threshold of* $\mathcal{TK}^3(4)$*-free hypergraphs is at least* $18/361$. □

Proof. The proof is very similar to the proof in Section 7.3, so we only sketch it here. Choose k, n, u, v, w, U, V, W as in the proof of the lower bound of Theorem 2.7 in Section 7.3; that is, k, n, u, v, w are integers with $n \ll u, v, w$ and U, V, W are disjoint sets of vertices of size u, v, w respectively. Divide U into $U_1, \ldots, U_n$ so that $|U_i| = u/n$, and divide V into $V_1, \ldots, V_n$ such that $|V_i| = v/n$. Let H be the hypergraph formed by taking $\mathrm{KN}^3(n,k)$ and adding the complete 3-partite hypergraph on U, V, W and the following edges. For $S \in V(\mathrm{KN}^3(n,k))$ and $x \in U_i$ and $y \in V_j$, make $\{S, x, y\}$ an edge if $i, j \in S$. The minimum degree is maximized when $u = v$ and $w = u/9$, which gives minimum degree approximately

$$uv/9 \approx \frac{18}{361} \cdot \binom{N}{2},$$

where $N = u + v + w + \binom{n}{k}$ is the number of vertices in the hypergraphs.

Let F be any hypergraph in $\mathcal{TK}^3(4)$ and assume that F is a subhypergraph of H in which c_1, c_2, c_3, c_4 are the four core vertices. Because any 3-partite hypergraph is $\mathcal{TK}^3(4)$-free, it is easy to see that some edge of F must lie in $\mathrm{KN}^3(n,k)$, and so there must be at least two core vertices in $\mathrm{KN}^3(n,k)$. If $c_1, c_2 \in \mathrm{KN}^3(n,k)$ and $c_3 \in U \cup V$, then c_3 is in either U_i or V_i for some i. But then $i \in c_1 \cap c_2$ (recall that vertices in $\mathrm{KN}^3(n,k)$ are k-sets), which contradicts the fact that c_1 and c_2 are contained together in some edge of $\mathrm{KN}^3(n,k)$. Thus all four core vertices must be in $\mathrm{KN}^3(n,k)$, which is not possible because $n < 4k$. □

This gives lower bounds on the chromatic thresholds of $\mathrm{TK}^3(4)$-free and $\mathcal{TK}^3(4)$-free hypergraphs and leads to the following questions.

Question 9.2. *What is the chromatic threshold for* $\mathrm{TK}^3(4)$*-free hypergraphs? It is between* $18/361$ *and* $2/9$*. What is the chromatic threshold for* $\mathcal{TK}^3(4)$*-free hypergraphs? It has the same lower bound as for* $\mathrm{TK}^3(4)$*-free hypergraphs, and because* $F_5 \in \mathcal{TK}^3(4)$ *the upper bound is* $(\sqrt{41}-5)/8$*.*

A similar construction provides a $\mathcal{TK}^3(s)$-free hypergraph for any $s \geqslant 5$. We have not optimized the values.

Lemma 9.3. *When* $s \geqslant 5$*, the chromatic threshold of* $\mathcal{TK}^3(s)$*-free hypergraphs is at least*

$$\frac{(s-2)(s-3)(s-4)^2}{(s^2-13)^2} = 1 - \frac{13}{s} + O\left(\frac{1}{s^2}\right).$$

Proof. Fix $t \geqslant 2$, $k \geqslant 2t$, and let $n = 3k + 2(t-1)$. Notice that $n < 4k$. By Theorem 7.5, the chromatic number of $\mathrm{KN}^3(n,k)$ is therefore at least t. Fix $N \gg \binom{n}{k}$.

Partition N vertices into one part of size u and $s-2$ parts of size x, for some u that is divisible by n. Include as an edge each triple that has at most one vertex in each part. Further partition the part of size u into n sets, $U_1, \ldots, U_n$, each of size u/n. From the remaining $s-2$ parts of size x, choose two and designate them W_1, W_2; label the remaining $s-4$ parts $V_1, \ldots, V_{s-4}$. Let H be the 3-uniform hypergraph formed by taking the disjoint union of $\mathrm{KN}^3(n,k)$ and the above complete $(s-1)$-partite hypergraph, and adding the following edges. If $S \in V(\mathrm{KN}^3(n,k))$, $v \in V_i$, and $v' \in V_j$ for $i \neq j$, add the edge $\{S, v, v'\}$. If $S \in V(\mathrm{KN}^3(n,k))$ and $u \in U_i$ and $v \in V_j$, then add the edge $\{S, u, v\}$ if and only if $i \in S$.

Notice that H has chromatic number at least t, and that $V(H) = N + \binom{n}{k}$.

Claim 1. H contains no element of $\mathcal{TK}^3(s)$ as a subgraph.

Proof. Suppose there is such a subgraph; then at least one core vertex must be contained in $V(\mathrm{KN}^3(n,k))$, because an $(s-1)$-partite graph is $\mathcal{TK}^s(3)$-free. In that case, no core vertex can be in $W_1 \cup W_2$ because there is no edge that contains a vertex from $W_1 \cup W_2$ as well as a vertex from $V(\mathrm{KN}^3(n,k))$. There must therefore be at least three core vertices in $V(\mathrm{KN}^3(n,k))$, which means that two of them must appear in an edge contained within $V(\mathrm{KN}^3(n,k))$. Suppose they are S_1, S_2. If another core vertex is in U, say $u \in U_i$, then there must be an edge of H containing u and S_1, and there must be an edge containing u and S_2. This implies that $i \in S_1 \cap S_2$, which contradicts the fact that S_1 and S_2 appear together in an edge of $\mathrm{KN}^3(n,k)$.

All core vertices must therefore be in $V(\mathrm{KN}^3(n,k)) \cup V$, which means that there must be at least four of them in $V(\mathrm{KN}^3(n,k))$. Because each pair of those four core vertices must appear together in an edge, and that edge must be in $\mathrm{KN}^3(n,k)$, those four sets must be pairwise disjoint. This is impossible because $n < 4k$. □

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

The minimum degree of this graph is approximately

$$\min\left\{\frac{1}{3}(s-4)ax+\binom{s-4}{2}x^2,\binom{s-2}{2}x^2,(s-3)ax+\binom{s-3}{2}x^2\right\}.$$

Notice that a vertex in $W_1 \cup W_2$ has degree strictly less than a vertex in $\mathrm{KN}^3(n,k)$, and so they do not enter into the above computation. This minimum is largest when

$$u=\frac{3(2s-7)x}{s-4},$$

which implies that

$$x=\left(\frac{s-4}{s^2-13}\right)N.$$

The minimum degree of H is then

$$\frac{(s-2)(s-3)}{2}\cdot\frac{(s-4)^2}{(s^2-13)^2}N^2=\left(1-\frac{13}{x}+O\left(\frac{1}{s^2}\right)\right)\frac{N^2}{2}. \qquad \square$$

The construction in Lemma 9.3 has one part of 'type' U (which is partitioned into n sets), $s-4$ parts of 'type' V (which are not partitioned, and whose vertices appear in edges that intersect K), and two parts of 'type' W (which are not partitioned and have no vertices that appear in edges intersecting K). Using this strategy, one can generate similar constructions for $\mathrm{TK}^r(s)$; the above proof applies whenever there are x parts of type U, $s-(r+1)$ parts of type V, and y parts of type W, where $x+y=r$ and $s-(r+1)+x \geqslant r-1$. This last condition is needed for the edges intersecting K.

Question 9.4. *What is the chromatic threshold for* $\mathrm{TK}^3(s)$*-free hypergraphs for* $s>3$*? It is between*

$$\frac{(s-2)(s-3)(s-4)^2}{(s^2-13)^2}=1-\frac{13}{s}+O\left(\frac{1}{s^2}\right)$$

and

$$\left(1-\frac{1}{s-1}\right)\left(1-\frac{2}{s-1}\right)=1-\frac{3}{s-1}+\frac{2}{(s-1)^2}.$$

The upper bound comes from $T_{r,s-1}(n)$.

9.2. $S(7)$-free hypergraphs

Next, consider the Fano plane $S(7)$. De Caen and Füredi [6] showed that

$$\mathrm{ex}(n,S(7))=\left(\frac{3}{4}+o(1)\right)\binom{n}{3}.$$

The extremal hypergraph for $S(7)$, proved to be extremal by Füredi and Simonovits [14] and also by Keevash and Sudakov [21], is the hypergraph formed by taking two almost equal vertex sets U and V and taking all edges that have at least one vertex in each of U and V. We can modify the hypergraph from Section 7.3 to obtain a lower bound on the chromatic threshold of $S(7)$-free hypergraphs.

Proposition 9.5. *The chromatic threshold of $S(7)$-free hypergraphs is at least $9/17$.* □

Proof. Fix $t \geqslant 2$ and $0 < \epsilon \ll 1$. Then by Lemma 7.5 there exists k sufficiently large that if $n = (3+\epsilon)k$ then $\mathrm{KN}^3(n,k)$ has chromatic number at least t. Fix such a k, and fix $N \gg \binom{n}{k}$.

Partition N vertices into two sets, U and V, with $|U| = 9N/17$ and $|V| = 8N/17$. Further partition U into n parts, $U_1, \dots, U_n$, each of size $|U|/n$. Include as an edge each triple that has at least one vertex in each of U, V. Let H be the hypergraph formed by taking the disjoint union of this hypergraph and $\mathrm{KN}^3(n,k)$ and adding the following edges. For $u \in U_i$, $u' \in U_j$, and $X \in V(\mathrm{KN}^3(n,k))$ include $\{X, u, u'\}$ as an edge if $i, j \in X$ (recall that vertices in $\mathrm{KN}^3(n,k)$ are subsets of $[n]$). Let $K = V(\mathrm{KN}^3(n,k))$. Notice that H has chromatic number at least t, and that $V(H) = N + \binom{n}{k}$.

Claim 1. H contains no subhypergraph isomorphic to $S(7)$.

Proof. First notice that $\mathrm{KN}^3(n,k)$ is $S(7)$-free because every pair of vertices in $S(7)$ are in an edge, which would require there to be 7 pairwise-disjoint k-subsets of $[n]$. Because $n = (3+\epsilon)k$, this would be a contradiction. It is easy to see, by considering the partition $U, (K \cup V)$, that if H contains a copy of $S(7)$ then it must involve an edge from $H[K]$ (otherwise the extremal $S(7)$-free hypergraph also contains a copy of $S(7)$). Call this edge $\{A, B, C\}$.

There are four vertices in $S(7) \setminus \{A, B, C\}$, and at least one must be outside K. No more than one can be in V because there is no edge with one vertex in K and two in V. No more than one can be in U otherwise one of $A \cap B$, $A \cap C$, $B \cap C$ is non-empty, which contradicts the assumption that $\{A, B, C\}$ is an edge of $H[K]$. Therefore, there must be either five or six vertices of $S(7)$ in K. Suppose v is a vertex of $S(7)$ that is outside K. Then v appears in three edges that overlap only at v, say $\{v, S_1, S_2\}$, $\{v, S_3, S_4\}$, and $\{v, S_5, S_6\}$. At least one of these edges must contain two vertices from K, but there is no such edge in H. □

The minimum degree of H is at least

$$\min\left\{|U||V| + \binom{|U|/3}{2}, |U||V| + \binom{|U|}{2}, |U||V| + \binom{|V|}{2}\right\} = \frac{9}{34}N^2 - \frac{3}{34}N. \qquad \square$$

Question 9.6. *What is the chromatic threshold of $S(7)$-free hypergraphs? It is at least $9/17$ and at most $3/4$, where the upper bound is from the extremal hypergraph of $S(7)$.*

9.3. T_5-free hypergraphs

Recall that the 3-uniform hypergraph T_5 has vertices A, B, C, D, E and edges

$$\{A, B, C\}, \quad \{A, D, E\}, \quad \{B, D, E\}, \quad \{C, D, E\}.$$

Let $B^3(n)$ be the 3-uniform hypergraph with the most edges among all n-vertex 3-graphs whose vertex set can be partitioned into X_1, X_2 such that each edge contains exactly one

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

vertex from X_2. Füredi, Pikhurko and Simonovits [13] proved that for n sufficiently large the extremal T_5-free hypergraph is $B^3(n)$. It follows that the chromatic threshold for the family of T_5-free hypergraphs is at most $4/9$.

Proposition 9.7. *The chromatic threshold of T_5-free hypergraphs is at least* $16/49$. □

Proof. Fix $t \geqslant 2$ and $0 < \epsilon \ll 1$. Then by Lemma 7.5 there exists k sufficiently large that if $n = (3/2+\epsilon)k$ then $\mathrm{KN}_2^3(n,k)$ has chromatic number at least t. Fix such a k, and fix $N \gg \binom{n}{k}$.

Partition N vertices into two parts, U and V, with $|U| = 4N/7$ and $|V| = 3N/7$. Further partition U into n parts, $U_1, \ldots, U_n$, each of size $|U|/n$. Include as an edge any triple with two vertices in U and one in V. Let H be the hypergraph formed by taking the disjoint union of this graph and $\mathrm{KN}_2^3(n,k)$ and including the following edges. If $X \in V(\mathrm{KN}_2^3(n,k))$ and $u \in U_i$ and $v \in V$, then let $\{u, v, X\}$ be an edge if $i \in X$ (recall that vertices of $\mathrm{KN}_2^3(n,k)$ are subsets of $[n]$). Let $K = V(\mathrm{KN}_2^3(n,k))$. Notice that H has chromatic number at least t, and that $V(H) = N + \binom{n}{k}$.

Claim 1. T_5 is not a subhypergraph of H.

Proof. Let H' be the hypergraph obtained from H by deleting all edges contained in K, and let $X_1 = K \cup U$ and $X_2 = V$. It is now easy to see that H' is a subhypergraph of the extremal T_5-free hypergraph; if H contains a copy of T_5 it must therefore involve an edge from K. If that edge is $\{A, D, E\}$ (see the labelling of T_5 above), then because $\{B, D, E\}$ and $\{C, D, E\}$ are edges of T_5 it must be the case that both of B, C are in K, but by Lemma 8.3 K does not span a copy of T_5. Similarly, neither $\{B, D, E\}$ nor $\{C, D, E\}$ can be contained in K.

We may therefore assume that $\{A, B, C\}$ is contained in K. Because $\{A, D, E\}$ is an edge, and by Lemma 8.3, at least one of D, E is in U. Suppose that $D \in U_i$. Then because $\{A, D, E\}$, $\{B, D, E\}$, and $\{C, D, E\}$ are all edges of T_5 it must be the case that $i \in A \cap B \cap C$. This contradicts the assumption that $\{A, B, C\}$ is an edge. □

The minimum degree of H is at least

$$\min\left\{\frac{2|U||V|}{3}, |U||V|, \binom{|U|}{2}\right\} = \frac{8}{49}N^2 - \frac{2}{7}N. \qquad \square$$

9.4. Co-chromatic thresholds

There is another possibility when generalizing the definition of chromatic threshold from graphs to hypergraphs: we can use the co-degree instead of the degree. Recall that if H is an r-uniform hypergraph and $\{x_1, \ldots, x_{r-1}\} \subseteq V(H)$, then the *co-degree* $d(x_1, \ldots, x_{r-1})$ *of* $x_1, \ldots, x_{r-1}$ is $|\{z : \{x_1, \ldots, x_{r_1}, z\} \in H\}|$. Let F be a family of r-uniform hypergraphs. The *co-chromatic threshold* of F is the infimum of the values $c \geqslant 0$ such that the subfamily of F consisting of hypergraphs H with minimum co-degree at least $c|V(H)|$ has bounded

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

chromatic number. More generally, the *k-degree* $d(x_1, \ldots, x_k)$ *of* $x_1, \ldots, x_k$ is

$$|\{\{z_{k+1}, \ldots, z_r\} : \{x_1, \ldots, x_k, z_{k+1}, \ldots, z_r\} \in H\}|$$

and we define the k-chromatic threshold similarly. Given a hypergraph H and subsets U, V, W of $V(H)$, we say that an edge $\{u, v, w\}$ is of type UVW if $u \in U, v \in V$ and $w \in W$.

The co-chromatic thresholds of F_5-free hypergraphs and $\mathrm{TK}^3(4)$-free hypergraphs are trivially zero because if the minimum co-degree of H is at least 10 then H contains a copy of $\mathrm{TK}^3(4)$ and a copy of F_5. For the Fano plane, the last author proved [26] that for every $\epsilon > 0$ there exists n_0 such that any 3-uniform hypergraph with $n > n_0$ vertices and minimum co-degree greater than $(1/2 + \epsilon)n$ contains a copy of $S(7)$. In 2009, Keevash [18] improved this by proving that any 3-uniform hypergraph with minimum co-degree greater than $n/2$ contains a copy of $S(7)$ for n sufficiently large. Notice that the lower bound construction for the chromatic threshold described above has non-zero minimum co-degree but the co-degree depends on the parameter t. We can modify the construction to prove a better lower bound on the co-chromatic threshold of $S(7)$-free hypergraphs.

Proposition 9.8. *The co-chromatic threshold of* $S(7)$*-free hypergraphs is at least* $2/5$. □

Proof. Fix $t \geqslant 2$ and $0 < \epsilon \ll 1$. Then by Lemma 8.2 there exists k large enough that if $n = (3/2 + \epsilon)k$ then $\mathrm{KN}_2^3(n, k)$ has chromatic number at least t. Fix $N \gg \binom{n}{k}$.

Partition N vertices into two parts, U and V, of size $\frac{3N}{5}$ and $\frac{2N}{5}$ respectively. Include as an edge any triple with at least one vertex in each part. Further partition U into n sets, $U_1, \ldots, U_n$, each of size $|U|/n$. Let H be the hypergraph formed by taking the disjoint union of this hypergraph with $\mathrm{KN}_2^3(n, k)$ and including the following edges. Include any edge of type KUV, where $K = V(\mathrm{KN}_2^3(n, k))$. For any $X, Y \in K$, if $|X \cap Y| < k - 4\epsilon k$ then include every edge of the form $\{X, Y, u\}$ where $u \in U_i$ for some $i \in X \cup Y$. If $|X \cap Y| \geqslant k - 4\epsilon k$ then include every edge of the form $\{X, Y, u\}$ where $u \in U_i$ for some $i \in X \cap Y$. Notice that H has chromatic number at least t and that $V(H) = N + \binom{n}{k}$.

Claim 1. The above hypergraph contains no subgraph isomorphic to $S(7)$.

Proof. First notice that the complete bipartite 3-uniform hypergraph contains no copy of $S(7)$. Therefore, by considering the partition $U, V \cup K$, we can see that any copy of $S(7)$ must contain an edge induced by K. Call this edge $\{A, B, C\}$. It also follows from Lemma 8.4 that there is no copy of $S(7)$ completely contained in K.

Claim 1a. Any copy of $S(7)$ intersects U (or V) in at most one vertex.

Proof. Notice that for any edge e in $S(7)$, every other edge intersects e in at exactly one vertex; therefore for any copy of $S(7)$ in H every edge contains one of A, B, C. If there were two vertices of $S(7)$ in U (or in V) then the edge of $S(7)$ joining them would be unable to intersect A, B, or C. □

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

Claim 1b. Any copy of $S(7)$ contains no vertex from V.

Proof. Suppose for contradiction that a copy of $S(7)$ contains some vertex from V; then by Claim 1a it intersects V in exactly one vertex. Every vertex of $S(7)$ is contained in three edges, but because there is at most one vertex from U involved in the copy of $S(7)$, there can be only one edge that contains the vertex from V. □

Any copy of $S(7)$ must therefore have exactly six vertices in K and exactly one vertex in U. Suppose they are $A,B,C,D,E,F \in K$ and $G \in U_i$. Suppose also that the edges of $S(7)$ induced by K are

$$\{A,B,C\}, \quad \{A,E,F\}, \quad \{C,D,E\}, \quad \{B,D,F\}.$$

Claim 1c. If $\{S_1,S_2,S_3\}$ is an edge in K then $|S_i \cap S_j| \leqslant k/2+\epsilon k$ for all $i \neq j$.

Proof. This follows from the definition of the hypergraph on K,

$$k=|S_1| \leqslant n-|S_2 \cap S_3| = (3/2+\epsilon)k - |S_2 \cap S_3|, \quad \text{so } |S_2 \cap S_3| \leqslant k/2+\epsilon k,$$

and the claim follows through symmetry. □

Claim 1d. The following intersections all have size at least $2k-4\epsilon k$: $A \cap D, B \cap E, C \cap F$.

Proof. We will prove that $|A \cap D| \geqslant 2k-4\epsilon k$; the rest follow through symmetry. Because $\{B,D,F\}$ is an edge,

$$D \subseteq (\overline{B} \cap F) \cup (B \cap \overline{F}) \cup (\overline{B} \cap \overline{F}).$$

Also, because $\{A,B,C\}$ is an edge,

$$|\overline{A} \cap \overline{B}| = |\overline{A}| - |\overline{A} \cap B| \leqslant (k/2+\epsilon k) - (k/2-\epsilon k) = 2\epsilon k.$$

Similarly, because $\{A,E,F\}$ is an edge, $|\overline{A} \cap \overline{F}| \leqslant 2\epsilon k$. Therefore,

$$|D \cap \overline{A}| \leqslant |\overline{A} \cap \overline{B} \cap F| + |\overline{A} \cap B \cap \overline{F}| + |\overline{A} \cap \overline{B} \cap \overline{F}| \leqslant |\overline{A} \cap \overline{B}| + |\overline{A} \cap \overline{F}| \leqslant 4\epsilon k,$$

and so $|D \cap A| \geqslant |D| - 4\epsilon k = k-4\epsilon k$. □

It follows from Claim 1d that $S(7)$ cannot be a subgraph of H. Otherwise, the edges

$$\{A,D,u\}, \quad \{B,E,u\}, \quad \{C,F,u\}$$

would all appear, and by the definition of H, because the intersections mentioned in Claim 1d are large, it follows that $i \in (A \cap D) \cap (B \cap E) \cap (C \cap F)$. In that case, however, $A \cap B \cap C$ is not empty and so $\{A,B,C\}$ is not an edge. □

It remains only to compute the minimum degree of H. Vertices $S_1,S_2 \in K$ have co-degree at least

$$\frac{k-4\epsilon k}{n}|U|$$

https://doi.org/10.1017/S0963548315000061 Published online by Cambridge University Press

if $|S_1 \cap S_2| \geqslant k - 4\epsilon k$, and at least

$$\frac{k + 4\epsilon k}{n}|U|$$

otherwise. Vertices $u_1, u_2 \in U$ have co-degree at least $|V|$ and vertices $v_1, v_2 \in V$ have co-degree at least $|U|$. All other pairs of vertices have co-degree at least $|U|$ or $|V|$. The minimum co-degree is therefore at least

$$\min\left\{\frac{k(1-4\epsilon)}{k(3/2+\epsilon)}|U|, |U|, |V|\right\} = \left\{\frac{2-8\epsilon}{3+2\epsilon} \cdot \frac{3}{5}N, \frac{3}{5}N, \frac{2}{5}N\right\}.$$

For some choice of ϵ, this is approximately $\frac{2}{5}|V(H)|$. □

Question 9.9. *What is the co-chromatic threshold of the Fano-free hypergraphs? It is between* $2/5$ *and* $1/2$.

Acknowledgements

We would like to thank the referee for detailed and insightful feedback, particularly for pointing out some flaws in our original proof of Theorem 3.1. Repairing this issue led us to a much improved proof approach.

References

[1] Allen, P., Böttcher, J., Griffiths, S., Kohayakawa, Y. and Morris, R. (2013) The chromatic thresholds of graphs. *Adv. Math.* **235** 261–295.

[2] Alon, N., Frankl, P. and Lovász, L. (1986) The chromatic number of Kneser hypergraphs. *Trans. Amer. Math. Soc.* **298** 359–370.

[3] Andrásfai, B., Erdős, P. and Sós, V. T. (1974) On the connection between chromatic number, maximal clique and minimal degree of a graph. *Discrete Math.* **8** 205–218.

[4] Balogh, J. and Lenz, J. Hypergraphs with zero chromatic threshold. Submitted.

[5] Brandt, S. and Thomassé, S. Dense triangle-free graphs are four colorable: A solution to the Erdős-Simonovits problem. *J. Combin. Theory Ser. B*, to appear.

[6] de Caen, D. and Füredi, Z. (2000) The maximum size of 3-uniform hypergraphs not containing a Fano plane. *J. Combin. Theory Ser. B* **78** 274–276.

[7] Erdős, P. and Simonovits, M. (1973) On a valence problem in extremal graph theory. *Discrete Math.* **5** 323–334.

[8] Erdős, P. and Simonovits, M. (1983) Supersaturated graphs and hypergraphs. *Combinatorica* **3** 181–192.

[9] Erdős, P. and Stone, A. H. (1946) On the structure of linear graphs. *Bull. Amer. Math. Soc.* **52** 1087–1091.

[10] Frankl, P. and Füredi, Z. (1983) A new generalization of the Erdős–Ko–Rado theorem. *Combinatorica* **3** 341–349.

[11] Frankl, P. and Füredi, Z. (1989) Extremal problems whose solutions are the blowups of the small Witt-designs. *J. Combin. Theory Ser. A* **52** 129–147.

[12] Frankl, P. and Tokushige, N. (2003) Weighted multiply intersecting families. *Studia Sci. Math. Hungar.* **40** 287–291.

[13] Füredi, Z., Pikhurko, O. and Simonovits, M. (2005) On triple systems with independent neighbourhoods. *Combin. Probab. Comput.* **14** 795–813.

[14] Füredi, Z. and Simonovits, M. (2005) Triple systems not containing a Fano configuration. *Combin. Probab. Comput.* **14** 467–484.
[15] Goddard, W. and Lyle, J. (2011) Dense graphs with small clique number. *J. Graph Theory* **66** 319–331.
[16] Goldwasser, J. On the Turán number of $\{123, 124, 345\}$. Manuscript.
[17] Gowers, W. T. (2007) Hypergraph regularity and the multidimensional Szemerédi theorem. *Ann. of Math.* (2) **166** 897–946.
[18] Keevash, P. (2009) A hypergraph regularity method for generalized Turán problems. *Random Struct. Alg.* **34** 123–164.
[19] Keevash, P. (2011) Hypergraph Turán problems. In *Surveys in Combinatorics 2011*, Vol. 392 of *London Math. Soc. Lecture Note Series*, Cambridge University Press, pp. 83–139.
[20] Keevash, P. and Mubayi, D. (2004) Stability theorems for cancellative hypergraphs. *J. Combin. Theory Ser. B* **92** 163–175.
[21] Keevash, P. and Sudakov, B. (2005) The Turán number of the Fano plane. *Combinatorica* **25** 561–574.
[22] Kleitman, D. J. (1966) Families of non-disjoint subsets. *J. Combin. Theory* **1** 153–155.
[23] Lange, C. E. M. C. and Ziegler, G. M. (2007) On generalized Kneser hypergraph colorings. *J. Combin. Theory Ser. A* **114** 159–166.
[24] Łuczak, T. (2006) On the structure of triangle-free graphs of large minimum degree. *Combinatorica* **26** 489–493.
[25] Łuczak, T. and Thomassé, S. Coloring dense graphs via VC-dimension. Submitted.
[26] Mubayi, D. (2005) The co-degree density of the Fano plane. *J. Combin. Theory Ser. B* **95** 333–337.
[27] Mubayi, D. (2006) A hypergraph extension of Turán's theorem. *J. Combin. Theory Ser. B* **96** 122–134.
[28] Nagle, B., Rödl, V. and Schacht, M. (2006) The counting lemma for regular k-uniform hypergraphs. *Random Struct. Alg.* **28** 113–179.
[29] Pikhurko, O. (2008) An exact Turán result for the generalized triangle. *Combinatorica* **28** 187–208.
[30] Pikhurko, O. (2013) Exact computation of the hypergraph Turán function for expanded complete 2-graphs. *J. Combin. Theory Ser. B* **103** 220–225.
[31] Rödl, V. and Skokan, J. (2004) Regularity lemma for k-uniform hypergraphs. *Random Struct. Alg.* **25** 1–42.
[32] Rödl, V. and Skokan, J. (2006) Applications of the regularity lemma for uniform hypergraphs. *Random Struct. Alg.* **28** 180–194.
[33] Sarkaria, K. S. (1990) A generalized Kneser conjecture. *J. Combin. Theory Ser. B* **49** 236–240.
[34] Sauer, N. (1972) On the density of families of sets. *J. Combin. Theory Ser. A* **13** 145–147.
[35] Simonovits, M. (1974) Extremal graph problems with symmetrical extremal graphs: Additional chromatic conditions. *Discrete Math.* **7** 349–376.
[36] Tao, T. (2006) A variant of the hypergraph removal lemma. *J. Combin. Theory Ser. A* **113** 1257–1280.
[37] Thomassen, C. (2002) On the chromatic number of triangle-free graphs of large minimum degree. *Combinatorica* **22** 591–596.
[38] Thomassen, C. (2007) On the chromatic number of pentagon-free graphs of large minimum degree. *Combinatorica* **27** 241–243.
[39] Vapnik, V. N. and Chervonenkis, A. Y. (1971) Theory of uniform convergence of frequencies of events to their probabilities and problems of search for an optimal solution from empirical data. *Avtomat. i Telemeh.* (2) 42–53.
[40] Ziegler, G. M. (2002) Generalized Kneser coloring theorems with combinatorial proofs. *Invent. Math.* **147** 671–691.