

SIMPLICIAL TURÁN PROBLEMS

DAVID CONLON, SIMÓN PIGA, AND BJARNE SCHÜLKE

ABSTRACT. A simplicial complex H consists of a pair of sets (V, E) where V is a set of vertices and $E \subseteq \mathcal{P}(V)$ is a collection of subsets of V closed under taking subsets. Given a simplicial complex F and $n \in \mathbb{N}$, the extremal number $\text{ex}(n, F)$ is the maximum number of edges that a simplicial complex on n vertices can have without containing a copy of F . We initiate the systematic study of extremal numbers in this context by asymptotically determining the extremal numbers of several natural simplicial complexes. In particular, we asymptotically determine the extremal number of a simplicial complex for which the extremal example has more than one incomplete layer.

§1. INTRODUCTION

Turán's paper [23] in which he showed that complete, balanced, $(r - 1)$ -partite graphs maximise the number of edges among all K_r -free graphs is often seen as one of the starting points of extremal combinatorics. Already in this paper, he asked for similar results for hypergraphs. More than 80 years later, still little is known in this regard.

A hypergraph $H = (V, E)$ consists of a vertex set V and an edge set $E \subseteq \mathcal{P}(V) = \{e \subseteq V\}$. We say that H is k -uniform (or a k -graph) if $E \subseteq V^{(k)} = \{e \subseteq V : |e| = k\}$. The extremal number for $n \in \mathbb{N}$ and a k -graph F is the maximum number of edges in an F -free k -graph on n vertices and is denoted by $\text{ex}^{(k)}(n, F)$. The Turán density of F is then $\pi(F) = \lim_{n \rightarrow \infty} \text{ex}^{(k)}(n, F) \binom{n}{k}^{-1}$ (this limit was shown to exist by monotonicity in [16]). The problem of determining the extremal numbers or the Turán densities of hypergraphs is often referred to as the Turán problem. For graphs, this problem is reasonably well understood, as a result of Erdős, Stone, and Simonovits [9, 10], generalising Turán's theorem, says that the Turán density of any graph F is $\frac{\chi(F)-2}{\chi(F)-1}$, where $\chi(F)$ is the chromatic number of F .

2020 *Mathematics Subject Classification.* 05C65, 05C35, 05D05, 05D99.

Key words and phrases. Turán problems, hypergraphs, simplicial complexes.

Data availability statement. There are no additional data beyond that contained within the main manuscript.

The research leading to these results was supported by NSF Award DMS-2054452 (D. Conlon) and EPSRC Grant No. EP/V002279/1 (S. Piga).

In contrast, for hypergraphs, only a few Turán densities are known, one prominent example being that of the Fano plane [2, 8, 15, 17]. Even the Turán density of the complete 3-graph on four vertices, $K_4^{(3)}$, is still not known, although it is widely conjectured to be $5/9$. In fact, the problem is already open for the 3-graph that consists of three edges on four vertices, denoted by $K_4^{(3)-}$. For more on the hypergraph Turán problem, we refer the interested reader to the survey by Keevash [18].

In this work, we consider the Turán problem for (abstract) simplicial complexes. A hypergraph $H = (V, E)$ is a simplicial complex if for all $e \in E$ and $e' \subseteq e$, we have $e' \in E$. In other words, a simplicial complex is a hypergraph whose edge set is closed under taking subsets (we also say that E is hereditary). The k -uniform layer of a simplicial complex $H = (V, E)$ is $H_k = (V, E^{(k)})$, where $E^{(k)} = \{e \in E : |e| = k\}$. In line with the geometric interpretation of simplicial complexes, we define the dimension of H to be

$$\dim(H) = \max\{k \geq 0 : E^{(k)} \neq \emptyset\} - 1.$$

For brevity, when describing the edge set of a simplicial complex, we will generally not mention the induced edges, i.e., the edges which are properly contained in another edge.

The extremal number for $n \in \mathbb{N}$ and a simplicial complex F is the maximum number of edges in an F -free simplicial complex on n vertices and is denoted by $\text{ex}(n, F)$. For a family of simplicial complexes \mathcal{F} , we define $\text{ex}(n, \mathcal{F})$ analogously. We are by no means the first to study the extremal problem for simplicial complexes. For example, already in 1983, Frankl [12] observed that $\text{ex}(n, F) = \sum_{i=0}^{s-1} \binom{n}{i}$ for $F = ([s], \mathcal{P}([s]))$ and derived the celebrated Sauer–Shelah–Vapnik–Chervonenkis lemma [20, 21, 24] as a corollary. However, we believe that we are the first to study the problem systematically, asymptotically determining the extremal numbers of several natural simplicial complexes.

As a first attempt, one might hope for some way to reduce the Turán problem for simplicial complexes to that for uniform hypergraphs. Indeed, given a simplicial complex F of dimension $k - 1$, let H be the simplicial complex with $|V(H)| = n$, $E^{(i)} = V(H)^{(i)}$ for $0 \leq i \leq k - 1$ and $E^{(k)}$ the edge set of a k -graph on $V(H)$ that is F_k -free. Then H is clearly F -free, yielding the following general lower bound.

Observation 1.1. If F is a simplicial complex of dimension $k - 1$, then

$$\text{ex}(n, F) \geq \text{ex}^{(k)}(n, F_k) + \sum_{i=0}^{k-1} \binom{n}{i}.$$

We will show that this bound is asymptotically optimal for several simplicial complexes. However, this is not always the case. To say more, we recall an old conjecture of Hajnal (originally formulated in terms of traces of sets rather than simplicial complexes). Given

positive integers k and d with $k > d$, let $\mathcal{F}(k, d)$ be the family of all simplicial complexes with vertex set $[k]$ and $1 + \sum_{i=0}^d \binom{k}{i}$ edges. The problem of determining $\text{ex}(n, \mathcal{F}(k, d))$ can be seen as a variant of the well-known Brown–Erdős–Sós problem [22] for simplicial complexes. Hajnal’s conjecture (see [6, 14]) is then that $\text{ex}(n, \mathcal{F}(k, d)) = \sum_{i=0}^d \binom{n}{i}$, in keeping with the lower bound in Observation 1.1. Bollobás and Radcliffe [5] proved this conjecture for $k = 4$ and $d = 2$. However, in full generality, it was disproved by Frankl [11] and later, in a very strong sense, by Bukh and Goac [7]. In particular, from their result it follows that for every $d \geq 2$ and every exponent $\ell \in \mathbb{N}$, there is a sufficiently large k such that $\text{ex}(n, \mathcal{F}(k, d)) = \Omega(n^\ell)$. However, for $k \geq 5$ and $d \geq 2$, the problem of determining the asymptotics of $\text{ex}(n, \mathcal{F}(k, d))$ remains wide open. Indeed, it seems to be significantly harder to determine the extremal number of a simplicial complex if it does not agree with the lower bound in Observation 1.1.

Our main result, Theorem 1.2 below, breaks through this barrier, asymptotically determining the extremal number of a simplicial complex B for which the extremal number is not given by the bound in Observation 1.1. To the best of our knowledge, this is the first result of this type and shows that the extremal behaviour of simplicial complexes can be genuinely different from that of the uniform hypergraph given by the largest layer.

Theorem 1.2. *For $B = (\{v_1, \dots, v_5\}, \{v_3v_4, v_1v_2v_3, v_1v_4v_5\})$, there is a constant $C \in \mathbb{N}$ such that*

$$\text{ex}(n, B) \leq \frac{9}{8} \binom{n}{2} + \frac{21}{16}n + C$$

for every $n \geq 5$.

As shown by the following example, this bound is best possible up to the constant.

Example 1.3. Let $n \in \mathbb{N}$ with $n \equiv 6 \pmod{8}$, let $V = V_1 \cup V_2$ for some sets V_1 and V_2 with $|V_2|$ even and $|V_1| + |V_2| = n$, and let \mathcal{M} be a perfect 2-uniform matching on V_2 . Furthermore, let

$$E' = \{abc : ab \in \mathcal{M} \text{ and } c \in V_1\} \cup V_1^{(2)}$$

and let E be the downward closure of E' . Then set $H = (V, E)$ and note that $B \not\subseteq H$ and

$$|E| = \frac{n - |V_1|}{2} |V_1| + (n - |V_1|) |V_1| + \binom{|V_1|}{2} + \frac{n - |V_1|}{2} + n + 1.$$

The right-hand side is maximised if $|V_1| = \frac{3}{4}n - \frac{1}{2}$, in which case it is equal to $\frac{9}{8} \binom{n}{2} + \frac{21}{16}n + \frac{5}{4}$.

The proof of Theorem 1.2 is the longest and most difficult part of this paper. We believe that the approach used for this proof might have further applications for hypergraph Turán-type problems. Roughly speaking, it proceeds by gaining more and more structure on an

extremal B -free simplicial complex, ultimately yielding the same structure as Example 1.3. This is done by introducing and analysing a parameter ρ , which is defined for every vertex and depends on its link, and symmetrising a subset of the vertices determined by ρ .

We now state our results where, up to lower order terms, the extremal number is given by the lower bound in Observation 1.1. We begin with the comparatively simple case of two disjoint 3-edges connected by a 2-edge.

Theorem 1.4. *For the simplicial complex $F = (\{v_1, \dots, v_6\}, \{v_1v_2v_3, v_4v_5v_6, v_1v_4\})$, there is $C \in \mathbb{N}$ such that $n^2 - n + 2 \leq \text{ex}(n, F) \leq n^2 - n + C$.*

Given a 3-graph F and $A \subseteq V(F)$, let F^+ be the simplicial complex obtained from F by adding a new vertex a and all 2-edges from a to A (and then taking the downward closure). Our next theorem follows from a more general result, Theorem 4.1, that we prove in Section 4. It gives a general condition under which the extremal number of F^+ is approximately equal to the bound in Observation 1.1 provided we also require that the host simplicial complex is 2-dimensional. Let us call a 3-graph *proper* if for every $\varepsilon > 0$, there is some $n_0 \in \mathbb{N}$ such that for every $n \geq n_0$, we have

$$(1 + \varepsilon) [\text{ex}^{(3)}(n, F) - \text{ex}^{(3)}(n - 1, F)] \geq \frac{\text{ex}^{(3)}(n, F)}{n}.$$

For instance, if $\text{ex}^{(3)}(n, F) \approx Cn^\alpha$ for some $C > 0$ and $\alpha \geq 1$, then F is proper. Moreover, by duplicating a vertex that has at least the average degree in an extremal example for F on $n - 1$ vertices (combined with a standard double counting argument from [16]), one can easily see that if the shadow of F is complete, then F is proper.

Theorem 1.5. *For every proper 3-graph F , $A \subseteq V(F)$ and $\eta, \varepsilon > 0$, there is $n_0 \in \mathbb{N}$ such that the following holds. If a 2-dimensional simplicial complex H on $n \geq n_0$ vertices satisfies*

$$\frac{\text{ex}^{(3)}(n, F)}{n^2} + 1 \geq \left(\frac{(|A| - 1)}{|A|} + \eta \right) \left[1 + \frac{\Delta^{(3)}}{2} \right], \quad (1.1)$$

where $\Delta^{(3)} = \max_{xy \in V(H)^{(2)}} d^{(3)}(xy)$, and $e(H) \geq (1 + \varepsilon) [\text{ex}^{(3)}(n, F) + \binom{n}{2}]$, then H contains a copy of F^+ .

The following corollary of Theorem 4.1 (using that the extremal number of the 3-graph with two edges on four vertices has been determined in [13]) asymptotically determines the extremal number of another natural simplicial complex on five vertices.

Corollary 1.6. *For the simplicial complex $F = (\{v_1, \dots, v_5\}, \{v_1v_2v_3, v_2v_3v_4, v_2v_5, v_3v_5\})$ and $\varepsilon > 0$, there is $n_0 \in \mathbb{N}$ such that $\text{ex}(n, F) = \left(\frac{4}{3} \pm \varepsilon\right) \binom{n}{2}$ for every $n \geq n_0$.*

Finally, we determine the asymptotics for the extremal number of the simplicial complex on four vertices given by two 3-edges that intersect in two vertices and a 2-edge between the two vertices which do not lie in the intersection.

Theorem 1.7. *For the simplicial complex $F = (\{v_1, v_2, v_3, v_4\}, \{v_1v_2v_3, v_2v_3v_4, v_1v_4\})$ and $\varepsilon > 0$, there is $n_0 \in \mathbb{N}$ such that $ex(n, F) = \left(\frac{2}{9} \pm \varepsilon\right) \binom{n}{3}$ for every $n \geq n_0$.*

Notation. For a simplicial complex $H = (V, E)$, $S \subseteq V$, and $i \in \mathbb{N}$, we define the i -uniform degree of S by $d^{(i)}(S) = |\{e \in E : S \subseteq e, |e| = i\}|$. Omitting parentheses around 1- and 2-sets, we have, in particular, that $d^{(i)}(v) = |\{e \in E : v \in e, |e| = i\}|$ for $v \in V$. We also write $d^{(\geq i)}(v) = \sum_{j \geq i} d^{(j)}(v)$. The degree of v counts all edges containing v except the singleton, that is, $d(v) = \sum_{i \in \mathbb{N}, i \geq 2} d^{(i)}(v)$. The neighbourhood of v is $N(v) = \{w \in V \setminus v : \exists e \in E : v, w \in e\}$ and, more generally, the neighbourhood of a set S is $N(S) = \{w \in V \setminus S : \exists e \in E : S \subseteq e, w \in e\}$. When considering simplicial complexes and the context is clear, we will usually only mention those edges whose downward closure is the edge set. For a vertex $v \in V$, let L_v be the link of v in H , i.e., the simplicial complex with vertex set $V - v$ and edge set $\{e - v : v \in e \in E\}$. Finally, we write $E^{(i)}(H) = \{e \in E : |e| = i\}$, $e^{(i)}(H) = |E^{(i)}|$, and $e(H) = |E|$. Similarly, we write $v(H) = |V|$.

§2. PROOF OF THEOREM 1.2

Let $H = (V, E)$ be a simplicial complex with $|V| = n$, $B \not\subseteq H$, and assume for the sake of contradiction that $|E| \geq \frac{9}{8} \binom{n}{2} + \frac{21}{16}n + C$, where C is some sufficiently large constant.

Roughly speaking, we proceed as follows. By induction on the number of vertices, we may assume some minimum degree. The proof is then centered around a parameter $\varrho(v)$, defined for every vertex v . Depending on the role that v plays in the link of another vertex w , we know how $\varrho(v)$ and $\varrho(w)$ relate. This allows us to determine the rough structure of H . By continuing to analyse ϱ and by symmetrising a subset of the vertices determined by ϱ , we can simplify this structure further. Finally, we obtain a B -free simplicial complex which equals the extremal example aside from maybe the sizes of the two partition classes and has at least as many edges as the initial H . A simple optimisation then finishes the proof.

Let us formalise the proof by performing an induction on n . Let $n_0 \in \mathbb{N}$ be sufficiently large and choose $C \in \mathbb{N}$ large enough that every simplicial complex on $n < n_0$ vertices with at least $\frac{9}{8} \binom{n}{2} + \frac{21}{16}n + C$ edges contains a copy of B . We take those cases as the base of our induction. The remainder of this section is devoted to the induction step.

For $n \geq n_0$, assume that every simplicial complex on $n - 1$ vertices with at least $\frac{9}{8} \binom{n-1}{2} + \frac{21}{16}(n-1) + C$ edges contains B . Let $H = (V, E)$ be a simplicial complex with $|V| = n$, $B \not\subseteq H$,

and assume for the sake of contradiction that $|E| \geq \frac{9}{8}\binom{n}{2} + \frac{21}{16}n + C$. If there is a vertex $v \in V$ with $d(v) \leq \frac{9}{8}(n-1) + \frac{5}{16}$, we have that $|E(H-v)| \geq \frac{9}{8}\binom{n}{2} + \frac{21}{16}n + C - d(v) - 1 \geq \frac{9}{8}\binom{n-1}{2} + \frac{21}{16}(n-1) + C$ and so, by induction, $H-v$ contains a copy of B . Therefore, we may assume that for every $v \in V$,

$$d(v) > \frac{9}{8}(n-1) + \frac{5}{16}. \quad (2.1)$$

Clearly, H cannot contain any edges of size at least five. Thus, the links consist only of 3-edges, 2-edges, and 1-edges (and the empty set). Furthermore, we have the following.

Claim 2.1. *For every $v \in V$, there is no 2-uniform path with three edges in L_v .*

Proof. Assume that there is a 2-uniform path $abcd$ in the link of some $v \in V$. Since H is a simplicial complex, we have $bc \in E$ and, by the definition of the link, we have $abv, cdv \in E$. Thus, there is a copy of B in $H[a, b, c, d, v]$, a contradiction. \square

For a simplicial complex F and a subcomplex $F' \subseteq F$, we say that F' is *isolated* in F if F does not contain an edge containing both a vertex of F' and a vertex in $V(F) \setminus V(F')$. Claim 2.1 means that every link consists only of: isolated vertices, isolated (2-)edges, isolated induced (2-uniform) stars with at least two edges, isolated (2-uniform) triangles, and potentially isolated 3-edges. However, the next claim states that H does not have any 4-edges, so the links do not in fact contain any 3-edges.

Claim 2.2. *There are no 4-edges in H .*

Proof. Assume that $e = v_1v_2v_3v_4$ is a 4-edge. Note that no 3-edge can intersect e in exactly two vertices, as otherwise $B \subseteq H$. For $i \in [4]$, let N_i be the set of vertices in $V \setminus e$ which are contained in a 3-edge with v_i . Note that for all distinct $i, j \in [4]$, we have that v_i has no 2-edge to any vertex in N_j (as otherwise, again, $B \subseteq H$), so in particular N_i and N_j are disjoint. Without loss of generality, we may assume that $|N_1| \leq \dots \leq |N_4|$. Since all 2-edges between v_1 and the vertices in $N_2 \cup N_3 \cup N_4$ are missing, $d^{(2)}(v_1) \leq (n-1) - (|N_2| + |N_3| + |N_4|)$. Together with (2.1), this implies that

$$d^{(3)}(v_1) + d^{(4)}(v_1) > |N_2| + |N_3| + |N_4| + \frac{n-1}{8} + \frac{5}{16}.$$

On the other hand, by Claim 2.1, $L_{v_1}[N_1]$ does not contain any 2-uniform path on three edges. It is easy to check that this means the number of 2- and 3-edges in L_{v_1} is at most $4|N_1|/3$. Adding the 3- and 4-edges in e containing v_1 , we obtain

$$d^{(3)}(v_1) + d^{(4)}(v_1) \leq 4 + \frac{4}{3}|N_1|,$$

which is a contradiction for n sufficiently large. \square

Definition and basic properties of ϱ . Let us now introduce the central parameter of the proof. Let $\varphi(v)$ denote the number of isolated 2-edges in L_v and define

$$\varrho(v) = e^{(2)}(L_v) + \varphi(v) - 2.$$

The idea behind this parameter is that it is (a lower bound on) the number of 2-edges that must be missing at a vertex w if w appears in an isolated 2-edge in L_v . We make this precise with the following claim.

Claim 2.3. *Let $v, w, x \in V$ be such that wx is an isolated edge in L_v . Then $d^{(2)}(w) \leq n - 1 - \varrho(v)$. Moreover,*

$$\varrho(w) \geq d^{(3)}(w) - 2 > \frac{n-1}{8} - \frac{27}{16} + \varrho(v).$$

Proof. Let $ab \in E^{(2)}(L_v) \setminus \{wx\}$ and note that since wx is an isolated edge in L_v , we have that a, b, w, x are distinct. If $aw \in E$, then $H[a, b, v, w, x]$ contains a copy of B , a contradiction. Thus, the only 2-edge in H formed by w and a vertex that is contained in a 2-edge of L_v is wx . In other words, w is ‘missing’ a 2-edge to every vertex spanned by a 2-edge in $E^{(2)}(L_v) \setminus \{wx\}$.

Recall that, by Claims 2.1 and 2.2, L_v consists only of isolated vertices, isolated 2-edges, isolated induced (2-uniform) stars with at least two edges, and isolated (2-uniform) triangles. Clearly, every isolated 2-edge in L_v spans exactly two vertices. Moreover, we have that isolated (2-uniform) triangles and stars span at least as many vertices as they have 2-edges. Therefore, the total number of vertices spanned by $E^{(2)} \setminus \{wx\}$ is at least $e^{(2)}(L_v) + \varphi(v) - 2 = \varrho(v)$. By the argument in the paragraph above, each of those vertices corresponds to a missing 2-edge in w . This yields the first inequality of the claim.

The moreover part simply follows from the first part, (2.1), and the definition of ϱ . \square

The following claim is in a sense the reverse of the previous one. For $v, w \in V$, we say that w is a *tip of a star in L_v* , and write $w \leftarrow v$, if there are $x, y \in V$ such that $wx, xy \in E(L_v)$.

Claim 2.4. *Let $v, w \in V$ be such that $w \leftarrow v$. Then $d^{(2)}(v) \leq n - \varrho(w)$. Moreover,*

$$\varrho(v) \geq d^{(3)}(v) - 2 > \frac{n-1}{8} - \frac{43}{16} + \varrho(w).$$

Proof. Let $x, y \in V$ be such that $wx, xy \in E(L_v)$ and note that $vx \in E(L_w)$. If vx is an isolated edge in L_w , the statement follows from Claim 2.3. If there is some $y' \in V \setminus \{v, x, y\}$ such that $vy' \in E(L_w)$ or $xy' \in E(L_w)$, then $H[v, w, x, y, y']$ contains B . Thus, we are left with the remaining case that both v and x lie in at most two 2-edges in L_w , namely, vx and possibly vy and xy . As in the proof of Claim 2.3, if there is $ab \in E^{(2)}(L_w)$

with $ab \cap \{v, x, y\} = \emptyset$ and $av \in E$, then $B \subseteq H[a, b, v, w, x]$. That is, v is missing a 2-edge to every vertex spanned by an edge in $E^{(2)}(L_w) \setminus \{vx, vy, xy\}$. Again, as in the proof of Claim 2.3, there are at least $e^{(2)}(L_w) - 3 + \varphi(w) = \varrho(w) - 1$ such vertices. This yields the first part of the claim.

The moreover part again follows from the first part, (2.1), and the definition of ϱ . \square

At this point, we already begin to see some resemblance of H to the extremal example, namely, in the structure of the links of those vertices with extremal ϱ .

Corollary 2.5. *Let $v \in V$.*

- (1) *If $\varrho(v) = \min_{w \in V} \varrho(w)$, then the 2-edges of L_v form a matching of size larger than $\frac{n-1}{8} + \frac{5}{16}$.*
- (2) *If $\varrho(v) = \max_{w \in V} \varrho(w)$, then the 2-edges of L_v do not contain any isolated edges.*
- (3) *The link of v consists only of isolated vertices, isolated 2-edges, and isolated induced (2-uniform) stars with at least two edges.*

Proof. For (1), note that if the 2-edges of L_v do not form a matching, then there would be a vertex w that is a tip of a star in L_v , implying that $\varrho(w) < \varrho(v)$ by the second inequality in Claim 2.4. Moreover, (2.1) implies that $d^{(3)}(v) > \frac{n-1}{8} + \frac{5}{16}$.

For (2), note that if there is an isolated 2-edge xy in L_v , then $\varrho(x) > \varrho(v)$ by Claim 2.3.

To see (3), recall that we already know that L_v consists only of isolated vertices, isolated 2-edges, isolated induced (2-uniform) stars with at least two edges, and isolated (2-uniform) triangles. Assume now that there is a triangle with vertices w, x , and y in L_v . Then considering the link of w yields $x \leftarrow w$ (since x and y are tips of a star centered in v in L_w) and thus $\varrho(x) < \varrho(w)$ by Claim 2.4. However, considering the link of x similarly gives $w \leftarrow x$ and so $\varrho(w) < \varrho(x)$, a contradiction. \square

Next, we bring the structure closer to the extremal example by performing some symmetrisations.

Symmetrisation of vertices. We begin by defining V_1 as the set of vertices in H for which the 2-edges in the link form a matching. Note that, by Corollary 2.5 (1), we have $V_1 \neq \emptyset$. Moreover, no two vertices of V_1 are contained in the same 3-edge. Indeed, if $v, w \in V_1$ are such that $wx \in L_v$ for some vertex $x \in V$, then, by Claim 2.3, we have $\varrho(w) > \varrho(v)$. But since $vx \in L_w$ as well, we have $\varrho(w) < \varrho(v)$, a contradiction.

Now pick $v \in V_1$ with $d(v)$ maximal among the vertices in V_1 . Clearly, if we replace each vertex in V_1 (one by one) by a copy of v and add all 2-edges between these vertices, then the total number of edges does not decrease. Moreover, recalling that the 2-edges of L_v form a matching, it is not hard to check that no new copies of B are created in this way.

Thus, from now on we can consider H to be the resulting simplicial complex. In particular, all vertices $v \in V_1$ have the same link L_v whose 2-edges form a fixed matching M .

Note that, by (2.1), we have $|M| > \frac{n-1}{8} + \frac{5}{16}$ and therefore, by the definition of $\varrho(\cdot)$, we have, for all $v \in V_1$, that

$$\varrho(v) > 2 \left(\frac{n-1}{8} + \frac{5}{16} \right) - 2 = \frac{n-1}{4} - \frac{11}{8}. \quad (2.2)$$

The following notation will help us as we make more use of Claims 2.3 and 2.4. For $v, w \in V$ and $i \in \mathbb{N}$, write $v \stackrel{i}{\leftarrow} w$ if there are vertices $v = v_0, \dots, v_i = w$ such that, for all $j \in [i]$, we have that $v_{j-1} \leftarrow v_j$. Furthermore, for $S \subseteq V$, write $S \stackrel{i}{\leftarrow} w$ if there is some $v \in S$ such that $v \stackrel{i}{\leftarrow} w$. Note that $v \stackrel{1}{\leftarrow} w$ just means $v \leftarrow w$.

Claim 2.6. *There is no $v \in V$ and $i \geq 3$ such that $V_1 \stackrel{i}{\leftarrow} v$.*

Proof. Assume for the sake of contradiction that $v \in V$ and $i \in \mathbb{N}$ with $i \geq 3$ such that $V_1 \stackrel{i}{\leftarrow} v$. Then there are $v_0 \in V_1$ and $v_1, v_2, v_3 \in V$ (with v_3 being v if $i = 3$) such that $v_0 \leftarrow v_1 \leftarrow v_2 \leftarrow v_3$. By (2.2), we know that $\varrho(v_0) > \frac{n-1}{4} - \frac{11}{8}$. Applying Claim 2.4 twice yields $\varrho(v_2) > \frac{n-1}{4} - \frac{11}{8} + 2 \left(\frac{n-1}{8} - \frac{43}{16} \right) = \frac{n-1}{2} - \frac{27}{4}$. This means, again by Claim 2.4, that $d^{(2)}(v_3) \leq n - \varrho(v_2) \geq n - \frac{n-1}{2} + \frac{27}{4} = \frac{n-1}{2} + \frac{31}{4}$ and

$$d^{(3)}(v_3) > \frac{n-1}{8} - \frac{43}{16} + 2 + \varrho(v_2) \geq \frac{5(n-1)}{8} - \frac{119}{16}.$$

It follows that $d^{(3)}(v_3) > d^{(2)}(v_3)$ for n sufficiently large. However, Corollary 2.5 (3) implies that $d^{(3)}(v_3) \leq d^{(2)}(v_3) - 1$, yielding the required contradiction. \square

Next, guided by the intuition coming from the extremal example, we show that the vertices in $V(M)$ (which we might also call V_2) are exactly the vertices v for which $V_1 \stackrel{1}{\leftarrow} v$.

Lemma 2.7. *For $v \in V$, the following are equivalent:*

- $v \in V(M)$.
- $V_1 \stackrel{1}{\leftarrow} v$ holds and $V_1 \stackrel{i}{\leftarrow} v$ does not hold for $i \geq 2$.

Proof. “ \Leftarrow ”: Let $x \in V_1$ be a tip of a star in L_v . This means in particular that v is contained in a 2-edge in L_x . But any such edge is in M , so that $v \in V(M)$.

“ \Rightarrow ”: Let $w \in V$ be such that $vw \in M$ and let $x \in V_1$. In particular, vw is an isolated edge in L_x , so that $\varrho(v) > \varrho(x) + \frac{n-1}{8} - \frac{27}{16}$ by Claim 2.3. Note that $wx \in E(L_v)$. If wx is an isolated edge in L_v , then another application of Claim 2.3 would yield $\varrho(x) > \varrho(v) + \frac{n-1}{8} - \frac{27}{16}$, a contradiction. Therefore, wx is not an isolated edge in L_v , meaning that there is another 2-edge intersecting wx . Since vw is an isolated edge in L_x , such an edge cannot contain x , so it must contain w , implying that $d_{L_v}^{(2)}(w) \geq 2$. In other words, x is a tip of a star in L_v , i.e., $x \stackrel{1}{\leftarrow} v$.

By Claim 2.6, it is enough to show that $V \stackrel{i}{\leftarrow} v$ does not hold for any $i \geq 2$. Assume for the sake of contradiction that it does hold, i.e., that there is some $x \in V_1$ and $r \in V$ such that $x \leftarrow r \leftarrow v$.

Note that $V_1 \stackrel{1}{\leftarrow} r$ holds and $V_1 \stackrel{i}{\leftarrow} r$ does not hold for $i \geq 2$, since otherwise we would have $V_1 \stackrel{i+1}{\leftarrow} v$ for some $i \geq 2$, contradicting Claim 2.6. Hence, by “ \Leftarrow ”, we know that $r \in V(M)$. Observe that any 2-edge in H between vertices in distinct edges of M would yield a copy of B and that $r \leftarrow v$ in particular implies that $rv \in E(H)$. Thus, $rv \in M$ (and so $r = w$), implying that $rx \in E(L_v)$. Finally, recall that r is a tip of a star in L_v and note that x cannot be a vertex of this star, since rv is an isolated 2-edge in L_x . Therefore, in L_v , the 2-edge rx is a pendant edge to this star. However, this gives a (2-uniform) path with three edges in L_v , contradicting Claim 2.1. \square

Gaining further structure. Recall that Corollary 2.5 (3) states that the link of any vertex consists only of isolated vertices, isolated 2-edges, and isolated (2-uniform) induced stars with at least two edges. The following key lemma says that vertices of $V(M)$ contain exactly one star in their links.

Lemma 2.8. *If $w \in M$, then L_w contains exactly one isolated (2-uniform) induced star with at least two edges and the centre of this star is w .*

We first prove the following statement.

Claim 2.9. *If $v \in V$ and $w \notin V_1$ with $w \leftarrow v$, then $V_1 \stackrel{i}{\leftarrow} v$ for some $i \geq 2$.*

Proof. Assume that the statement is not true and let v and w be a counterexample where $\varrho(w)$ is minimal. As $w \notin V_1$, there is some vertex with 2-degree at least 2 in L_w and, therefore, some $w' \in V$ with $w' \leftarrow w$. Note that by Claim 2.4, we have $\varrho(w') < \varrho(w)$. If $w' \in V_1$, then $V_1 \stackrel{2}{\leftarrow} v$, contrary to our assumption that v and w are a counterexample to the statement. Thus, $w' \notin V_1$ and since $\varrho(w)$ was chosen to be minimal among all counterexamples, we have that w and w' are not a counterexample, meaning that there is some $i \geq 2$ such that $V_1 \stackrel{i}{\leftarrow} w$. But then $V_1 \stackrel{i+1}{\leftarrow} v$, contrary to our assumption. \square

Let us now prove Lemma 2.8.

Proof of Lemma 2.8. First, note that a vertex in $V(M)$ must have a (2-uniform) star with at least two edges in its link for otherwise it would be in V_1 . To show that v has at most one (2-uniform) star with at least two edges in its link, assume the opposite. Then there are $x_i, y_i \in V$ for $i \in [3]$ such that $x_1 \neq y_1$ and $x_1x_2, x_1x_3, y_1y_2, y_1y_3 \in E(L_v)$.

First, let us observe that $x_2, x_3, y_2, y_3 \in V_1$. Indeed, assume that one of x_2, x_3, y_2, y_3 is not contained in V_1 , say, without loss of generality, x_2 . By Claim 2.9 and the fact that $x_2 \leftarrow v$,

we would then have $V_1 \stackrel{i}{\leftarrow} v$ for some $i \geq 2$. But, by Lemma 2.7, this contradicts $v \in V(M)$. Therefore, $x_2, x_3, y_2, y_3 \in V_1$. But then $vx_1 \in E(L_{x_2})$ and $vy_1 \in E(L_{y_2})$ contradict the facts that x_2 and y_2 have the same link and v lies in a matching edge in this link.

Lastly, let us show that w is the centre of a star with at least two edges. Recall that the vertices in V_1 are symmetrised, so that, in particular, every vertex in V_1 forms an edge with w in L_v . Hence, it is enough to show that $|V_1| \geq 2$. For this consider a vertex $a \in V(M)$ and recall that there must be a 2-uniform star with at least two edges in L_a . If one of the leaves in this star is not in V_1 , then Claim 2.9 implies that $V_1 \stackrel{i}{\leftarrow} a$ for some $i \geq 2$. But, by Lemma 2.7, this would contradict $a \in V(M)$. Thus, we indeed have $|V_1| \geq 2$ and w is the centre of a star with at least two edges. \square

Remark 2.10. In the remainder of this subsection, we will show how to symmetrise the pairs that form edges in M . This is not strictly necessary for the proof, so the reader may already jump to the next subsection. However, we believe that this technique of iterative partial symmetrisation, where we symmetrise certain vertex subsets and certain subsets of the pairs of vertices at appropriate junctures, has the potential to be more broadly useful in the study of hypergraph Turán problems, so we decided to include it regardless.

We symmetrise the pairs that form edges in M by essentially replacing any pair by a pair of larger degree as long as there still exist pairs which do not ‘look the same’. More precisely, for distinct $x, y \in V$, we set

$$\begin{aligned} E_x(x, y) &= \{e \setminus x : e \in E(H), e \cap \{x, y\} = x\}, \\ E_y(x, y) &= \{e \setminus y : e \in E(H), e \cap \{x, y\} = y\}, \text{ and} \\ E_{xy}(x, y) &= \{e \setminus \{x, y\} : e \in E(H), e \cap \{x, y\} = \{x, y\}\}. \end{aligned}$$

When the pair x, y is clear from context we will omit the ‘ (x, y) ’ from the notation. If there are $xy, x'y' \in M$ such that $E_x \neq E_{x'}$, $E_y \neq E_{y'}$, or $E_{xy} \neq E_{x'y'}$, we perform the following operation on H . Without loss of generality, assume that

$$|E_x| + |E_y| + |E_{xy}| \geq |E_{x'}| + |E_{y'}| + |E_{x'y'}|. \quad (2.3)$$

Delete the vertices x' and y' and all edges containing at least one of them (these are precisely $E_{x'}$, $E_{y'}$, and $E_{x'y'}$). Then add vertices u and v as well as the sets of edges

$$E_u^+ = \{e \cup u : e \in E_x\}, \quad E_v^+ = \{e \cup v : e \in E_y\}, \quad \text{and} \quad E_{uv}^+ = \{e \cup \{u, v\} : e \in E_{xy}\}$$

and call the resulting simplicial complex H' . Intuitively, this means that we replace the pair $x'y'$ by a copy of the pair xy . We will say that u is a copy of x and that v is a copy

of y .¹ To see that this operation is well defined, note that no edge in E_x , E_y , or E_{xy} contains a vertex from the pair $x'y'$ and, therefore, we never try to add new edges containing the deleted vertices x' and y' . Indeed, if, for example, x' is contained in an edge of E_x , then x and x' are in a common edge of H , so, since H is hereditary, $xx' \in E(H)$. But then, since $xy, x'y' \in M$, any vertex in V_1 together with x, x', y, y' forms a copy of B , a contradiction. Note that, because of (2.3), we do not decrease the number of edges, i.e., we have $|E(H')| \geq |E(H)|$ (and $|V(H')| = |V(H)| = n$).

We now show that this symmetrisation does not introduce a copy of B , i.e., $B \not\subseteq H'$. Indeed, assume that $B \subseteq H'$ on some vertices v_1, \dots, v_5 with $v_1v_2v_3, v_1v_4v_5, v_3v_4 \in E(H')$. Since $B \not\subseteq H$, we have $\{x, y\} \cap \{v_1, \dots, v_5\} \neq \emptyset$ and $\{u, v\} \cap \{v_1, \dots, v_5\} \neq \emptyset$.

First assume that $x, u \in \{v_1, \dots, v_5\}$ or $y, v \in \{v_1, \dots, v_5\}$, say, without loss of generality, $x, u \in \{v_1, \dots, v_5\}$. Because there is no edge in H' that contains both a vertex from xy and one from uv , we may assume that $v_2 = x$ and $u \in \{v_4, v_5\}$. For the same reason, $v_1 \notin \{x, y, u, v\}$. We will now argue that in H the vertex v_1 is the centre of a 2-uniform star with at least two edges in L_x . But, since $v_1 \neq y$, this will contradict Lemma 2.8. If $\{u, v\} = \{v_4, v_5\}$, then $v_3v_4 \in E(H')$ implies that $v_3 \neq y$. Recalling that uv is a copy of xy , we see that $v_1xy \in E(H)$, whereby $v_1 \neq y$ is the centre of a 2-uniform star with at least two edges in L_x (namely, v_1v_3 and v_1y). This yields the aforementioned contradiction and so we infer that $v \notin \{v_4, v_5\}$. But then, again using that u is a copy of x , we see that $v_1 \neq y$ is still the centre of a 2-uniform star with at least two edges in L_x (namely, v_1v_3 and v_1z , where $z \in \{v_4, v_5\} \setminus u$). Once again, this yields a contradiction.

The remaining cases are where $x, v \in \{v_1, \dots, v_5\}$ and $y, u \notin \{v_1, \dots, v_5\}$ or $y, u \in \{v_1, \dots, v_5\}$ and $x, v \notin \{v_1, \dots, v_5\}$. Without loss of generality, we may assume that $x, v \in \{v_1, \dots, v_5\}$ and $y, u \notin \{v_1, \dots, v_5\}$. As before, we may also assume that $v_2 = x$ and $v \in \{v_4, v_5\}$ and we still know that $v_1 \notin \{x, y, u, v\}$. From $u \notin \{v_1, \dots, v_5\}$, it follows that for $z \in \{v_4, v_5\} \setminus v$, we have $z \neq u$ and so, since v is a copy of y and $v_1vz \in E(H')$, we have that $v_1yz \in E(H)$. But then $y \notin \{v_1, \dots, v_5\}$ and $xy \in E(H)$ yield $B \subseteq H[v_1, x, v_3, y, z]$, a contradiction.

Thus, H' is a B -free simplicial complex on n vertices with at least as many edges as H . We repeat this operation iteratively so long as there are $xy, x'y' \in M$ such that $E_x \neq E_{x'}$, $E_y \neq E_{y'}$, or $E_{xy} \neq E_{x'y'}$. We may therefore indeed assume that in H the pairs of M are ‘symmetrised’, i.e., for all $xy, x'y' \in M$, we have $E_x = E_{x'}$, $E_y = E_{y'}$, and $E_{xy} = E_{x'y'}$.

¹Observe, however, that u is *not* a copy of x in the usual sense. For instance, there are no edges containing both u and y .

The vertices outside $V_1 \cup V(M)$. Observe that $H[V_1 \cup V(M)]$ already has the same structure as Example 1.3, except perhaps for the orders of V_1 and $V(M)$. In this subsection, we conclude the proof by showing that $V_3 := V \setminus (V_1 \cup V(M)) = \emptyset$. First note that, for every vertex $v \in V_3$, we have

$$V_1 \stackrel{2}{\leftarrow} v. \quad (2.4)$$

Indeed, since $v \notin V_1$, there is a tip of a star in L_v , so Claim 2.9 gives $V_1 \stackrel{i}{\leftarrow} v$ for some $i \geq 1$. Since $v \notin V(M)$, v is not contained in a 3-edge together with any vertex in V_1 and so $i \geq 2$. Lastly, Claim 2.6 yields $i = 2$.

Now we show that there are no 3-edges with all vertices in V_3 .

Claim 2.11. $E \cap V_3^{(3)} = \emptyset$.

Proof. Assume that $abc \in E$ with $a, b, c \in V_3$. Clearly, $ab \in L_c$ and, by Corollary 2.5 (3), in L_c the pair ab is either an isolated 2-edge or part of an isolated induced (2-uniform) star with at least two edges. A similar conclusion holds for the two other pairs in $\{a, b, c\}$. Suppose that two pairs of vertices, say ab and ac , appear as isolated 2-edges in L_c and L_b , respectively. Then, by Claim 2.3, this would imply that $\varrho(c) < \varrho(b)$ and $\varrho(b) < \varrho(c)$, respectively, which is a contradiction. Therefore, at least one of the pairs in $\{a, b, c\}$ appears as part of an isolated induced (2-uniform) star with at least two edges in the relevant link. But this means that at least one of the vertices a, b, c , say a , appears as a tip of a star in the link of one of the other two, say b . That is, $a \leftarrow b$. But then $V_1 \stackrel{2}{\leftarrow} a \leftarrow b$, so that $V_1 \stackrel{3}{\leftarrow} b$, contradicting Claim 2.6. \square

Next, we show that if V_3 is non-empty, then $V(M)$ must be very large.

Claim 2.12. *If $V_3 \neq \emptyset$, then $|V(M)| > n - \frac{21}{2}$.*

Proof. Let $v \in V_3$. If there is a 3-edge $uvw \in E$ with $u \in V_1$, then vw is an edge in L_u and since L_u only contains isolated 2-edges, this would give $v \in V(M)$, a contradiction. Together with Claim 2.11, this implies that every 3-edge containing v contains a vertex of $V(M)$. On the other hand, Lemma 2.8 states that for every $xy \in M$, L_x contains exactly one (2-uniform) star with at least two edges and that the centre of this star is y . The leaves of this star are all in V_1 by Claim 2.9 and Lemma 2.7. Since $v \notin V_1$, this implies that if v lies in a 3-edge with x , then it lies in exactly one such 3-edge (which appears as an isolated 2-edge in L_x). Note also that any 3-edge in H intersects at most one $xy \in M$, since H does not contain any 2-edge between two vertices in distinct $xy, x'y' \in M$.

Recall that vxy cannot be an edge (because then, in L_x , v would be a leaf of the star centred in y). Hence, assuming that v lies in a 3-edge with each of x and y , there

are $u, w \in V_3$ such that $uvx, wvy \in E$. If $u \neq w$, we directly obtain a copy of B . If $u = w$, we obtain a copy of B by using a vertex in V_1 (for instance, with x being the vertex that is contained in both 3-edges of this copy of B). Therefore, v cannot lie in a 3-edge with both x and y . Combining this with the above, we obtain that

$$d^{(3)}(v) \leq |V(M)|/2. \quad (2.5)$$

On the other hand, by (2.2), we have that $\varrho(a) > \frac{2(n-1)}{8} - \frac{11}{8}$ for all $a \in V_1$, so that $d^{(3)}(v) > \frac{4(n-1)}{8} - \frac{19}{4}$ by (2.4) and two applications of Claim 2.4. Together with (2.5), this yields the claim. \square

Now we can finish the proof. Since there is a vertex in V_1 by Corollary 2.5 (1), the minimum degree condition (2.1) implies that there is some $vw \in M$. Claim 2.3 and (2.2) together imply that $d^{(3)}(v) > \frac{3(n-1)}{8} - \frac{17}{8}$.

However, we can get an upper bound as follows. Recall that by Corollary 2.5 (3), the 2-edges in L_v are isolated 2-edges and isolated induced 2-uniform stars with at least two edges. Lemma 2.8 yields that there is exactly one star with at least two edges, its centre is w and, by Claim 2.9 and Lemma 2.7, its leaves are precisely the vertices in V_1 . Hence, the 2-edges in L_v are precisely given by a star with at least two edges and centre w whose leaves are precisely the vertices in V_1 as well as a matching of isolated 2-edges all of whose vertices are in V_3 . Therefore, we have

$$d^{(3)}(v) \leq |V_1| + \frac{|V_3|}{2} \leq n - |V(M)|. \quad (2.6)$$

If $V_3 \neq \emptyset$, (2.6) and Claim 2.12 yield $d^{(3)}(v) \leq \frac{3(n-1)}{8} - \frac{17}{8}$, which contradicts the lower bound above. Hence, $V_3 = \emptyset$, which means that H looks like Example 1.3 up to the sizes of the partition classes. A simple optimisation and the extremality of H then yields that H is the hypergraph in Example 1.3.

§3. PROOF OF THEOREM 1.4

For the lower bound, take an n -vertex simplicial complex containing all pairs as 2-edges and an intersecting family of 3-edges. The total number of edges is $1 + n + \binom{n}{2} + \binom{n-1}{2} = n^2 - n + 2$.

For the upper bound, suppose C is sufficiently large. We will show by induction on n that every simplicial complex on n vertices with at least $n^2 - n + C$ edges contains a copy of F .

The base case of the induction follows simply from C being large. Now assume that for every $m \leq n - 1$ every simplicial complex on m vertices with at least $m^2 - m + C$ edges contains a copy of F and let H be a simplicial complex on n vertices with at least $n^2 - n + C$

edges. Since C is large, we can assume that n is large as well. First note that if there is a vertex v with $d(v) \leq 2n - 3$, then

$$e(H - v) \geq n^2 - n + C - (2n - 3) - 1 = (n - 1)^2 - (n - 1) + C,$$

so that, by induction, $H - v$ and, thus, H contain a copy of F . We may therefore assume that $d(v) \geq 2n - 2$ for every $v \in V(H)$.

Next we want to argue that if H contains two intersecting 4-edges, then it contains a copy of F . Let us therefore assume that H contains an edge $e = wxyz$. Then it is easy to see that if any 3-edge intersects e in exactly one vertex, then H contains a copy of F . In particular, this implies that if any other 4-edge intersects e , it does so in exactly three vertices. Assume that there is some such $e' = w'xyz$. Then w' (and w) cannot be contained in any 3-edge that contains a vertex outside of $e \cup e'$. But this yields that $d(w') \leq 1 + n - 1 + 6 \leq n + 6$, a contradiction. Thus, no two 4-edges intersect, which also implies that there are no edges of size larger than 4 and that every vertex is contained in at most one 4-edge.

We now show that $d^{(3)}(v) \leq 3n$ for all $v \in V(H)$. Let $v \in V(H)$ be arbitrary and let $a \in V(H)$ be such that $va \in E(H)$. We first show that there is a 3-edge containing a but not v . This follows directly if a is contained in a 4-uniform edge, so we may assume that $d^{(4)}(v) = 0$. Therefore, we have

$$d^{(3)}(a) = d^{(\geq 2)}(a) - d^{(2)}(a) \geq 2n - 2 - (n - 1) \geq n - 1,$$

so there must be a 3-edge e containing a but not v . But then every 3-edge that contains v has to intersect e , so that $d^{(3)}(v) \leq 3n$.

We now fix $v \in V(H)$ and let $\alpha = \frac{|N(v)|}{n} < 1$. Consider some 3-edge uvw and note that every 3-edge of H which intersects $N(v)$ has to intersect uvw . Thus, the above bound on the 3-degree applied to u , v , and w in turn yields that there are at most $9n$ such edges. Therefore,

$$\begin{aligned} e(H - N(v)) &\geq n^2 - n + C - \left(2\alpha n + \binom{\alpha n}{2} + \alpha(1 - \alpha)n^2 + 9n\right) \\ &\geq \left(1 - \alpha(1 - \alpha) - \frac{\alpha^2}{2}\right)n^2 - (1 + 2\alpha + 9)n + C \geq ((1 - \alpha)n)^2 - (1 - \alpha)n + C. \end{aligned}$$

The result therefore follows by induction.

§4. PROOF OF THEOREM 1.5

We will actually prove the following more general theorem, of which Theorem 1.5 is a special case.

Theorem 4.1. *For every proper 3-graph F , $A \subseteq V(F)$ and $\varepsilon > 0$, there is $n_0 \in \mathbb{N}$ such that the following holds. If a simplicial complex H on $n \geq n_0$ vertices satisfies*

$$\frac{e^3 n^{n-2}}{(n+3)^n} [\text{ex}^{(3)}(n, F) + n^2] \geq \left[\frac{(|A| - 1)}{|A|} + \frac{1}{n} \frac{v(F) - |A|}{|A|} \right] \left[1 + \sum_{i \geq 2} \frac{\Delta^{(i+1)}}{i} \right], \quad (4.1)$$

where $\Delta^{(i)} = \max_{xy \in V(H)^{(2)}} d^{(i)}(xy)$, and $e(H) \geq (1 + \varepsilon) [\text{ex}^{(3)}(n, F) + \binom{n}{2}]$, then H contains a copy of F^+ if either:

- (1) H does not contain any edge of size 4 or
- (2) F is contained in a 4-edge.

To see that Theorem 4.1 indeed implies Theorem 1.5, observe that $(\frac{n}{n+3})^n$ is a decreasing function and tends to e^{-3} as n tends to infinity. Therefore, we have that (1.1) yields (4.1).

Proof of Theorem 4.1. Let F be a proper 3-graph, $A \subseteq V(F)$ and $\varepsilon > 0$. Let C be a sufficiently large constant ($C^{-1} \ll \varepsilon, 1/v(F)$). We will show by induction on n that if H is a simplicial complex on $n \geq v(F^+)$ vertices satisfying (4.1) and $e(H) \geq (1 + \varepsilon/2) [\text{ex}^{(3)}(n, F) + \binom{n}{2}] + C$, then H contains a copy of F^+ . It then follows that there is some n_0 such that every simplicial complex H on $n \geq n_0$ vertices satisfying (4.1) and $e(H) \geq (1 + \varepsilon) [\text{ex}^{(3)}(n, F) + \binom{n}{2}]$ contains a copy of F^+ .

The base of the induction and the statement for small n both follow from C being large. Now let n be large and assume that for every $n' \leq n - 1$, every simplicial complex H' on n' vertices satisfying (4.1) (with n' in place of n) and $e(H') \geq (1 + \varepsilon/2) [\text{ex}^{(3)}(n', F) + \binom{n'}{2}] + C$ contains a copy of F^+ . Let H be a simplicial complex on $n \geq n_0$ vertices satisfying (4.1) with $e(H) \geq (1 + \varepsilon/2) [\text{ex}^{(3)}(n, F) + \binom{n}{2}] + C$. Assume for the sake of contradiction that H does not contain a copy of F^+ . We first derive a minimum degree condition. Suppose that there is a vertex v in H with $d(v) \leq \frac{e^3 n^n}{(n+3)^n} \left[\frac{\text{ex}^{(3)}(n, F)}{n} + n \right]$. Note that for some fixed $\eta > 0$ and large n , we have $\frac{(n+3)^n}{e^3 n^n} (1 + \varepsilon/2) > 1 + \eta$. By the definition of proper (applied with η instead of ε) and since n is large, we therefore have

$$\frac{(1 + \varepsilon/2)(n+3)^n}{e^3 n^n} [\text{ex}^{(3)}(n, F) - \text{ex}^{(3)}(n-1, F)] \geq \frac{\text{ex}^{(3)}(n, F)}{n}$$

or, after rearranging,

$$(1 + \varepsilon/2) [\text{ex}^{(3)}(n, F) - \text{ex}^{(3)}(n-1, F)] \geq \frac{e^3 n^n}{(n+3)^n} \frac{\text{ex}^{(3)}(n, F)}{n}.$$

Keeping in mind that n is large and so $\frac{e^3 n^n}{(n+3)^n} < 1 + \frac{\varepsilon}{4}$, this in turn implies that

$$e(H-v) \geq \left(1 + \frac{\varepsilon}{2}\right) [\text{ex}^{(3)}(n, F) + \binom{n}{2}] + C - \frac{e^3 n^n}{(n+3)^n} \left[\frac{\text{ex}^{(3)}(n, F)}{n} + n \right] - 1$$

$$\geq \left(1 + \frac{\varepsilon}{2}\right) \left[\text{ex}^{(3)}(n-1, F) + \binom{n-1}{2} \right] + C.$$

In order to apply induction, we must also check that (4.1) holds for $H-v$ and $n-1$ instead of H and n . But, by a double counting argument from [16], we have $\frac{\text{ex}^{(3)}(n-1, F)}{n-3} \geq \frac{\text{ex}^{(3)}(n, F)}{n}$, so, since n is large, we have

$$\begin{aligned} & \frac{e^3(n-1)^{n-3}}{(n+2)^{n-1}} \left[\text{ex}^{(3)}(n-1, F) + (n-1)^2 \right] \\ &= \frac{e^3(n-1)^{n-3}(n-3)}{(n+2)^{n-1}} \frac{\text{ex}^{(3)}(n-1, F)}{n-3} + \frac{e^3(n-1)^{n-1}(n+3)^n}{(n+2)^{n-1}e^3n^n} \frac{e^3n^n}{(n+3)^n} \\ &\geq \frac{e^3(n-1)^{n-3}(n-3)}{(n+2)^{n-1}} \frac{\text{ex}^{(3)}(n, F)}{n} + \frac{n}{n-1} \frac{e^3n^n}{(n+3)^n} \\ &\geq \frac{n}{n-1} \frac{e^3n^{n-2}}{(n+3)^n} \left[\text{ex}^{(3)}(n, F) + n^2 \right] \\ &\geq \frac{n}{n-1} \left[\frac{(|A|-1)}{|A|} + \frac{1}{n} \frac{v(F) - |A|}{|A|} \right] \left[1 + \sum_{i \geq 2} \frac{\Delta_H^{(i+1)}}{i} \right] \\ &\geq \left[\frac{(|A|-1)}{|A|} + \frac{1}{n-1} \frac{v(F) - |A|}{|A|} \right] \left[1 + \sum_{i \geq 2} \frac{\Delta_{H-v}^{(i+1)}}{i} \right], \end{aligned}$$

where the second inequality follows from $\frac{e^3(n-1)^{n-3}(n-3)}{(n+2)^{n-1}} \geq \frac{n}{n-1} \frac{e^3n^{n-2}}{(n+3)^n}$ for large n . Hence, we can apply induction to $H-v$, yielding $F^+ \subseteq H-v \subseteq H$, a contradiction. Thus, we may assume that $d(v) > \frac{e^3n^n}{(n+3)^n} \left[\frac{\text{ex}^{(3)}(n, F)}{n} + n \right]$ for every $v \in V(H)$.

If either of the conditions (1) or (2) holds, we know that H contains a copy of F . Abusing notation slightly, we now use F and A to denote the images of F and A under some embedding. For H not to contain a copy of F^+ , there must be a missing 2-edge between every vertex in $V(H-F)$ and some vertex in A . Therefore, there is some vertex $a \in A$ with at least $\frac{n-v(F)}{|A|}$ missing 2-edges. Note also that since $d^{(i)}(xy) \leq \Delta^{(i)}$ for all $xy \in V(H)^{(2)}$ and H is hereditary, we have that, for every $v \in V$,

$$d(v) \leq d^{(2)}(v) \left[1 + \sum_{i \geq 2} \frac{\Delta^{(i+1)}}{i} \right].$$

Combining this with our lower bound on the degree (of any vertex) and our upper bound on $d^{(2)}(a)$, we get that

$$\begin{aligned} \frac{e^3n^n}{(n+3)^n} \left[\frac{\text{ex}^{(3)}(n, F)}{n} + n \right] &< d(a) \leq d^{(2)}(a) \left[1 + \sum_{i \geq 2} \frac{\Delta^{(i+1)}}{i} \right] \\ &\leq \left(n-1 - \frac{n-v(F)}{|A|} \right) \left[1 + \sum_{i \geq 2} \frac{\Delta^{(i+1)}}{i} \right] \end{aligned}$$

$$\leq \frac{(|A| - 1)n + v(F) - |A|}{|A|} \left[1 + \sum_{i \geq 2} \frac{\Delta^{(i+1)}}{i} \right]$$

and, hence,

$$\frac{e^3 n^n}{(n+3)^n} \left[\frac{\text{ex}^{(3)}(n, F)}{n^2} + 1 \right] < \left[\frac{(|A| - 1)}{|A|} + \frac{1}{n} \frac{v(F) - |A|}{|A|} \right] \left[1 + \sum_{i \geq 2} \frac{\Delta^{(i+1)}}{i} \right].$$

But this contradicts (4.1), meaning that H does contain a copy of F^+ . \square

§5. PROOF OF THEOREM 1.7

Proof of Theorem 1.7. For the lower bound, simply consider the simplicial complex on n vertices that has three (almost) balanced partition classes and all 2-edges and 3-edges which intersect each partition class in at most one vertex.

For the upper bound, first note that if a simplicial complex contains an edge of size at least four, then it contains a copy of F . Thus, we only need to show that given $\varepsilon > 0$, there is some n_0 such that every simplicial complex H on $n \geq n_0$ vertices with at least $(\frac{2}{9} + \varepsilon) \binom{n}{3}$ edges, all of which are of size at most three, contains a copy of F . Next, note that F is contained in the simplicial complex

$$F^+ = (\{v_1, v_2, v_3, v_4, v_5\}, \{v_1v_2v_3, v_2v_3v_4, v_1v_4v_5\}).$$

Bollobás [4] showed that for every $\varepsilon > 0$, there is some n_1 such that every 3-graph on $n \geq n_1$ vertices with at least $(\frac{2}{9} + \frac{\varepsilon}{2}) \binom{n}{3}$ edges contains the 3-graph

$$F_3 = (\{v_1, v_2, v_3, v_4, v_5\}, \{v_1v_2v_3, v_2v_3v_4, v_1v_4v_5\}).$$

From this, we can infer our result as follows. For every $\varepsilon > 0$, there is some n_2 such that every simplicial complex H on $n \geq n_2$ vertices with at least $(\frac{2}{9} + \varepsilon) \binom{n}{3}$ edges not containing any edge of size larger than three has at least $(\frac{2}{9} + \frac{\varepsilon}{2}) \binom{n}{3}$ 3-edges. Thus, the 3-uniform layer of any simplicial complex H on $n \geq n_0 = \max\{n_1, n_2\}$ vertices with at least $(\frac{2}{9} + \varepsilon) \binom{n}{3}$ edges contains a copy of F_3 , which in the simplicial complex H induces a copy of F^+ and, thereby, a copy of F . \square

§6. CONCLUDING REMARKS

In this work, we studied the Turán problem for simplicial complexes. Since this problem seems to be of comparable difficulty to the Turán problem for uniform hypergraphs, any non-trivial result on the following problem would be of interest.

Problem 6.1. Given a simplicial complex F and a positive integer n , determine $\text{ex}(n, F)$ or its asymptotics.

Theorem 1.2 shows that this problem is distinct from the problem for uniform hypergraphs in that, for some F , the extremal construction for F has more than one incomplete layer. On the other hand, we showed that for several simplicial complexes F , the extremal number is asymptotically given by the trivial bound $\text{ex}^{(k)}(n, F_k) + \sum_{i=0}^{k-1} \binom{n}{i}$, where $k = \dim(F) + 1$. This suggests the following problem.

Problem 6.2. Given an integer $k \geq 3$, characterise the $(k - 1)$ -dimensional simplicial complexes F with $\text{ex}(n, F) = \text{ex}^{(k)}(n, F_k) + \sum_{i=0}^{k-1} \binom{n}{i}$.

As mentioned in the introduction, Hajnal (see [6, 15]) initiated the investigation of a variant of the Brown–Erdős–Sós problem [22] for simplicial complexes. We believe this special case of the Turán problem for simplicial complexes to be of particular interest. Recall that $\mathcal{F}(k, d)$ denotes the family of all simplicial complexes with vertex set $[k]$ and $1 + \sum_{i=0}^d \binom{k}{i}$ edges.

Problem 6.3. For $k \geq 5$ and $d \geq 2$, determine $\text{ex}(n, \mathcal{F}(k, d))$ or its asymptotics.

We have already noted that Bollobás and Radcliffe [5] solved Problem 6.3 for $k = 4$ and $d = 2$. A simple inductive argument gives the following upper bound for the case $k = 5$ and $d = 2$.

Proposition 6.4. For every $n \geq 5$, $\text{ex}(n, \mathcal{F}(5, 2)) \leq \frac{1}{2}(n^{5/2} + 3n^2)$.

Proof. We proceed by induction, noting that the base case $n = 5$ follows easily since $\frac{1}{2}(n^{5/2} + 3n^2) \geq 55 > 32 = \sum_{i=0}^5 \binom{5}{i}$. Suppose now that H is a simplicial complex on n vertices with $e(H) > \frac{1}{2}(n^{5/2} + 3n^2)$.

If there is a vertex v such that $d(v) \leq \frac{1}{2}(n^{3/2} + 3n - 2)$, then

$$e(H - v) > \frac{1}{2}(n^{5/2} + 3n^2 - n^{3/2} - 3n + 2) - 1 \geq \frac{1}{2}((n - 1)^{5/2} + 3(n - 1)^2),$$

so, by induction, $H - v$ contains some $F \in \mathcal{F}(5, 2)$. Therefore, we may assume that

$$d(v) > \frac{1}{2}(n^{3/2} + 3n - 2) \tag{6.1}$$

for every vertex $v \in V(H)$.

Suppose that there is an edge $e \in E^{(4)}(H)$. Then it is easy to see that for any vertex $v \in V(H) \setminus e$ the set $e \cup \{v\}$ spans at least 17 edges. Hence, we may assume that $E^{(4)} = \emptyset$ and so (6.1) yields that $d^{(3)}(v) \geq \frac{1}{2}(n^{3/2} + n)$ for every vertex $v \in V(H)$. Therefore, for any vertex $v \in V(H)$, the classical theorem of Kővari, Sós and Turán [19] implies that there is a (2-uniform) copy of C_4 in the link of v . It is not hard to check that the vertices of such a C_4 together with v span at least 18 edges. \square

Question 6.5. Is it true that $\text{ex}(n, \mathcal{F}(5, 2)) = \Omega(n^{5/2})$?

Let $LC_4^{(3)}$ be the 3-uniform hypergraph with $V(LC_4^{(3)}) = \{u, v_1, \dots, v_4\}$ and $E(LC_4^{(3)}) = \{uv_i v_{i+1} : i \in [4]\}$, where the indices are viewed as in $\mathbb{Z}/4\mathbb{Z}$. In other words, $LC_4^{(3)}$ is the 3-uniform hypergraph consisting of a vertex u with a (2-uniform) C_4 in its link. By following the proof of Proposition 6.4, we get that $\text{ex}(n, \mathcal{F}(5, 2)) \leq \text{ex}^{(3)}(n, LC_4^{(3)}) + \binom{n}{2} + n + 1$. In particular, improving the bound $\text{ex}^{(3)}(n, LC_4^{(3)}) = O(n^{5/2})$ would answer Question 6.5 in the negative. However, a result of Brown, Erdős and Sós [22, Theorem 3] shows that $\text{ex}^{(3)}(n, LC_4^{(3)}) = \Omega(n^{5/2})$, perhaps giving some evidence towards a positive answer for Question 6.5.

Let us conclude by briefly discussing the problem of determining the extremal number of the simplicial complex induced by the k -uniform clique on t vertices for $t \geq k + 2$. Denote the simplicial complex given by the downward closure of $K_t^{(k)}$ by $K(t, k)$.

Problem 6.6. Given $t \geq k + 2$, determine the asymptotics of $\text{ex}(n, K(t, k))$.

A simple symmetrisation argument yields the following result for the case $k = 2$.

Observation 6.7. For $t \geq 3$, $\varepsilon > 0$ and n sufficiently large, $\text{ex}(n, K(t + 1, 2)) = (1 \pm \varepsilon) \frac{t!}{t^t} \binom{n}{t}$.

Proof. Let H be a simplicial complex on n vertices without a $K(t + 1, 2)$ and, subject to this condition, with the maximum number of edges. Assume that there are three vertices x, y and z such that xy and yz are not in $E(H)$ but xz is. If $d(y) \geq d(x), d(z)$, then we can replace x and z by copies of y . This would yield a simplicial complex without a $K(t + 1, 2)$ but with more edges than H , a contradiction. Otherwise, for one of x and z , say, without loss of generality, for x , we have $d(x) > d(y)$. In this case we can replace y by a copy of x , again yielding a simplicial complex without a $K(t + 1, 2)$ but with more edges than H . Thus, we may assume that non-adjacency is an equivalence relation, so that H is complete r -partite for some $r \leq t$, that is, the vertex set can be partitioned into r sets such that every edge contains at most one vertex from each of these sets. Therefore, H has $\max_{r \in [t]} (1 \pm \varepsilon) \frac{r!}{r^r} \binom{n}{r} = (1 \pm \varepsilon) \frac{t!}{t^t} \binom{n}{t}$ edges. \square

This problem is related to another much-studied problem in extremal combinatorics. Following Alon and Shikhelman [1], given two k -graphs F and G , we write $\text{ex}(n, F, G)$ for the maximum number of copies of F in a G -free k -graph on n vertices.

Observation 6.8. For $t \geq k + 2$, $\varepsilon > 0$ and n sufficiently large,

$$\text{ex}(n, K(t, k)) = (1 \pm \varepsilon) \text{ex}(n, K_{t-1}^{(k)}, K_t^{(k)}).$$

Proof. Consider a simplicial complex H on n vertices with at least $(1 + \varepsilon) \text{ex}(n, K_{t-1}^{(k)}, K_t^{(k)})$ edges. Any t -edge in H would immediately yield a $K(t, k)$, so we can assume that H does

not contain any edges of size larger than $t - 1$. Thus, since n is taken sufficiently large, H contains at least $(1 + \varepsilon/2)\text{ex}(n, K_{t-1}^{(k)}, K_t^{(k)})$ edges of size $t - 1$. Each of these induces a $K_{t-1}^{(k)}$. But, by definition, this implies that H contains a $K_t^{(k)}$ and, thereby, a $K(t, k)$.

Suppose now that H'' is a k -uniform hypergraph on n vertices that maximises the number of copies of $K_{t-1}^{(k)}$ without containing a $K_t^{(k)}$. Let H' be the $(t - 1)$ -uniform hypergraph on the same vertex set that is obtained by placing a $(t - 1)$ -edge for every $K_{t-1}^{(k)}$ of H'' and let H be the downward closure of H' . Clearly, since n is large, H has more than $\text{ex}(n, K_{t-1}^{(k)}, K_t^{(k)})$ edges. It is also easy to check that H does not contain a copy of $K(t, k)$. \square

Combining Observation 6.8 with [1, Proposition 2.2], which implies that $\text{ex}(n, K_{t-1}, K_t) = (1 + o(1))\binom{n}{t}^t$, gives another proof of Observation 6.7. Moreover, using flag algebras, Bodnar [3] recently proved that $\text{ex}(n, K_4^{(3)}, K_5^{(3)}) = (\frac{3}{8} + o(1))\binom{n}{4}$. By Observation 6.8, this implies that $\text{ex}(n, K(5, 3)) = (\frac{3}{8} + o(1))\binom{n}{4}$. To the best of our knowledge, Problem 6.6 remains open for all other (t, k) with $t \geq 6$ or $k \geq 3$.

REFERENCES

- [1] N. Alon and C. Shikhelman, *Many T copies in H -free graphs*, J. Combin. Theory Ser. B **121** (2016), 146–172, DOI [10.1016/j.jctb.2016.03.004](https://doi.org/10.1016/j.jctb.2016.03.004). MR3548290 [↑6, 6](#)
- [2] L. Bellmann and Chr. Reiher, *Turán's theorem for the Fano plane*, Combinatorica **39** (2019), no. 5, 961–982, DOI [10.1007/s00493-019-3981-8](https://doi.org/10.1007/s00493-019-3981-8). MR4039597 [↑1](#)
- [3] L. Bodnar, *Generalized Turán problem for complete hypergraphs*, preprint available at arXiv:2302.07571 [math.CO]. [↑6](#)
- [4] B. Bollobás, *Three-graphs without two triples whose symmetric difference is contained in a third*, Discrete Math. **8** (1974), 21–24, DOI [10.1016/0012-365X\(74\)90105-8](https://doi.org/10.1016/0012-365X(74)90105-8). MR345869 [↑5](#)
- [5] B. Bollobás and A. J. Radcliffe, *Defect Sauer results*, J. Combin. Theory Ser. A **72** (1995), no. 2, 189–208, DOI [10.1016/0097-3165\(95\)90060-8](https://doi.org/10.1016/0097-3165(95)90060-8). MR1357769 [↑1, 6](#)
- [6] J. A. Bondy, *Induced subsets*, J. Combin. Theory Ser. B **12** (1972), no. 2, 201–202, DOI [10.1016/0095-8956\(72\)90025-1](https://doi.org/10.1016/0095-8956(72)90025-1). MR319773 [↑1, 6](#)
- [7] B. Bukh and X. Goaoc, *Shatter functions with polynomial growth rates*, SIAM J. Discrete Math. **33** (2019), no. 2, 784–794, DOI [10.1137/17M1113680](https://doi.org/10.1137/17M1113680). MR3948236 [↑1](#)
- [8] D. De Caen and Z. Füredi, *The maximum size of 3-uniform hypergraphs not containing a Fano plane*, J. Combin. Theory Ser. B **78** (2000), no. 2, 274–276, DOI [10.1006/jctb.1999.1938](https://doi.org/10.1006/jctb.1999.1938). MR1750899 [↑1](#)
- [9] P. Erdős and M. Simonovits, *A limit theorem in graph theory*, Studia Sci. Math. Hungar. **1** (1966), 51–57. MR205876 [↑1](#)
- [10] P. Erdős and A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc. **52** (1946), 1087–1091, DOI [10.1090/S0002-9904-1946-08715-7](https://doi.org/10.1090/S0002-9904-1946-08715-7). MR0018807 [↑1](#)
- [11] P. Frankl, *On a problem of Bondy and Hajnal*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 159 (English, with French summary). MR539967 [↑1](#)

- [12] P. Frankl, *On the trace of finite sets*, J. Combin. Theory Ser. A **34** (1983), no. 1, 41–45, DOI [10.1016/0097-3165\(83\)90038-9](https://doi.org/10.1016/0097-3165(83)90038-9). MR685210 [↑1](#)
- [13] P. Frankl and Z. Füredi, *Exact solution of some Turán-type problems*, J. Combin. Theory Ser. A **45** (1987), no. 2, 226–262, DOI [10.1016/0097-3165\(87\)90016-1](https://doi.org/10.1016/0097-3165(87)90016-1). MR0894820 [↑1](#)
- [14] Z. Füredi and J. Pach, *Traces of finite sets: extremal problems and geometric applications*, Extremal problems for finite sets (Visegrád, 1991), Bolyai Soc. Math. Stud., vol. 3, János Bolyai Math. Soc., Budapest, 1994, pp. 251–282. MR1319167 [↑1](#)
- [15] Z. Füredi and M. Simonovits, *Triple systems not containing a Fano configuration*, Combin. Probab. Comput. **14** (2005), no. 4, 467–484, DOI [10.1017/S0963548305006784](https://doi.org/10.1017/S0963548305006784). MR2160414 [↑1, 6](#)
- [16] G. Katona, T. Nemetz, and M. Simonovits, *On a problem of Turán in the theory of graphs*, Mat. Lapok **15** (1964), 228–238 (Hungarian, with English and Russian summaries). MR0172263 [↑1, 1, 4](#)
- [17] P. Keevash and B. Sudakov, *The Turán number of the Fano plane*, Combinatorica **25** (2005), no. 5, 561–574, DOI [10.1007/s00493-005-0034-2](https://doi.org/10.1007/s00493-005-0034-2). MR2176425 [↑1](#)
- [18] P. Keevash, *Hypergraph Turán problems*, Surveys in combinatorics 2011, London Math. Soc. Lecture Note Ser., vol. 392, Cambridge Univ. Press, Cambridge, 2011, pp. 83–139. MR2866732 [↑1](#)
- [19] T. Kővári, V. T. Sós, and P. Turán, *On a problem of K. Zarankiewicz*, Colloq. Math. **3** (1954), 50–57. [↑6](#)
- [20] N. Sauer, *On the density of families of sets*, J. Combin. Theory Ser. A **13** (1972), 145–147, DOI [10.1016/0097-3165\(72\)90019-2](https://doi.org/10.1016/0097-3165(72)90019-2). MR307902 [↑1](#)
- [21] S. Shelah, *A combinatorial problem; stability and order for models and theories in infinitary languages*, Pacific J. Math. **41** (1972), 247–261. MR307903 [↑1](#)
- [22] V. T. Sós, P. Erdős, and W. G. Brown, *On the existence of triangulated spheres in 3-graphs, and related problems*, Period. Math. Hungar. **3** (1973), no. 3-4, 221–228, DOI [10.1007/BF02018585](https://doi.org/10.1007/BF02018585). MR0323647 [↑1, 6, 6](#)
- [23] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452 (Hungarian, with German summary). MR18405 [↑1](#)
- [24] V. N. Vapnik and A. Ja. Červonenkis, *The uniform convergence of frequencies of the appearance of events to their probabilities*, Dokl. Akad. Nauk SSSR **181** (1968), 781–783 (Russian). MR0231431 [↑1](#)

DEPARTMENT OF MATHEMATICS, CALIFORNIA INSTITUTE OF TECHNOLOGY, USA

Email address: dconlon@caltech.edu

SCHOOL OF MATHEMATICS, UNIVERSITY OF BIRMINGHAM, UK

Email address: s.piga@bham.ac.uk

DEPARTMENT OF MATHEMATICS, CALIFORNIA INSTITUTE OF TECHNOLOGY, USA

Email address: schuelke@caltech.edu