Original citation:
Falgas--Ravry, Victor, Marchant, Edward, Pikhurko, Oleg and Vaughan, Emil R.. (2015) The codegree threshold for 3-graphs with independent neighborhoods. SIAM Journal on Discrete Mathematics, 29 (3). pp. 1504-1539.

Permanent WRAP url:
http://wrap.warwick.ac.uk/72484

SIAM J. DISCRETE MATH.
Vol. 29, No. 3, pp. 1504–1539

THE CODEGREE THRESHOLD FOR 3-GRAPHS WITH INDEPENDENT NEIGHBORHOODS*

VICTOR FALGAS–RAVRY[†], EDWARD MARCHANT[‡], OLEG PIKHURKO[§], AND EMIL R. VAUGHAN[¶]

Abstract. Given a family of 3-graphs $\mathcal{F}$, we define its codegree threshold $\operatorname{coex}(n, \mathcal{F})$ to be the largest number $d = d(n)$ such that there exists an n-vertex 3-graph in which every pair of vertices is contained in at least d 3-edges but which contains no member of $\mathcal{F}$ as a subgraph. Let $F_{3,2}$ be the 3-graph on $\{a, b, c, d, e\}$ with 3-edges abc, abd, abe, and cde. In this paper, we give two proofs that $\operatorname{coex}(n, \{F_{3,2}\}) = \left(\frac{1}{3} + o(1)\right)n$, the first by a direct combinatorial argument and the second via a flag algebra computation. Information extracted from the latter proof is then used to obtain a stability result, from which in turn we derive the exact codegree threshold for all sufficiently large n: $\operatorname{coex}(n, \{F_{3,2}\}) = \lfloor n/3 \rfloor - 1$ if n is congruent to 1 modulo 3, and $\lfloor n/3 \rfloor$ otherwise. In addition we determine the set of codegree-extremal configurations for all sufficiently large n.

Key words. codegree, Turán density, Turán function, 3-graphs

AMS subject classifications. 05D05, 05C35, 05C65

DOI. 10.1137/130926997

1. Introduction.

1.1. Turán-type problems. We begin with some standard definitions. Let $r, n \in \mathbb{N}$. We write $[n]$ for the discrete interval $\{1, 2, \dots, n\}$. Also, given a set S, we denote by $S^{(r)}$ the collection of all r-subsets from S.

An *r-graph* is a pair of sets $G = (V, E)$, where $V = V(G)$ is a set of *vertices* and $E = E(G)$ is a collection of r-sets from V, which constitute the *r-edges* of G. An r-graph G is *nonempty* if $E(G) \neq \emptyset$. A *subgraph* of G is an r-graph H with $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. Given a family of r-graphs $\mathcal{F}$, we say that G is $\mathcal{F}$*-free* if no member of $\mathcal{F}$ is isomorphic to a subgraph of G.

One of the central problems in extremal combinatorics is determining the maximum number $\operatorname{ex}(n, \mathcal{F})$ of r-edges that an r-graph on n vertices may contain while remaining $\mathcal{F}$-free, where $\mathcal{F}$ is a family of nonempty r-graphs. The function $n \mapsto \operatorname{ex}(n, \mathcal{F})$ is known as the *Turán number* of $\mathcal{F}$.

PROBLEM 1. *Let $\mathcal{F}$ be a family of nonempty r-graphs. Determine the Turán number of $\mathcal{F}$.*

Often, computing the Turán number exactly may be difficult, and so, lowering our sights, we are interested in the asymptotic behavior of the Turán function: what is the asymptotically maximal proportion of all possible edges that an $\mathcal{F}$-free

*Received by the editors July 1, 2013; accepted for publication (in revised form) May 28, 2015; published electronically August 18, 2015.
http://www.siam.org/journals/sidma/29-3/92699.html

[†]Department of Mathematics, Vanderbilt University, Nashville, TN 37240, and Institutionen för matematik och matematisk statistik, Umeå Universitet, 901 87 Umeå, Sweden (victor.falgas-ravry@vanderbilt.edu). This author's research was supported by the Kempe Foundation.

[‡]29 Woodside Close, HP6 5EF Amersham, UK (ejmarchant@gmail.com). This author's research was funded by Trinity College, Cambridge, UK.

[§]Mathematics Institute and DIMAP, University of Warwick, CV4 7AL Coventry, UK (O.Pikhurko@warwick.ac.uk). This author's research was supported by ERC grant 306493 and EPSRC grant EP/K012045/1.

[¶]Centre for Discrete Mathematics, Queen Mary University of London, E1 4NS London, UK (emil79@gmail.com).

r-graph may contain? An easy averaging argument shows that the nonnegative sequence $\mathrm{ex}(n, \mathcal{F})/\binom{n}{r}$ is nonincreasing and hence converges to a limit as n tends to infinity. This limit is known as the *Turán density* of $\mathcal{F}$ and is denoted by $\pi(\mathcal{F})$.

PROBLEM 2. *Let $\mathcal{F}$ be a family of nonempty r-graphs. Determine the Turán density of $\mathcal{F}$.*

These two problems have been studied very successfully in the case $r = 2$, corresponding to ordinary (2-)graphs. Turán determined the Turán number of complete graphs [37], while Erdős and Stone [9] fully resolved Problem 2 in a seminal result relating the Turán density of a family of graphs to its chromatic number.

Despite recent progress, this stands in some contrast to the situation when $r \geq 3$. Indeed few Turán densities are known even for 3-graphs, and the problem of determining them is known to be hard in general. Let us introduce here a few of the 3-graphs relevant to our discussion. As a convention, we will write xyz for the 3-edge $\{x, y, z\}$ and $\pi(F_1, F_2, \dots, F_t)$ for the Turán density $\pi(\{F_1, F_2, \dots, F_t\})$.

Let K_4 denote the complete 3-graph on four vertices, and let K_4^- denote the 3-graph obtained from K_4 by deleting one of its edges. Let $F_{3,2}$ be the 3-graph $([5], \{123, 124, 125, 345\})$. Finally, let F_7 be the *Fano plane*, namely the (unique up to isomorphism) 3-graph on seven vertices in which every pair of vertices is contained in exactly one 3-edge.

Almost no Turán densities or Turán numbers for 3-graphs were known until de Caen and Füredi [6] established that $\pi(F_7) = 3/4$. (A notable exception is a result of Bollobás [4].) The Turán number of the Fano plane was independently determined shortly afterwards by Keevash and Sudakov [23] and Füredi and Simonovits [16]. Around the same time, Füredi, Pikhurko, and Simonovits determined first the Turán density [14] and then the Turán number [15] of $F_{3,2}$.

The next major development as far as computing Turán densities is concerned was the advent of Razborov's semidefinite method [35]. With the assistance of computers, this method has been used in recent years to significantly increase the number of known Turán densities for 3-graphs [2, 13].

1.2. The codegree problem. Given a 3-graph G and a vertex $x \in V(G)$, the *degree* $d(x)$ of x in G is the number of 3-edges of G containing x. The *minimum degree* of G is $\delta(G) = \min_{x \in V(G)} d(x)$. It is not hard to see that the Turán density problem for 3-graphs is equivalent to determining asymptotically what minimum degree condition forces a 3-graph on n vertices to contain a copy of a member of a given family $\mathcal{F}$ as a subgraph.

A natural variant is to consider what minimum *codegree* condition is required to force an $\mathcal{F}$-subgraph. Here, the *codegree* $d(x, y)$ of two distinct vertices x, y in a 3-graph G is the number of 3-edges of G which contain the pair $\{x, y\}$. (We may sometimes write this as $d_G(x, y)$ to emphasize that we are taking the codegree in G and not some other 3-graph.) The *minimum codegree* $\delta_2(G)$ of G is, as the name suggests, the minimum of $d(x, y)$ over all pairs of vertices from $V(G)$.

We may then define for a family of nonempty 3-graphs $\mathcal{F}$ the *codegree threshold* $\mathrm{coex}(n, \mathcal{F})$ to be the maximum of $\delta_2(G)$ over all $\mathcal{F}$-free 3-graphs G on n vertices. This is the codegree analogue of the Turán number.

PROBLEM 3. *Let $\mathcal{F}$ be a family of nonempty 3-graphs. Determine the codegree threshold of $\mathcal{F}$.*

Again it may be that, in general, computing the codegree threshold proves difficult and that we would first be interested in determining the asymptotic behavior of $\mathrm{coex}(n, \mathcal{F})$. Following the analogy with the Turán-type problems, it is natural to consider the sequence $\mathrm{coex}(n, \mathcal{F})/(n-2)$ or some close relative. Here, however, we do not in general have monotonicity: Lo and Markström [25] showed that neither of $\mathrm{coex}(n, K_4)/n$ and $\mathrm{coex}(n, K_4)/(n-2)$ is nonincreasing. The limit of $\mathrm{coex}(n, \mathcal{F})/n$ does exist, however, as first shown by Mubayi and Zhao [31]. Thus we may define the *codegree density* of $\mathcal{F}$ to be

$$\gamma(\mathcal{F}) := \lim_{n\to\infty} \frac{\mathrm{coex}(n, \mathcal{F})}{n-2}.$$

(Obviously, choosing n or $n-2$ in the denominator does not affect the limit.)

This gives us a codegree analogue of the Turán density for 3-graphs.

PROBLEM 4. *Let $\mathcal{F}$ be a family of nonempty 3-graphs. Determine the codegree density $\gamma(\mathcal{F})$.*

What is the relationship between $\pi(\mathcal{F})$ and $\gamma(\mathcal{F})$? By counting 3-edges in two ways it is easy to show that $\gamma(\mathcal{F}) \leq \pi(\mathcal{F})$.

The first result on codegree density is due to Mubayi [30], who showed that $\gamma(F_7) = 1/2$. This gave an example where $\gamma(\mathcal{F})$ is strictly less than $\pi(\mathcal{F})$ (since de Caen and Füredi had shown that $\pi(F_7) = 3/4$). The codegree threshold for the Fano plane was determined for all sufficiently large n by Keevash [21], who used hypergraph regularity and quasirandomness to get a stability result from which he was able to proceed to the exact result via more standard combinatorial arguments. His method gave slightly more than just the codegree threshold, as it also identified exactly which 3-graphs could attain it, namely complete bipartite 3-graphs. DeBiasio and Jiang [7] later gave a simpler proof that $\mathrm{coex}(n, \mathcal{F}) = \lfloor n/2 \rfloor$ for n sufficiently large which avoided the use of regularity.

Except for the Fano plane, almost no codegree results are known for 3-graphs. Keevash and Zhao [24] studied the codegree density of projective geometries, following on earlier work of Keevash [20] on their Turán densities. Nagle [32] conjectured that $\gamma(K_4^-) = 1/4$, while Czygrinow and Nagle [5] conjectured that $\gamma(K_4) = 1/2$, with lower-bound constructions coming in both cases from random tournaments. Falgas–Ravry [10] gave nonisomorphic lower bound constructions for $\gamma(K_t)$ for general t. Recently, a subset of the authors proved $\gamma(K_4^-) = 1/4$ using flag algebras [12].

1.3. 3-graphs with independent neighborhoods. Given a 3-graph G and a pair of distinct vertices $x, y \in V(G)$, their *joint neighborhood* in G is

$$\Gamma(x, y) = \{z \in V(G) : \{x, y, z\} \in E(G)\}.$$

In an $F_{3,2}$-free 3-graph, the joint neighborhoods form independent (edge-free) subsets of the vertex set. Such 3-graphs are thus said to have *independent neighborhoods.*

As mentioned in section 1.1, the Turán density and Turán number of $F_{3,2}$ were determined by Füredi, Pikhurko, and Simonovits [14, 15], who showed that the extremal configurations were "one-way bipartite" 3-graphs.

CONSTRUCTION 1. *Given a vertex set V and a bipartition $V = A \sqcup B$, we define a one-way bipartite 3-graph $D_{A,B}$ on V by taking as the 3-edges all triples $\{a_1, a_2, b\}$ with $a_1, a_2 \in A$ and $b \in B$ (see Figure 1).*

It is easy to see that $D_{A,B}$ has independent neighborhoods and that the number of 3-edges in $D_{A,B}$ is maximized when $|A| = 2|B| + O(1)$.

THEOREM (see Füredi, Pikhurko, and Simonovits [15]). *There exists $n_0 \in \mathbb{N}$ such that if G is a 3-graph on $n \geq n_0$ vertices with independent neighborhoods and $|E(G)| = \mathrm{ex}(n, F_{3,2})$, then there exists a partition $V(G) = A \sqcup B$ of its vertex set such that $G = D_{A,B}$.*

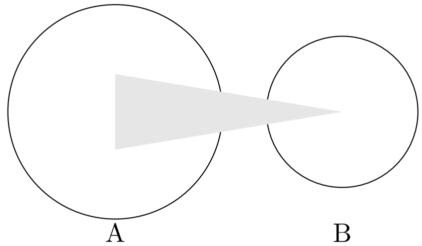

Fig. 1. *Construction* 1.

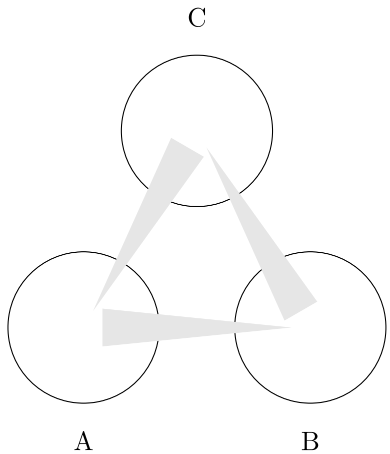

Fig. 2. *Construction* 2.

Bohman et al. [3] conjectured that a natural modification of Construction 1 was optimal for the codegree problem for $F_{3,2}$.

Construction 2. *Given a vertex set V and a tripartition $V = A \sqcup B \sqcup C$, we define a 3-graph $T_{A,B,C}$ on V by taking the union of $D_{A,B}$, $D_{B,C}$, and $D_{C,A}$ (see Figure 2).*

Again we have that $T_{A,B,C}$ has independent neighborhoods, and

$$\delta_2(T_{A,B,C}) = \min(|A|, |B|, |C|) - 1,$$

which is maximized when the three parts A, B, C are *balanced*, that is, have sizes as equal as possible. Thus $\mathrm{coex}(n, F_{3,2}) \geq \lfloor n/3 \rfloor - 1$. Bohman et al. [3] conjectured that this provides a tight lower bound for the codegree density.

Conjecture 1 (see Bohman et al. [3]).

$$\gamma(F_{3,2}) = \frac{1}{3}.$$

1.4. Results and structure of the paper. In this paper we show that

$$\mathrm{coex}(n, \{F_{3,2}\}) = \begin{cases} \lfloor n/3 \rfloor - 1 & \text{if } n \text{ is congruent to 1 modulo 3,} \\ \lfloor n/3 \rfloor & \text{otherwise} \end{cases}$$

for all n sufficiently large and determine the set of extremal configurations (which are close to but distinct from balanced $T_{A,B,C}$ configurations in general). This settles

Conjecture 1 in the affirmative and fully resolves Problems 3 and 4 for the family $\mathcal{F} = \{F_{3,2}\}$ and n sufficiently large.

We first give two proofs that the codegree density of $F_{3,2}$ is $1/3$.

THEOREM 1 (Codegree density).

$$\gamma(F_{3,2}) = \frac{1}{3}.$$

In section 2, we give a purely combinatorial proof of Theorem 1 due to Marchant, which appeared in his Ph.D. thesis [26]. In section 3, we adapt the semidefinite method of Razborov to the codegree setting to give a second proof of Theorem 1. While this second proof, a computer-assisted flag algebra calculation, is not nearly as elegant, it gives us some information about the structure of near-extremal 3-graphs. This information can be used together with a hypergraph removal lemma to prove a stability result. To state this formally, we need to make one more definition.

DEFINITION 1. *Let G and H be 3-graphs on vertex sets of size n. The* edit distance *between G and H is the minimum number of changes needed to make G into an isomorphic copy of H, where a change consists in replacing an edge by a nonedge, or vice versa.*

THEOREM 2 (Stability). *For all $\varepsilon > 0$ there exist $\delta > 0$ and $n_0 \in \mathbb{N}$ such that if G is an $F_{3,2}$-free 3-graph on $n \geq n_0$ vertices with*

$$\delta_2(G) \geq \left(\frac{1}{3} - \delta\right) n,$$

then G lies at edit distance at most $\varepsilon\binom{n}{3}$ from a balanced $T_{A,B,C}$ construction.

We use Theorem 2 in section 4 to prove our result on the codegree threshold.

THEOREM 3 (Codegree threshold). *For all n sufficiently large,*

$$\operatorname{coex}(n, \{F_{3,2}\}) = \begin{cases} \lfloor n/3 \rfloor - 1 & \text{if } n \text{ is congruent to } 1 \text{ modulo } 3, \\ \lfloor n/3 \rfloor & \text{otherwise.} \end{cases}$$

In addition we determine the set of extremal configurations. Since this set depends on the congruence class of n modulo 3 and in one case has a slightly technical description, we postpone the corresponding theorems to section 4 (Theorems 37, 39, 46, and 51).

We end the paper with a discussion of "mixed problems": given c: $0 \leq c \leq 1/3$, what is the asymptotically maximal 3-edge density ρ_c in $F_{3,2}$-free 3-graphs with codegree density at least c? We make a conjecture regarding the value of ρ_c.

2. Codegree density via extensions.

In this section, we prove that $\gamma(F_{3,2}) = 1/3$. Our strategy is similar in spirit to the one espoused by de Caen and Füredi [6] in their work on the Turán density of the Fano plane: we show that if $\delta_2(G)$ is large, then G contains either a copy of $F_{3,2}$ or a copy of some "nice subgraph" H. In the latter case we repeat the procedure using the extra assumption that H is a subgraph of G: we find again either a copy of $F_{3,2}$ or a copy of an even "nicer" subgraph, H', and so on.

Our approach is based on Lemma 4, proved in the next subsection, which establishes the existence of "nice" extensions of a subgraph in a 3-graph with high codegree. In section 2.2, we define conditional codegree density—loosely speaking, the codegree density subject to the constraint of containing a particular subgraph H. This concept then allows us to apply Lemma 4 in a very streamlined fashion in the final subsection to prove Theorem 1.

2.1. Extensions. We prove here a useful lemma, which tells us that if we have a small subgraph H inside a 3-graph G which has a high minimum codegree $\delta_2(G)$, then we can extend H to a slightly larger "nice" subgraph H' of G.

We begin with some definitions.

DEFINITION 2. *Let H be a 3-graph. A* (simple) extension *of H is a 3-graph H' with $V(H') = V(H) \cup \{z\}$ for some $z \notin V(H)$ and $E(H') \supseteq E(H)$. We denote by $L(H'; H)$ the* link graph *of the new vertex z,*

$$L(H'; H) = \{xy \in V(H)^{(2)} : \; xyz \in E(H')\}.$$

DEFINITION 3. *A sequence of 3-graphs $(G_n)_{n\in\mathbb{N}}$* tends to infinity *if $|V(G_n)| \to \infty$ as $n \to \infty$. Also, given a 3-graph H, we say that a sequence $(G_n)_{n\in\mathbb{N}}$* contains H *if all but finitely many of the 3-graphs G_n contain H as a subgraph.*

Given a set S, write $\Delta(S)$ for the $(|S| - 1)$-dimensional simplex

$$\left\{\underline{\alpha} \in [0,1]^S : \sum_{s\in S} \alpha_s = 1\right\}.$$

If H is a 3-graph and $\underline{\alpha} \in \Delta(V(H)^{(2)})$, then $\underline{\alpha}$ is a weighting on the pairs of vertices of H. We can now state and prove our key lemma.

LEMMA 4. *Let H be a 3-graph. Suppose $(G_n)_{n\in\mathbb{N}}$ is a sequence of 3-graphs tending to infinity with*

$$c = \liminf_{n\to\infty} \frac{\delta_2(G_n)}{|V(G_n)|}$$

and that $(G_n)_{n\in\mathbb{N}}$ contains H. Then, for any $\underline{\alpha} \in \Delta(V(H)^{(2)})$, there are a simple extension H' of H with

$$\sum_{xy\in L(H';H)} \alpha_{xy} \geq c$$

and a subsequence $(G_{n_k})_{k\in\mathbb{N}}$ of $(G_n)_{n\in\mathbb{N}}$ such that $(G_{n_k})_{k\in\mathbb{N}}$ contains H'.

Proof. Let $(G_n) = (G_n)_{n\in\mathbb{N}}$ be a 3-graph sequence tending to infinity with

$$c = \liminf_{n\to\infty} \frac{\delta_2(G_n)}{|V(G_n)|}.$$

Suppose H is a 3-graph contained in (G_n), and let $\underline{\alpha} \in \Delta(V(H)^{(2)})$.

We claim that for every $\varepsilon > 0$ there exists an extension H' of H such that H' is contained as a subgraph in infinitely many of the 3-graphs G_n and the weaker condition

$$\sum_{xy\in L(H';H)} \alpha_{xy} \geq c - 2\varepsilon$$

holds. This is sufficient to prove the lemma, as there are up to isomorphism only finitely many possible simple extensions of H, and so one of them must satisfy the weaker condition for all $\varepsilon > 0$.

Fix $0 < \varepsilon < 1$ and choose $N \in \mathbb{N}$ sufficiently large such that for $n \geq N$ all of the following hold:

(i) $\delta_2(G_n)/|V(G_n)| \geq c - \varepsilon$,
(ii) $|V(G_n)| \geq |V(H)|/\varepsilon$, and
(iii) H is a subgraph of G_n.

Consider a 3-graph G_n from our sequence with $n \geq N$. Fix a copy of H within G_n (we know by (iii) above that such a copy exists), and consider the weighted sum

$$s = \sum_{xy \in V(H)^{(2)}} \alpha_{xy} |\Gamma(x,y)|\,.$$

We have $s \geq (c-\varepsilon)|V(G_n)|$ by (i) above. Also,

$$\begin{aligned} s &= \sum_{z \in V(G_n)} \; \sum_{xy \in V(H)^{(2)}:\; xyz \in E(G_n)} \alpha_{xy} \\ &\leq \left(\sum_{z \in V(G_n) \setminus V(H)} \; \sum_{xy \in V(H)^{(2)}:\; xyz \in E(G_n)} \alpha_{xy} \right) + |V(H)|\,. \end{aligned}$$

Hence by averaging there exists a vertex $z \notin V(H)$ such that

$$\begin{aligned} \sum_{xy \in V(H)^{(2)}:\; xyz \in E(G^n)} \alpha_{xy} &\geq \frac{|V(G_n)|}{|V(G_n) \setminus V(H)|}(c-\varepsilon) - \frac{|V(H)|}{|V(G_n) \setminus V(H)|} \\ &\geq \frac{|V(G_n)|}{|V(G_n) \setminus V(H)|}(c - 2\varepsilon) \qquad \text{(by (ii) above)} \\ &> c - 2\varepsilon\,. \end{aligned}$$

Therefore the simple extension H' of H with vertex set $V(H) \cup \{z\}$ and 3-edges $E(H) \cup \{xyz:\ xy \in V(H)^{(2)},\, xyz \in E(G_n)\}$ satisfies our weaker condition and is a subgraph of G_n. Since there are up to isomorphism only finitely many extensions of H, one of them must satisfy the weaker condition and be contained in infinitely many of the 3-graphs in our sequence $(G_n)_{n\in\mathbb{N}}$. This concludes the proof of our claim and with it the proof of the lemma. ∎

We shall sometimes write $w_{\underline{\alpha}}(L(H';H))$, or simply $w(L)$, for $\sum_{xy \in L(H';H)} \alpha_{xy}$. This quantity $w(L)$ is exactly the total weight of the pairs picked up by the new vertex in the extension with respect to the weighting $\underline{\alpha}$.

2.2. Conditional codegree density. Our arguments in the proof of Theorem 1 are of the form "if G contains H and $\delta_2(G)$ is large, then G must contain a copy of a member of $\mathcal{F}$." It is thus natural to make the following definition.

DEFINITION 4. *Let H be a 3-graph, and let $\mathcal{F}$ be a family of nonempty 3-graphs. The* conditional codegree threshold *of $\mathcal{F}$ given H, denoted by* $\mathrm{coex}(n, \mathcal{F}|H)$*, is the maximum of $\delta_2(G)$ over all n-vertex, $\mathcal{F}$-free 3-graphs G which contain a copy of H as a subgraph.*

Our aim in this subsection is to show that we can define a conditional codegree density from this, in other words that the sequence $\mathrm{coex}(n, \mathcal{F}|H)/n$ tends to a limit as $n \to \infty$. This will be very similar to the proof that the usual codegree density is well defined [31].

LEMMA 5. *Let H be a 3-graph, and let $\varepsilon > 0$. Then there exists an integer $N = N(\varepsilon, H)$ such that for all $n, n' \in \mathbb{N}$ with $N \leq n' \leq n$ every 3-graph G on n*

vertices containing a copy of H has a subgraph G' on n' vertices also containing a copy of H and satisfying

$$\frac{\delta_2(G')}{n'} > \frac{\delta_2(G)}{n} - \varepsilon$$

(this is just saying that G' has "codegree density" almost as large as G).

Proof. Let H be a 3-graph on h vertices, and let $\varepsilon > 0$. Suppose G is a 3-graph on n vertices containing a copy of H. We form an n'-vertex subgraph of G by fixing a copy of H in G and extending it by adding $n' - h$ vertices selected uniformly at random from the rest of G. Let G' denote the resulting (random) induced subgraph of G. Clearly, G' contains a copy of H and has the right order. Now let us show that, provided n and n' are sufficiently large, G' also has a good chance of having a reasonably high minimal codegree.

Let $P_1, P_2, \dots, P_{\binom{n'}{2}}$ be a random enumeration of the pairs of vertices from $V(G')$. Note that, conditional on $P_i = xy$, the set $V(G') \setminus (P_i \cup V(H))$ is distributed as a uniformly chosen random subset of $V(G) \setminus (P_i \cup V(H))$ of size $n' - |V(H) \cup P_i| \geq n' - h - 2$.

For each $i:\ 1 \leq i \leq \binom{n'}{2}$ and $t \in \mathbb{N}$, we have

$$\begin{aligned}\mathbb{P}(d_{G'}(P_i) \leq t) &\leq \sum_{xy \in V(G)^{(2)}} \mathbb{P}(P_i = xy)\mathbb{P}\Big(\big|\big(V(G') \cap \Gamma(x,y)\big) \setminus (P_i \cup V(H))\big| \leq t \Big| P_i = xy\Big)\\ &\leq \mathbb{P}(X \leq t),\end{aligned}$$

where X is the hypergeometric random variable

$$X \sim \text{Hypergeometric}\left(n' - 2 - h, \delta_2(G) - h, n - h\right).$$

(Recall that the Hypergeometric(s, t, N) distribution with parameters $s, t \leq N$ is obtained as follows: fix a t-subset A of an N-set. Then pick an s-set B from the same N-set uniformly at random; the Hypergeometric(s, t, N) distribution is the distribution of the number of elements of A included in B.)

Now, provided n, n' are both sufficiently large,

$$\mathbb{E}(X) \geq \frac{n'}{n}\delta_2(G) - \frac{\varepsilon}{2}n'.$$

We can now use a standard Chernoff-type bound for the hypergeometric distribution (see, for example, Lemma 2 in [18]) to show that the probability that P_i is a low codegree pair in G' is small.

$$\begin{aligned}\mathbb{P}\left(d_{G'}(P_i) \leq \frac{n'}{n}\delta_2(G) - \varepsilon n'\right) &\leq \mathbb{P}\left(X \leq \mathbb{E}(X) - \frac{\varepsilon n'}{2}\right)\\ &\leq \exp\left(\frac{-(\varepsilon n'/2)^2}{\mathbb{E}(X)/2}\right)\\ &\leq \exp\left(\frac{-\varepsilon^2 n'}{2}\right).\end{aligned}$$

Summing over all $\binom{n'}{2}$ pairs P_i from $V(G')$ and using the union bound, we deduce that

$$\mathbb{P}\left(\delta_2(G') \leq \frac{n'}{n}\delta_2(G) - \epsilon n'\right) \leq \binom{n'}{2}\exp\left(\frac{-\varepsilon^2 n'}{2}\right).$$

For n' sufficiently large, this is strictly less than 1. Thus with strictly positive probability G' satisfies $\delta_2(G')/n' > \delta_2(G)/n - \varepsilon$ as required, and in particular a good choice of G' exists. ∎

With Lemma 5 in hand, we can now prove the main result of this section.

PROPOSITION 6. *For all* 3*-graphs* H *and all families of nonempty* 3*-graphs* $\mathcal{F}$ *not containing* H*, the sequence* $\mathrm{coex}(n, \mathcal{F}|H)/n$ *tends to a limit as* $n \to \infty$*.*

Proof. Let H be a 3-graph, and let $\mathcal{F}$ be a family of nonempty 3-graphs which does not contain H. Set

$$a_n = \frac{\mathrm{coex}(n, \mathcal{F}|H)}{n}.$$

We shall show that $(a_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and hence is convergent in $[0, 1]$.

Pick $\varepsilon > 0$, and let $N = N(\varepsilon, H)$ be the integer whose existence is guaranteed by Lemma 5. Let $n, n' \in \mathbb{N}$ be integers with $n \geq n' \geq N$. Suppose G is an n-vertex $\mathcal{F}$-free 3-graph containing a copy of H with $\delta_2(G) = \mathrm{coex}(n, \mathcal{F}|H)$. By Lemma 5, G has an n'-vertex subgraph G' which contains a copy of H and satisfies $\delta_2(G')/n' \geq \delta_2(G)/n - \varepsilon$. Since G is $\mathcal{F}$-free, so is G', and we thus must have

$$a_n - a_{n'} \leq a_n - \frac{\delta_2(G')}{n'} \leq a_n - \frac{\delta_2(G)}{n} + \varepsilon = \varepsilon.$$

We claim that there also exists an integer $M = M(\varepsilon, H) \geq N$ such that for all integers $n \geq M$ we have $a_M - a_n \leq \varepsilon$. Indeed, either $M_1 = N$ is a good choice of M or there exists an integer $M_2 > N$ with $a_{M_2} < a_N - \varepsilon$. Then either M_2 is a good choice of M or there exists an integer $M_3 > M_2$ with $a_{M_3} < a_{M_2} - \varepsilon$, in which case we iterate the argument. As the sequence $a_{M_1}, a_{M_2}, \ldots$ consists of real numbers from $[0, 1]$, is strictly decreasing, and has gaps between successive terms of at least ε, it can have length at most $1 + \lceil 1/\varepsilon \rceil$. Thus, after a bounded number of iterations of our argument, we find a good choice of M.

Then for any $n \geq M$ we have $|a_n - a_M| \leq \varepsilon$. It follows that $(a_n)_{n\in\mathbb{N}}$ is Cauchy as claimed and hence converges to a limit in $[0, 1]$. ∎

We may thus define the conditional codegree density of $\mathcal{F}$ given H.

DEFINITION 5. *Let* $\mathcal{F}$ *be a family of nonempty* 3*-graphs, and let* H *be a* 3*-graph not belonging to* $\mathcal{F}$*. The* conditional codegree density $\gamma(\mathcal{F}|H)$ *of* $\mathcal{F}$ *given* H *is the limit*

$$\gamma(\mathcal{F}|H) = \lim_{n\to\infty} \frac{\mathrm{coex}(n, \mathcal{F}|H)}{n}.$$

The following simple observation encapsulates the usefulness of conditional codegree densities in bounding codegree densities.

LEMMA 7. *Let* $\mathcal{F}$ *be a family of nonempty* 3*-graphs, and let* H *be a* 3*-graph not contained in* $\mathcal{F}$*. Then*

$$\gamma(\mathcal{F}) = \max\{\gamma(\mathcal{F}|H), \gamma(\mathcal{F} \cup \{H\})\}.$$

Proof. Let $c = \max\{\gamma(\mathcal{F}|H), \gamma(\mathcal{F} \cup \{H\})\}$. Clearly, we have that $\gamma(\mathcal{F}) \geq \gamma(\mathcal{F}|H)$ and $\gamma(\mathcal{F}) \geq \gamma(\mathcal{F} \cup \{H\})$, so $\gamma(\mathcal{F}) \geq c$.

Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of 3-graphs tending to infinity with $\liminf_{n\to\infty} \frac{\delta_2(G_n)}{|V(G_n)|} > c$, and let n be sufficiently large. Then, since $\gamma(\mathcal{F} \cup \{H\}) \leq c$, G_n must contain a member of $\mathcal{F}$ or H. As $\gamma(\mathcal{F}|H) \leq c$, if G_n contains H, then it must also contain a member of $\mathcal{F}$. In particular, G_n contains a member of $\mathcal{F}$. It follows that $\gamma(\mathcal{F}) \leq c$, as claimed. ∎

2.3. Proof of Theorem 1. For an integer t, the *blow-up* $F(t)$ of a 3-graph F is the 3-graph formed by replacing each vertex v of F by a set S_v of t new vertices and placing for each 3-edge $\{x, y, z\} \in E(F)$ all t^3 triples meeting each of S_x, S_y, and S_z in one vertex. If $\mathcal{F}$ is a family of 3-graphs, then its *blow-up* $\mathcal{F}(t)$ is defined to be the family $\{F(t) : F \in \mathcal{F}\}$.

Just like the ordinary Turán density, the codegree density γ exhibits *blow-up invariance*: the codegree density of a finite family is the same as the codegree density of its blow-up. This fact was reproved by several researchers; see, e.g., [24, 25, 31].

LEMMA 8 (see [24, 25, 31]). *Let $\mathcal{F}$ be a finite family of 3-graphs, and let $t \in \mathbb{N}$. Then*

$$\gamma(\mathcal{F}(t)) = \gamma(\mathcal{F}).$$

Having stated this lemma, let us now define some 3-graphs we shall need in our proof of Theorem 1. Recall from the introduction that K_4 is the complete 3-graph on four vertices, and K_4^- is the 3-graph obtained from K_4 by deleting one of its 3-edges. Further, let S_k denote the *star* on $k+1$ vertices, that is, the 3-graph with vertex set $\{x, y_1, \ldots, y_k\}$ and 3-edges $\{xy_iy_j : 1 \leq i < j \leq k\}$. Note that S_3 is (isomorphic to) K_4^-.

Finally, let S'_k denote the 3-graph on $k+2$ vertices obtained by duplicating the central vertex x of the star S_k. Thus S'_k has vertex set $\{x_1, x_2, y_1, \ldots, y_k\}$ and 3-edges $\{x_1y_iy_j : 1 \leq i < j \leq k\} \cup \{x_2y_iy_j : 1 \leq i < j \leq k\}$.

Our strategy in the proof of Theorem 1 is to show that if a 3-graph G has codegree $\delta_2(G) > \left(\frac{1}{3} + \varepsilon\right)|V(G)|$ and $|V(G)|$ is large, then G contains a copy of $F_{3,2}$ or it is forced to contain copies of larger and larger stars. We make this gradual ascension towards Theorem 1 in a series of lemmas on conditional codegree density, each of which relies on applying the key lemma (Lemma 4) with a suitable weighting $\underline{\alpha}$. We shall repeatedly look for and find copies of $F_{3,2}$ inside larger 3-graphs, and it will be convenient to write "$ab|cde$" to mean that abc, abd, abe, and cde are all 3-edges (and thus that $\{abcde\}$ spans a copy of $F_{3,2}$).

LEMMA 9. $\gamma(F_{3,2}, S'_3) \leq \frac{1}{3}$.

Proof. Clearly, $\gamma(F_{3,2}, S'_3) \leq \gamma(S'_3)$, and since S'_3 is a subgraph of $K_4^-(2)$, it is enough by Lemma 8 to show that $\gamma(K_4^-) \leq 1/3$. And indeed $\mathrm{coex}(n, K_4^-) \leq n/3$ since if we take any edge xyz in a K_4^--free 3-graph, the neighborhoods $\Gamma(x, y)$, $\Gamma(x, z)$, $\Gamma(y, z)$ must be disjoint. Thus $\gamma(K_4^-) \leq 1/3$ as claimed. ◻

LEMMA 10. *Let $k \geq 3$. Then $\gamma(F_{3,2}|S'_k) \leq k/(3k-1)$.*

Proof. Suppose $(G_n)_{n\in\mathbb{N}}$ is a 3-graph sequence tending to infinity and containing S'_k with

$$\liminf_{n\to\infty} \frac{\delta_2(G_n)}{|V(G_n)|} > \frac{k}{3k-1}.$$

Denote the vertices of S'_k by $V(S'_k) = \{x_1, x_2, y_1, \ldots y_k\}$ as before, and partition the collection of pairs $V(S'_k)^{(2)}$ into the three sets $P_1 = \{x_1x_2\}$, $P_2 = \{x_iy_j : 1 \leq i \leq 2,\ 1 \leq j \leq k\}$, and $P_3 = \{y_iy_j : 1 \leq i < j \leq k\}$.

We shall apply Lemma 4 using the following weight vector $\underline{\alpha} \in \Delta(V(S'_k)^{(2)})$:

$$\alpha_{uv} = \begin{cases} \frac{k-1}{3k-1} & \text{if } uv \in P_1\,, \\ \frac{1}{6k-2} & \text{if } uv \in P_2\,, \\ \frac{2}{(k-1)(3k-1)} & \text{if } uv \in P_3\,. \end{cases}$$

Lemma 4 guarantees that there is an extension H of S'_k for which

$$w_{\underline{\alpha}}(L(H;S'_k)) = \sum_{uv\in L(H;S'_k)} \alpha_{uv} \geq \liminf_{n\to\infty} \frac{\delta_2(G_n)}{|V(G_n)|} > \frac{k}{3k-1}$$

and an infinite subsequence $(G_{n_k})_{k\in\mathbb{N}}$ such that $(G_{n_k})_{k\in\mathbb{N}}$ contains H.

We now show that H must contain $F_{3,2}$ to conclude the proof of the lemma. This is essentially case-checking. Write L for the set $L(H;S'_k)$, w for $w_{\underline{\alpha}}$, and z for the vertex added to S'_k to form H.

Case 1. Suppose that L contains the single pair x_1x_2 from P_1. If L contains any pair y_iy_j from P_3, then $y_iy_j|x_1x_2z$, so that we have a copy of $F_{3,2}$, as claimed. On the other hand, if P_3 contains no edge of L, then consider $|L\cap P_2|$. If this is at least three, then at least one of the vertices x_1, x_2, without loss of generality x_1, must be incident to at least two edges of $L\cap P_2$. Let two such edges be x_1y_i and x_1y_j. Then $zx_1|x_2y_iy_j$, so that again we have a copy of $F_{3,2}$, as claimed. Finally, note that if $L\cap P_3=\emptyset$ and $|L\cap P_2|\leq 2$, then

$$w(L) \leq \frac{(k-1)|L\cap P_1|}{3k-1} + \frac{|L\cap P_2|}{2(3k-1)} \leq \frac{k}{3k-1},$$

contradicting the fact that $w(L) > k/(3k-1)$. Thus we are done in this case.

Case 2. Suppose that L does not contain x_1x_2 but contains at least one edge from P_2. Without loss of generality, let x_1y_i be one such edge.

If y_i is incident to two edges $y_iy_{j_1}$ and $y_iy_{j_2}$ of $L\cap P_3$, then $zy_i|x_1y_{j_1}y_{j_2}$ and we have a copy of $F_{3,2}$, as required. On the other hand, if $L\cap P_3$ contains at least one edge $y_{j_1}y_{j_2}$ not incident to y_i, then $x_1y_i|zy_{j_1}y_{j_2}$, again spanning a copy of $F_{3,2}$.

Now if L contains exactly one edge y_iy_j from P_3, then all edges in $L\cap P_2$ are incident with one of y_i, y_j. In particular, $|L\cap P_2|\leq 4$ and

$$\begin{aligned} w(L) &= \frac{|L\cap P_2|}{2(3k-1)} + \frac{2|L\cap P_3|}{(k-1)(3k-1)} \\ &\leq \frac{2}{3k-1} + \frac{2}{(k-1)(3k-1)} \\ &= \frac{k}{(3k-1)}\frac{2}{(k-1)} \leq \frac{k}{3k-1} \qquad \text{(since } k\geq 3\text{)}, \end{aligned}$$

a contradiction. On the other hand, if L contained no edge from P_3, then

$$w(L) = \frac{|L\cap P_2|}{2(3k-1)} \leq \frac{k}{3k-1}\,,$$

again a contradiction of our assumption that $w(L) > k/(3k-1)$.

Case 3. Finally, suppose that L contains no edge from P_1 or P_2. Then $L\subseteq P_3$, and

$$w(L) \leq \frac{2|P_3|}{(k-1)(3k-1)} = \frac{k}{3k-1},$$

contradicting our assumption that $w(L) > k/(3k-1)$.

It follows that H must contain a copy of $F_{3,2}$, as claimed. ∎

LEMMA 11. *Let $k\geq 3$. Then $\gamma(F_{3,2}, S_{k+1}, K_4|S'_k)\leq 1/3$.*

Proof. This is very similar to the proof of Lemma 10. Suppose $(G_n)_{n\in\mathbb{N}}$ is a 3-graph sequence tending to infinity which contains S'_k and satisfies

$$\liminf_{n\to\infty} \frac{\delta_2(G_n)}{|V(G_n)|} > \frac{1}{3}\,.$$

Denote the vertices of S'_k by $V(S'_k) = \{x_1, x_2, y_1, \ldots, y_k\}$ as before, and partition $V(S'_k)^{(2)}$ into the three sets $P_1 = \{x_1x_2\}$, $P_2 = \{x_iy_j : 1 \le i \le 2,\, 1 \le j \le k\}$, and $P_3 = \{y_iy_j : 1 \le i < j \le k\}$.

We apply Lemma 4 with a slightly different weighting. Let $\underline{\alpha}$ be defined by

$$\alpha_{uv} = \begin{cases} \frac{k-2}{3(k-1)} & \text{if } uv \in P_1\,, \\ \frac{1}{6(k-1)} & \text{if } uv \in P_2\,, \\ \frac{2}{3k(k-1)} & \text{if } uv \in P_3\,. \end{cases}$$

Lemma 4 guarantees the existence of an extension H of S'_k with

$$w_{\underline{\alpha}}(L(H;S'_k)) = \sum_{uv\in L(H;S'_k)} \alpha_{uv} \ge \liminf_{n\to\infty} \frac{\delta_2(G_n)}{|V(G_n)|} > \frac{1}{3}$$

and of an infinite subsequence $(G_{n_k})_{k\in\mathbb{N}}$ such that $(G_{n_k})_{k\in\mathbb{N}}$ contains H.

We now show that any such extension H must contain either $F_{3,2}$, S_{k+1}, or K_4. As in the previous lemma, this is just a matter of case-checking. Write L as before for the set $L(H;S'_k)$, w for $w_{\underline{\alpha}}$, and z for the vertex added to S'_k to form H.

Case 1. Suppose $x_1x_2 \in L$. By the analysis in Case 1 of Lemma 10, we know that if L contains any edge from P_3 or at least three edges from P_2, then H contains a copy of $F_{3,2}$ and we are done. On the other hand, if neither of these happens, then

$$w(L) = \frac{(k-2)|L\cap P_1|}{3(k-1)} + \frac{|L\cap P_2|}{6(k-1)} \le \frac{k-2}{3(k-1)} + \frac{1}{3(k-1)} = \frac{1}{3},$$

contradicting our assumption that $w(L) > 1/3$.

Case 2. Suppose $x_1x_2 \notin L$, but $L\cap P_2 \neq \emptyset$. By the analysis in Case 2 of Lemma 10, we know that if L contains an edge from P_2 incident to two edges from P_3 or an edge from P_2 and a disjoint edge from P_3, then H contains a copy of $F_{3,2}$ and we are done.

Also if L contains an edge $y_{j_1}y_{j_2}$ of P_3 and two edges $x_iy_{j_1}$, $x_iy_{j_2}$ from P_2, then $zx_iy_{j_1}y_{j_2}$ forms a copy of K_4 and we are done. In addition if, for some $i \in \{1,2\}$, L contains all k edges of the form x_iy_j, then $x_i, z, y_1, \ldots .y_k$ forms a copy of S_{k+1} and we are done.

Now let us suppose none of these things happens. If L contains an edge from P_3, then $|L\cap P_2| \le 2$ and $|L\cap P_3| \le 1$ (else we have a copy of K_4 or $F_{3,2}$), and thus

$$\begin{aligned} w(L) &\le \frac{2}{6(k-1)} + \frac{2}{3k(k-1)} \\ &< 1/3 \qquad\qquad \text{(since } k \ge 3\text{)}, \end{aligned}$$

a contradiction. On the other hand, if L contains no edge from P_3, then $|L\cap P_2| \le 2(k-1)$ (else we have a copy of S_{k+1}) and

$$w(L) \le \frac{2(k-1)}{6(k-1)} = 1/3\,,$$

again a contradiction.

Case 3. Finally, suppose L contains no edge from P_1 or P_2. Then $L \subseteq P_3$ and

$$w(L) \leq \frac{2\binom{k}{2}}{3k(k-1)} = 1/3\,,$$

contradicting yet again our assumption that $w(H) > 1/3$.

It follows that H must contain a copy of one of $F_{3,2}$, K_4, or S_{k+1}, as claimed. ∎

LEMMA 12. *$\gamma(F_{3,2}|K_4(2)) \leq 1/3$.*

Proof. We shall in fact prove the slightly stronger statement that $\gamma(F_{3,2}|K_4'') \leq 1/3$, where K_4'' is the 3-graph on six vertices $\{a, b, c_1, c_2, d_1, d_2\}$ with edges $\{abc_i : i \in [2]\} \cup \{abd_i : i \in [2]\} \cup \{ac_id_j : i, j \in [2]\} \cup \{bc_id_j : i, j \in [2]\}$. In other words, K_4'' is the 3-graph formed by duplicating two distinct vertices of K_4 (and hence a subgraph of $K_4(2)$).

Suppose that $(G_n)_{n\in\mathbb{N}}$ is a 3-graph sequence tending to infinity which contains K_4'' and satisfies

$$\liminf_{n\to\infty} \frac{\delta_2(G_n)}{|V(G_n)|} > \frac{1}{3}\,.$$

We apply Lemma 4 once more, with the following weighting $\underline{\alpha}$:

$$\alpha_{uv} = \begin{cases} \frac{1}{6} & \text{if } uv \in \{ac_1, ad_1, bc_1, bd_1, c_1c_2, d_1d_2\}\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Lemma 4 guarantees the existence of an extension H of K_4'' with

$$w_{\underline{\alpha}}(L(H; K_4'')) = \sum_{uv\in L(H;K_4'')} \alpha_{uv} \geq \liminf_{n\to\infty} \frac{\delta_2(G_n)}{|V(G_n)|} > \frac{1}{3}$$

and of an infinite subsequence $(G_{n_k})_{k\in\mathbb{N}}$ such that $(G_{n_k})_{k\in\mathbb{N}}$ contains H.

We now show that any such extension H contains a copy of $F_{3,2}$ as a subgraph. Write again L for the set $L(H; K_4'')$, w for $w_{\underline{\alpha}}$, and z for the vertex added to K_4'' to form H.

Since $w(L) > 1/3$, at least three of the edges in $\{ac_1, ad_1, bc_1, bd_1, c_1c_2, d_1d_2\}$ must be contained in the link graph L. If the three edges in that set which are incident to c_1 are in L, then $zc_1|c_2ab$ and we have a copy of $F_{3,2}$. Also if $c_1c_2 \in L$ and L contains either ad_1 or bd_1, then we have either $ad_1|c_1c_2z$ or $bd_1|c_1c_2z$, and thus we have a copy of $F_{3,2}$. Similarly if $d_1d_2 \in L$ and either ac_1 or bc_1 is in L, then we have $ac_1|d_1d_2z$ or $bc_1|d_1d_2z$.

It follows in particular that if L contains c_1c_2, then we have a copy of $F_{3,2}$. In exactly the same way we are done if $d_1d_2 \in L$. So finally suppose that neither of c_1c_2 and d_1d_2 is contained in L. Then at least three of the four edges ac_1, ad_1, bc_1, bd_1 must be in. In particular we must contain a pair of nonincident edges from that set. Assume without loss of generality that ad_1 and bc_1 are both in. Then $ad_1|bc_1z$, so that we again have a copy of $F_{3,2}$, as claimed. ∎

With Lemmas 9, 10, 11, and 12 in hand, we can finally prove our codegree density result.

Proof of Theorem 1. We first show by induction on k that $\gamma(F_{3,2}, S_k') \leq 1/3$ for all $k \geq 3$.

For the base case, we know from Lemma 9 that $\gamma(F_{3,2}, S'_3) \leq 1/3$. For the inductive step, suppose we knew that $\gamma(F_{3,2}, S'_K) \leq 1/3$ for some $K \geq 3$. We know from Lemma 11 that $\gamma(F_{3,2}, K_4, S_{K+1}|S'_K) \leq 1/3$. It then follows by Lemma 7 that

$$\begin{aligned}\gamma(F_{3,2}, K_4, S_{K+1}) &= \max\Big(\gamma\left(F_{3,2}, K_4, S_{K+1}, S'_K\right), \gamma\left(F_{3,2}, K_4, S_{K+1}|S'_K\right)\Big)\\ &\leq \max\left(\gamma\left(F_{3,2}, S'_K\right), \frac{1}{3}\right) \leq \frac{1}{3}.\end{aligned}$$

Using blow-up invariance (Lemma 8), we deduce that $\gamma(F_{3,2}, K_4(2), S'_{K+1}) \leq 1/3$. Combining this with the result of Lemma 12 that $\gamma(F_{3,2}|K_4(2)) \leq 1/3$, we have by one more application of Lemma 7 that $\gamma(F_{3,2}, S'_{K+1}) \leq 1/3$.

It follows that $\gamma(F_{3,2}, S'_k) \leq 1/3$ for all $k \geq 3$, as claimed. Our codegree density result is straightforward from this: for any $k \geq 3$ we have by Lemma 7 that

$$\gamma(F_{3,2}) = \max\left(\gamma(F_{3,2}|S'_k), \gamma(F_{3,2}, S'_k)\right) .$$

We also know from Lemma 10 that $\gamma(F_{3,2}|S'_k) \leq k/(3k-1)$. Since as shown inductively above we have $\gamma(F_{3,2}, S'_k) \leq 1/3$ for all $k \geq 3$, it follows that

$$\gamma(F_{3,2}) \leq \inf_{k\geq 3}\left(\max\left(\frac{k}{3k-1}, \frac{1}{3}\right)\right) = \frac{1}{3},$$

as desired. $\square$

3. Codegree density and stability via flag algebras. In this section, we use the flag algebra method of Razborov [34, 35] to give a second proof of Theorem 1 and to obtain the stability result claimed in Theorem 2. Several good expositions of flag algebras from an extremal combinatorics perspective have already appeared in the literature [1, 19, 13, 22]. We shall therefore be rather brief, directing the reader to the aforementioned papers for details. Our proof is generated by a computer using Vaughan's *Flagmatic* package (version 2.0) [39]. A proof certificate is stored under the name `F32Codegree.js` in the ancillary folder of the arXiv version of this paper [11], which also contains the flagmatic code `F32Codegree.sage` that generated the certificate. In section 3.1 we describe the structure of the file `F32Codegree.js` and show how the information contained therein implies the desired bound $\gamma(F_{3,2}) \leq \frac{1}{3}$. Since the file is large (over 2MB) and contains integers with dozens of digits, verification of the proof requires a computer as well. In order to verify all stated properties of the proof certificate, the reader can write her own script or use the script `inspect_certificate.py` included in *Flagmatic* to do some of the verifications for her.

3.1. Structure of the proof certificate. First of all, we refer the reader to the *Flagmatic* User's Guide [38], which, among many other things, describes how combinatorial structures (including types and flags that are defined below) are stored in proof certificates.

The certificate consists of various parts. Here we describe only those that are directly needed for verifying the validity of our proof.

Part `"admissible_graphs"` lists all $F_{3,2}$-free 3-graphs on $N = 6$ vertices up to isomorphism. There are exactly 426 of them; let us denote them by $G_1, \ldots, G_{426}$.

Part `"types"` lists *types* with $2\ell < N$ vertices, i.e. (vertex-labeled) $F_{3,2}$-free 3-graphs with vertex sets $\emptyset$, [2], and [4]. For our application, we need only one representative from each class of isomorphic 3-graphs; thus the number of listed types of orders 0, 2, and 4 is, respectively, 1, 1, and 5. Let us denote them by $\tau_1, \ldots, \tau_7$, using the same ordering as in *Flagmatic*: first by the number of vertices and then lexicographically by the list of 3-edges. For example, τ_2 is the type with two (labeled) vertices and no 3-edges, while τ_7 is a vertex-labeled K_4^3.

For a type τ on $[k]$, a *τ-flag* is a $(k+1)$-tuple $(F, x_1, \ldots, x_k)$ where F is an $F_{3,2}$-free 3-graph and $x_1, \ldots, x_k \in V(F)$ are distinct vertices of F such that the map $i \mapsto x_i$ is an isomorphism between τ and the induced subgraph $F[\{x_1, \ldots, x_k\}]$. We can view a flag as a 3-graph with k labeled roots that induce a copy of τ (while the remaining vertices are treated as unlabeled). This leads to the natural definition of an *isomorphism* f between two τ-flags $(F, x_1, \ldots, x_k)$ and $(H, y_1, \ldots, y_k)$, namely an isomorphism f between the unlabeled 3-graphs F and H such that the roots are preserved, that is, $f(x_i) = y_i$ for every $i \in [k]$.

Part `"flags"` contains for each $t \in [7]$ the list of all τ_t-flags $F_1^{\tau_t}, \ldots, F_{g_t}^{\tau_t}$ with $(N + |V(\tau_t)|)/2$ vertices up to flag isomorphism. For example, if $t = 1$, then τ_t is the type with no vertices, and we have to list all unlabeled 3-graphs of order 3; clearly, there are exactly two of them (edge and nonedge). If $t = 2$, then τ_t is the (unique) 2-vertex type, and we have to list all 4-vertex 3-graphs G with two roots; for $e(G) = 0, 1, 2, 3, 4$, there are, respectively, $1, 3, 4, 3, 1$ nonisomorphic ways of placing the roots. Thus $g_2 = 12$.

For each $i \in [7]$, the certificate (indirectly) contains a symmetric $(g_i \times g_i)$-matrix Q^{τ_i}. More precisely, $Q^{\tau_i} = RQ'R^T$, where Q' is a diagonal matrix all of whose diagonal entries are positive rational numbers (listed in part `"qdash_matrices"`) and R is a rational matrix (listed in part `"r_matrices"`). This representation automatically implies that the matrix Q^{τ_i} is positive semidefinite.

Part `"axiom_flags"` lists all τ_2-flags with five vertices. Recall that τ_2 is the (unique) type with two labeled vertices. There are 154 such flags. Let us denote them by $M_1, \ldots, M_{154}$. Part `"density_coefficients"` lists nonnegative rational numbers $c_1, \ldots, c_{154}$, one for each flag M_i.

Let τ be a type on $[k]$. For two τ-flags $(F, x_1, \ldots, x_k)$ and $(H, x_1, \ldots, x_k)$, let

$$P((F, x_1, \ldots, x_k), (H, y_1, \ldots, y_k))$$

be the number of $|V(F)|$-sets X such that $\{y_1, \ldots, y_k\} \subseteq X \subseteq V(H)$ and the induced τ-flag $(H[X], y_1, \ldots, y_k)$ is isomorphic to the τ-flag $(F, x_1, \ldots, x_k)$. For example, $P((K_3^3, x_1, x_2), (G, y, z))$ is the codegree of (y, z) in G, where (K_3^3, x_1, x_2) is the single 3-edge with two roots.

Let G be an arbitrary $F_{3,2}$-free 3-graph of (large) order n.

First, we compute two parameters σ_1 and σ_2 of G using the information above. We let

$$\sigma_1 = \sum_{x_1, x_2} \Big(P\big((K_3^3, x_1, x_2), (G, x_1, x_2)\big) - \frac{n}{3} \Big) \sum_{i=1}^{154} c_i P\big(M_i, (G, x_1, x_2)\big), \tag{1}$$

where the sum is over all $n(n-1)$ choices of distinct ordered pairs (x_1, x_2) from $V(G)$. Note that if the minimum codegree of G is at least $n/3$, then $\sigma_1 \geq 0$.

The definition of σ_2 is slightly more complicated. Initially, set $\sigma_2 = 0$. Then for each $k \in \{0, 2, 4\}$ let us do the following. Enumerate all $n(n-1)\ldots(n-k+1)$ sequences $(x_1, \ldots, x_k)$ of distinct vertices in $V(G)$. If the induced type $(G[\{x_1, \ldots, x_k\}], x_1, \ldots, x_k)$ is isomorphic to some τ_i, then we add $\mathbf{p} Q^{\tau_i} \mathbf{p}^T$ to σ_2, where

$$\mathbf{p} = \big(P(F_1^{\tau_i}, (G, x_1, \ldots, x_k)), \ldots, P(F_{g_i}^{\tau_i}, (G, x_1, \ldots, x_k))\big). \tag{2}$$

Since each Q^{τ_i} is positive semidefinite, we have that $\mathbf{p}Q^{\tau_i}\mathbf{p}^T \geq 0$. Thus σ_2 is nonnegative.

Let us take some type τ on $[k]$ and two τ-flags F_1 and F_2 with, respectively, ℓ_1 and ℓ_2 vertices. Let $\ell = \ell_1 + \ell_2 - k$. Consider the sum

$$\sum_{x_1,\dots,x_k} P(F_1, (G, x_1, \dots, x_k))\, P(F_2, (G, x_1, \dots, x_k)) \tag{3}$$

over all choices of k-tuples $(x_1, \dots, x_k)$ that induce a copy of τ in G. Each term $P(F_i, (G, x_1, \dots, x_k))$ in (3) can be expanded as the sum over ℓ_i-sets X_i with $\{x_1, \dots, x_k\} \subseteq X_i \subseteq V(G)$ of the indicator function that X_i induces a τ-flag isomorphic to F_i. Ignoring the choices when X_1 and X_2 intersect outside of $\{x_1, \dots, x_k\}$, the remaining terms can be generated by choosing an ℓ-set $X = X_1 \cup X_2$ first, then choosing distinct $x_1, \dots, x_k \in X$ to form $X_1 \cap X_2$, and finally splitting the remaining vertices of X between X_1 and X_2 so that $|X_i| = \ell_i$. Clearly, the terms that we ignore contribute at most $O(n^{\ell-1})$ in total. Also, the contribution of each ℓ-set X depends only on the isomorphism class of $G[X]$. Thus the sum in (3) can be written as an explicit linear combination of the subgraph counts $P(H, G)$, where H runs over unlabeled 3-graphs with ℓ vertices, modulo an additive error term $O(n^{\ell-1})$. An explicit formula for computing this linear combination can be found in, e.g., [34, Lemma 2.3].

Thus if we expand each quadratic form $\mathbf{p}Q^{\tau_i}\mathbf{p}^T$ and take the sum over all suitable $x_1, \dots, x_k \in V(G)$, where $k = |V(\tau_i)|$, then we obtain a (fixed) linear combination of $P(G_1, G), \dots, P(G_{426}, G)$ with an additive error term of $O(n^5)$. The analogous claim holds for each term in the right-hand side of (1). Thus both σ_1 and σ_2 can be represented in this form, that is,

$$\sigma_1 + \sigma_2 = \sum_{i=1}^{426} \alpha_i P(G_i, G) + O(n^5), \tag{4}$$

where each α_i is a rational number that does not depend on n and that can be computed given the information above (namely the matrices Q^{τ_j} and the coefficients c_j). An explicit formula for α_i is rather messy, so we do not state it.

The crucial properties that our certificate possesses is that each α_i is nonpositive and that $c_2 > 0$ for the τ_2-flag `"5:123(2)"` (listed as M_2 in Part `"axiom_flags"`), which in *Flagmatic* notation denotes the 5-vertex 3-graph with one 3-edge and two vertices of that 3-edge labeled. These properties (involving rational numbers) can be verified by the scripts that come with *Flagmatic* and use exact arithmetic. Explicitly, the α_i are stored in an array by *Flagmatic*, called `problem._bounds`. Asking sage to list all strictly positive elements in that array returns the empty set. As for the value of c_2, this can be read out by using the `varproblem` script. We refer the reader to the file `F32Codegree.sage` that contains such a verification at the end.

Assuming the above properties, we are ready to prove that $\gamma(F_{3,2}) \leq \frac{1}{3}$. Suppose on the contrary that $\gamma(F_{3,2}) > 1/3 + c$ for some $c > 0$.

Let ε be an arbitrary real with $0 < \varepsilon < \frac{1}{20}$, and let n be sufficiently large. Pick an $F_{3,2}$-free 3-graph G of order n and minimum codegree at least $(\frac{1}{3} + c)n$. Given G, compute σ_1 and σ_2 as above. We already know that $\sigma_2 \geq 0$. Also, as remarked earlier, the codegree assumption implies that each summand in (1) is nonnegative, so that $\sigma_1 \geq 0$.

LEMMA 13. *Let $j \in [154]$ be such that $c_j > 0$. Write M_j^0 for the unlabeled version of M_j. Then $P(M_j^0, G) < \varepsilon\binom{n}{5}$.*

Proof. Let us derive a contradiction from assuming that $P(M_j^0, G) \geq \varepsilon\binom{n}{5}$. For each 5-set $X \subseteq V(G)$ that induces M_j^0, choose $x_1, x_2 \in X$ such that the induced τ_2-flag $(G[X], x_1, x_2)$ is isomorphic to M_j. The number of pairs (x_1, x_2) that appear for at least $\varepsilon^2\binom{n-2}{3}$ different choices of X is at least $\varepsilon^2\binom{n}{2}$: indeed, otherwise the number of sets X as above is at most

$$\varepsilon^2\binom{n}{2} \times \binom{n}{3} + \binom{n}{2} \times \varepsilon^2\binom{n-2}{3} < \varepsilon\binom{n}{5}$$

for n sufficiently large (since $\varepsilon < \frac{1}{20}$), a contradiction. Each of these $\varepsilon^2\binom{n}{2}$ pairs (x_1, x_2) contributes at least $cn \times c_j\varepsilon^2\binom{n-2}{3}$ to (1). Thus $\sigma_1 = \Omega(n^6)$, which contradicts (4). (Recall that $\sigma_2 \geq 0$, while each $\alpha_j \leq 0$.) □

Since $\varepsilon > 0$ was arbitrary, it follows that our hypothetical counterexample G satisfies $P(M_j^0, G) = o(n^5)$ for each $j \in [154]$ with $c_j > 0$. In particular, $P(H, G) = o(n^5)$, where H is the 5-vertex 3-graph with exactly one edge.

We now use the *random sparsification* trick, as in [17, section 4.3]. Namely, fix p with $0 < p < \min\left(\frac{c}{4}, \frac{1}{2}\right)$ and let G' be obtained from G by deleting each edge with probability p. Then it is not hard to show (cf. Lemma 5) that with high probability $\delta_2(G') \geq (1/3 + c - 2p)n > (1/3 + c/2)n$. We know that G' is $F_{3,2}$-free (since G is). Also, as $|E(G)| = \Omega(n^3)$, G has $\Omega(n^5)$ 5-sets that span at least one edge. Each such set produces a copy of H in G' with probability at least $p^{\binom{5}{3}}$, which is small but strictly positive. In particular, with high probability $P(H, G') = \Omega(n^5)$: a typical outcome G' leads to a contradiction. Thus $\gamma(F_{3,2}) \leq \frac{1}{3}$, as claimed.

3.2. Generating the certificate. Although we have formally verified that $\gamma(F_{3,2}) \leq \frac{1}{3}$, let us briefly describe the steps that led to the certificate. As we already noted, the ancillary folder of [11] also contains the flagmatic code `F32Codegree.sage` that generated it as well as the transcript of the whole session (file `F32Codegree.txt`).

The method of using positive semidefinite matrices Q^{τ_i} to obtain inequalities between subgraph densities is fairly standard by now and has been used for a number of other problems. The new ingredient is the (rather obvious) idea of using (1) for deriving consequences of the codegree assumption $\delta_2(G) \geq \frac{1}{3}n$, namely that $\sigma_1 \geq 0$ for any choice of nonnegative coefficients c_i. The verification that each α_i can be made nonpositive can be done via semidefinite programming. More specifically, one can create an unknown block-diagonal matrix $X \succeq 0$ whose blocks are $Q^{\tau_1}, \ldots, Q^{\tau_7}$, followed by $c_1, \ldots, c_{154}$ as diagonal entries. Also, we added the extra restriction $c_1 + \cdots + c_{154} = 1$, to avoid the trivial solution when all unknowns are zero. This is done automatically by the function `make_codegree_problem`. The full support of general "axioms" (such as the codegree assumption) is not implemented in version 2.0 of *Flagmatic*. Hopefully, this will be done in future releases.

The choice $N = 6$ came from experimenting with the above approach (as $N = 5$ was not enough). Our experiments also suggested that the types τ_1 (empty vertex set) and τ_5 (two 3-edges on four vertices) are not really needed, that is, we can let Q^{τ_1} and Q^{τ_5} be the zero matrices (thus making the rounding step easier, as we will have fewer parameters). This was done by the command `set_inactive_types`.

A crucial observation for the rounding procedure is that any flag algebra proof as above has to satisfy some relations. Namely, if we run our flag algebra argument on an almost extremal example $G = T_{V_1,V_2,V_3}$ with $|V_i| = n/3$, then all the inequalities we obtain are tight up to an $O(n^5)$ additive error. This has a number of consequences.

Call a 3-graph G_i of order 6 *sharp* if $\alpha_i = 0$. The following lemma tells us a number of graphs must necessarily be sharp.

LEMMA 14. *If a 6-vertex 3-graph G_i is isomorphic to an induced subgraph of some $T_{A,B,C}$ construction, then G_i is sharp.*

Proof. Let G be a balanced $T_{A,B,C}$ construction on n vertices. Since G_i is an induced 6-vertex subgraph of a $T_{A,B,C}$ construction, it readily follows that $P(G_i,G) = \Omega(n^6)$. Now the minimum codegree in G is at least $n/3-2$, whence $\sigma_1(G) \geq -O(n^5)$. By definition, $\sigma_2(G) \geq 0$. Thus we have $\sigma_1(G)+\sigma_2(G) \geq -O(n^5)$. Since $\alpha_j \leq 0$ for all $j \in [426]$, equality (4) then implies that $-O(n^5) \leq \alpha_i P(G_i,G)$. As $P(G_i,G) = \Omega(n^6)$, we must have $\alpha_i = 0$, as claimed. ◻

LEMMA 15. *Let τ_i be a type on $k \in \{0,2,4\}$ vertices $x_1,\ldots,x_k$ which appears as an induced subgraph in a $T_{A,B,C}$ construction.*

Form $\mathbf{p}$ *as in* (2)*, with G a balanced $T_{A,B,C}$ construction on n vertices, and write $\|\mathbf{p}\|$ for its ℓ_2-norm. Then the limit of $\mathbf{p}/\|\mathbf{p}\|$ as $n\to\infty$ is a zero eigenvector of Q^{τ_i}.*

Proof. Let G be a balanced T_{V_1,V_2,V_3} construction on n vertices. The codegrees of pairs from $V(G)$ vary between $\lfloor n/3\rfloor - 1$ and $\lceil n/3\rceil$, so that $|\sigma_1(G)| = O(n^5)$. Now for all G_i which are 6-vertex subgraphs of G we have by Lemma 14 that $\alpha_i = 0$, while for all other 6-vertex 3-graphs G_i we have $P(G_i,G) = 0$. Equality (4) thus tells us that $O(n^5) + \sigma_2(G) = O(n^5)$, whence we deduce that $\sigma_2(G) = O(n^5)$.

Now for each $k \in \{0,2,4\}$ there are 3^k sequences $\boldsymbol{\epsilon} = (\epsilon_1,\epsilon_2,\ldots,\epsilon_k)$ with $\epsilon_i \in \{1,2,3\}$. Call a sequence of vertices $(x_1,\ldots x_k)$ an $\boldsymbol{\epsilon}$*-sequence* if $x_i \in V_{\epsilon_i}$ for every i. For every $\boldsymbol{\epsilon} \in \{1,2,3\}^k$, there exists a unique type τ_i (which, obviously, embeds into $T_{A,B,C}$ constructions) such that for every $\boldsymbol{\epsilon}$-sequence $(x_1,\ldots x_k)$, $(G[\{x_1,\ldots x_k\}],x_1, \ldots,x_k)$ is isomorphic to τ_i. What is more, for every such $\boldsymbol{\epsilon}$-sequence the vector $\mathbf{p}$ formed as in (2) is identical (depends on $\boldsymbol{\epsilon}$ but not on the choice of the x_i).

Fix $\boldsymbol{\epsilon} \in \{1,2,3\}^k$. By the nonnegativity of the summands contributing to $\sigma_2(G)$, we deduce that the sum of $\mathbf{p}Q^{\tau_i}\mathbf{p}^T$ over all $\boldsymbol{\epsilon}$-sequences is at most $O(n^5)$. Now this latter sum consists of $\Omega(n^k)$ identical terms, and $\|\mathbf{p}\| = \Omega(n^{3-\frac{k}{2}})$. It follows that

$$\begin{aligned} 0 \leq \frac{\mathbf{p}}{\|\mathbf{p}\|}Q^{\tau_i}\frac{\mathbf{p}^T}{\|\mathbf{p}\|} &= \mathbf{p}Q^{\tau_i}\mathbf{p}^T \times O(n^{k-6}) \\ &\leq O\left(\frac{\sigma_2(G)}{n^k}\right) \times O(n^{k-6}) \\ &= O(n^{-1}) = o(1). \end{aligned}$$

It is straightforward to see that for each $\boldsymbol{\epsilon} \in \{1,2,3\}^k$ the (unique) vector $\mathbf{p}/\|\mathbf{p}\|$ which can be formed from $\boldsymbol{\epsilon}$-sequences converges to a limit as $n\to\infty$. It follows from the inequality above and the positive semidefiniteness of Q^{τ_i} that this limit is a zero eigenvector of Q^{τ_i}, as claimed. ◻

In addition to the above, some further "forced" identities can be derived.

LEMMA 16. *Let T' be obtained from a T_{V_1,V_2,V_3} construction with $|V_i| \geq 6$ for each i by adding an extra "tripartite" 3-edge $\{u_1,u_2,u_3\}$ with $u_i \in V_i$. If a 6-vertex 3-graph G_i is isomorphic to an induced subgraph of T', then G_i is sharp.*

Proof. We may assume that G_i contains the tripartite 3-edge $\{u_1,u_2,u_3\}$, for otherwise it is isomorphic to an induced subgraph of T_{V_1,V_2,V_3} and we are done by Lemma 14.

Now let G be obtained from T_{V_1,V_2,V_3} with $|V_1| = |V_2| = |V_3| = n/3$ by adding the complete 3-partite 3-graph with parts $U_1 \cup U_2 \cup U_3$, where $U_i \subseteq V_i$ has size εn for some small $\varepsilon > 0$. This 3-graph is not $F_{3,2}$-free, but nothing prevents us from computing σ_1 and σ_2 (which are still nonnegative) using the same formulae as before. When we expand $\sigma_1+\sigma_2$ as in (4), the coefficients $\alpha_1,\ldots,\alpha_{426}$ will be the same, but we will have an extra sum $\sum_H \beta_H P(H,G)$ where H runs over 6-vertex 3-graphs, each containing a copy of $F_{3,2}$. While we have no control over the sign of each β_H, we know that they are constants independent of n. Also, we have $P(H,G) \leq (3\varepsilon)^4 n^6$. (Indeed, each H-subgraph of G has to use at least four vertices from $U = U_1 \cup U_2 \cup U_3$ because each copy of $F_{3,2} \subseteq G$ uses at least two added edges.)

Since ε can be arbitrarily small, the terms of order $O(\varepsilon^3 n^6)$ in the new version of (4) should have correct signs to avoid a contradiction. (There are no new terms of order εn^6 or $\varepsilon^2 n^6$, as we need to hit at least three vertices of U to detect an added 3-edge.) For our G_i, we have that $P(G_i, G) = \Omega(\varepsilon^3 n^6)$. Indeed, take an arbitrary embedding $f : V(G_i) \to V(G)$ and modify it to obtain an embedding f' such that for every $x \in V(G_i)$, $f'(x), f(x)$ are always in the same part V_i and $f'(x) \in U_i$ if and only if $f(x) \in U_i$. The resulting map $f' : V(G_i) \to V(G)$ gives us another embedding of G_i into G. Clearly, there are at least $(1 - o(1))(\varepsilon n)^3 (n/3)^3$ possible ways to choose f'. Thus necessarily $\alpha_i = 0$ (otherwise we would violate the nonnegativity of $\sigma_1 + \sigma_2$), and G_i is sharp, as claimed. □

We call the additional 3-edge $\{u_1, u_2, u_3\}$ in Lemma 16 a *phantom edge*. Such edges can appear in an extremal configuration but with density $o(1)$. Although sparse, they also force further sharp graphs, as shown in Lemma 16. Similarly it can be shown that they force some further zero eigenvectors in addition to those given by Lemma 15.

This phenomenon was first observed in [33, section 3.4]. A new idea here is that the "test" 3-graph G in the proof of Lemma 16 is not admissible.

The option `phantom_edge` (new in *Flagmatic* 2.0) tells the computer to use these extra identities at the rounding step.

There happened to be some further zero eigenvectors beyond those given by the observations above. Here we just guessed their values by inspecting the floating point solution and passed the information on to *Flagmatic* via its `add_zero_eigenvectors` function.

3.3. Stability. In this section we prove Theorem 2. Let G be an arbitrary $F_{3,2}$-free 3-graph on $[n]$ with minimum codegree $(1/3+o(1))n$. We shall use the information from our flag algebraic proof of Theorem 1 to establish that G lies within edit distance $o(n^3)$ of a balanced $T_{A,B,C}$ construction. First, let us show that almost all 6-vertex subgraphs of G are sharp 3-graphs.

LEMMA 17. *If a 6-vertex 3-graph G_i is not sharp, then $P(G_i, G) = o(n^6)$.*

Proof. Since $\delta_2(G) = n/3 + o(n)$, we have $\sigma_1(G) \geq -o(n^6)$. We know that $\sigma_2(G) \geq 0$ and that $\alpha_j \leq 0$ for all $j \in [426]$. Equality (4) thus implies that $-o(n^6) \leq \alpha_i P(G_i, G)$. Since G_i is not sharp, we have $\alpha_i < 0$, from which we deduce that $P(G_i, G) = o(n^6)$, as claimed. □

By applying a version of an Induced Removal Lemma (see [36] for a very strong version as well as a historical account), we can therefore change $o(n^3)$ edges of G and destroy all induced copies of nonsharp 3-graphs, without creating a copy of $F_{3,2}$. Let G' denote the 3-graph thus obtained; by definition, all of the 6-vertex subgraphs of G' are sharp 3-graphs.

Now the transcript of our flag algebraic proof of Theorem 1 shows that the number of sharp 3-graphs and the number of 6-vertex 3-graphs that embed into $T_{A,B,C}$ plus a tripartite 3-edge are both 13. By Lemma 16, these two families of 6-vertex 3-graphs must therefore coincide. In fact, it is routine to check by hand that there are nine 6-vertex 3-graphs that can appear in $T_{A,B,C}$ as induced subgraphs and that by adding one tripartite 3-edge to $T_{A,B,C}$ we increase this number by four.

We deduce from this the following.

LEMMA 18. *Every* 6*-vertex set* $X \subseteq V(G')$ *admits a partition* $X = A \cup B \cup C$ *such that* $G'[X]$ *is* $T_{A,B,C}$ *with at most one tripartite* 3*-edge added.*

By removing $o(n^3)$ edges from G, we may have destroyed our minimum codegree condition, but it will still hold on average: at most $o(n^2)$ pairs can have codegree less than $(1/3 + o(1))n$ in G'.

Let us now consider the type τ_6 which is a labeling of K_4^-.

LEMMA 19. $P(K_4^-, G') = \Omega(n^4)$.

Proof. The 3-graph G' contains at least $\left(\frac{1}{3} + o(1)\right)\binom{n}{3}$ 3-edges, while it is known that $\pi(K_4^-) < \frac{1}{3}$, as shown by Matthias [27] and Mubayi [29] (the current best known upper bound is $\pi(K_4^-) \leq 0.2871$, proved by Baber and Talbot [1] using flag algebras). Our claim is thus immediate from the Removal Lemma or from supersaturation (see Erdős and Simonovits [8]). ∎

For every quadruple of vertices $abcd$ that induces K_4^- in G' (with $abc, abd, acd \in E(G')$), form the vector $\mathbf{p} = \mathbf{p}_{abcd}$ as in (2). The transcript shows that there are 24 τ_6-flags with five vertices; thus $\mathbf{p}_{abcd} \in \mathbb{R}^{24}$. Also, the transcript shows that the rank of $Q = Q^{\tau_6}$ is 23; thus the nullspace of Q is 1-dimensional. From Lemma 15 we know that the (unique up to a scaling) forced zero eigenvector $\mathbf{z}$ of Q consists of 21 entries 0 and three equal entries that correspond to the three τ_6-flags with the unlabeled vertex having the following links in $abcd$: (1) ab, ac, ad; (2) bc, bd, cd; (3) empty. Indeed, the only way we see τ_6 in T_{V_1,V_2,V_3} is when $a \in V_i$ and $b, c, d \in V_{i-1}$ for some $i \in \mathbb{Z}_3$; by choosing the unlabeled vertex x in, respectively, V_{i-1}, V_i, V_{i+1}, we get these link graphs (each appearing about $n/3$ times when each $|V_j| = n/3$). Scale $\mathbf{z}$ so that it has unit ℓ_2-norm $\|\mathbf{z}\| = 1$.

Take a spectral decomposition $Q = \sum_{i=1}^{23} \lambda_i \mathbf{f}_i^T \mathbf{f}_i$, where the $\mathbf{f}_i$ are eigenvectors of Q such that $\{\mathbf{f}_1, \dots, \mathbf{f}_{23}, \mathbf{z}\}$ forms an orthonormal basis of $\mathbb{R}^{24}$. Since $Q \succeq 0$ has rank 23, we have that each $\lambda_i > 0$. Let $\lambda = \min(\lambda_1, \dots, \lambda_{23}) > 0$, a positive constant independent of n. Since $(\mathbf{p}, \mathbf{p}) = (\mathbf{p}, \mathbf{z})^2 + \sum_{i=1}^{23} (\mathbf{p}, \mathbf{f}_i)^2$, we have

$$\mathbf{p} Q \mathbf{p}^T = \sum_{i=1}^{23} \lambda_i (\mathbf{p}, \mathbf{f}_i)^2 \geq \lambda((\mathbf{p}, \mathbf{p}) - (\mathbf{p}, \mathbf{z})^2). \tag{5}$$

Note that, for all $abcd$ inducing τ_6, we have $\|\mathbf{p}_{abcd}\|^2 = \Omega(n^2)$. We know that $\sum_{abcd} \mathbf{p}_{abcd} Q \mathbf{p}_{abcd}^T = O(n^5)$. Thus, by Lemma 19, the right-hand side of (5) is $O(n) = o(\|\mathbf{p}_{abcd}\|^2)$ for all but $o(n^4)$ quadruples $abcd$ inducing τ_6. Fix one such "typical" quadruple $abcd$, and consider $\mathbf{p} = \mathbf{p}_{abcd}$. By the cosine formula, the approximate equality

$$(\mathbf{p}, \mathbf{z})^2 = (\mathbf{p}, \mathbf{p}) + O(n) = \|\mathbf{p}\|^2 \|\mathbf{z}\|^2 (1 + o(1))$$

implies that $\mathbf{p}$ and $\mathbf{z}$ are almost collinear. It follows that $\mathbf{p} \in \mathbb{R}^{24}$ has 21 coordinates with values $o(n)$ and three coordinates taking values $(1/3 + o(1))n$ corresponding to the τ_6-flags (1)–(3) defined above. So, if we define

$$V_1 = \{x \in V(G') \mid {G'}_x[abcd] = \{ab, ac, ad\}\},$$
$$V_2 = \{x \in V(G') \mid {G'}_x[abcd] = \{bc, bd, cd\}\},$$
$$V_3 = \{x \in V(G') \mid {G'}_x[abcd] = \emptyset\},$$

then for each $i \in [3]$ we have $|V_i| = (1/3 + o(1))n$. Let $W = [n] \setminus \bigcup_{i=1}^3 V_i$. Since $|W| = o(n)$, it is sufficient to show that the induced subgraph $G'[\bigcup_{i=1}^3 V_i]$ lies within edit distance $o(n^3)$ of the 3-graph T_{V_1,V_2,V_3} to conclude our proof of Theorem 2. We shall do this via a succession of easy lemmas. We again use "$x_1x_2|y_1y_2y_3$" as a notational shorthand for the statement that the 3-edges $x_1x_2y_1, x_1x_2y_2, x_1x_2y_3$, and $y_1y_2y_3$ are all present in our graph (and thus that $\{x_1x_2y_1y_2y_3\}$ spans a copy of $F_{3,2}$, contradicting our assumption that G' is $F_{3,2}$-free).

Lemma 20. *$G'[V_1]$ and $G'[V_2]$ are empty 3-graphs.*

Proof. Indeed, if $xyz \in G'[V_1]$, then $ab|xyz$, while if $xyz \in G'[V_2]$, then $bc|xyz$, both of which are contradictions. ∎

Lemma 21. *G' has no 3-edges of the form $V_1V_2V_2$, that is, 3-edges with two vertices in V_2 and one in V_1.*

Proof. Take any $z \in V_1$ and distinct $x, y \in V_2$. Consider $G'[abcdxz]$. By Lemma 18, we have that $G'[abcdxz] = T_{A,B,C}$ plus at most one tripartite edge for some partition $abcdxz = A \cup B \cup C$. Since $G'[abcd] \cong K_4^-$, it follows that bcd are in one part, say A, and a lies in the next part, B. Since $xbc, xbd, xcd \in E(G')$, we must have $x \in B$. Likewise, $z \in A$. Thus necessarily $xzb, xzc, xzd \in E(G')$.

Likewise, $yzb, yzc, yzd \in E(G')$. So if $xyz \in E(G')$ also, then $zy|bdx$, a contradiction. ∎

Lemma 22. *All but $o(n^3)$ 3-edges of the form $V_2V_2V_3$ are in G'.*

Proof. By our observation that most (all but $o(n^2)$) pairs in G' have codegree at least $(1 + o(1))n/3$, by the fact that $|W| = o(n)$, and by Lemma 20, the 3-graph $G'[\bigcup_{i=1}^3 V_i]$ must have at least $(1 - o(1))\binom{n/3}{2} \times n/3$ 3-edges that intersect the independent set V_2 in at least two vertices. By Lemma 21, all these 3-edges are of the form $V_2V_2V_3$, giving the required result. ∎

Lemma 23. *V_3 spans $o(n^3)$ 3-edges in G'.*

Proof. By Lemma 22, for all but $o(n^2)$ $x, y \in V_2$ we have that $|V_3 \setminus \Gamma(x, y)| = o(n)$. But $\Gamma(x, y)$ is an independent set, as G' is $F_{3,2}$-free. The lemma follows. ∎

Let $i \in \{1, 2, 3\}$. We write V_{i+1} for the part coming after V_i in the cyclic order on $\{1, 2, 3\}$, so that $V_{3+1} = V_1$, $V_{1-1} = V_3$, etc.

Lemma 24. *If all but $o(n^3)$ 3-edges $V_iV_iV_{i+1}$ are in G', then all but $o(n^3)$ 3-edges $V_iV_{i+1}V_{i+1}$ are not in G'.*

Proof. By the assumption of the lemma, for all but $o(n^5)$ 5-tuples of vertices $z, z', z'' \in V_i$ and $x, y \in V_{i+1}$ we have $xzz', xzz'', yz'z'' \in E(G')$. To prevent $xz|yz'z''$, we must have $xyz \notin E(G')$. ∎

By Lemmas 22 and 24, we conclude that all but at most $o(n^3)$ 3-edges of the form $V_2V_3V_3$ are not in $E(G')$. This together with Lemma 23 implies that almost all 3-edges of the form $V_3V_3V_1$ are in G' in the same way as we showed that almost all $V_2V_2V_3$ 3-edges are in G' in Lemma 22. Now, by Lemma 24 again, we have that only $o(n^3)$ 3-edges of the form $V_1V_1V_3$ belong to $E(G')$.

Finally, to finish the proof of stability, it remains that at most $o(n^3)$ 3-edges are of the form $V_1V_2V_3$. For all but $o(n^5)$ 5-tuples $x, x' \in V_1$, $y \in V_2$, and $z, z' \in V_3$, we have $xx'y, x'zz' \in E(G')$. Thus at least one of xyz, xyz' is missing from G' (to prevent $xy|x'zz'$). However, if we had $\Omega(n^3)$ 3-edges of the form $V_1V_2V_3$, then we would have $\Omega(n^4)$ choices of x, y, z, z' with both xyz, xyz' being in $E(G')$, a contradiction.

It follows that G' (and hence G) lies within edit distance $o(n^3)$ of a balanced T_{V_1,V_2,V_3} configuration. This concludes the proof of Theorem 2.

4. The codegree threshold. In this section, we determine the codegree threshold of $F_{3,2}$ for all sufficiently large n. This is a simple (but long) chain of arguments from stability, with a slight twist at the end when we deal with the fact that the extremal constructions are not unique and depend on the congruence class of n modulo 3.

We know from Theorem 2 that almost extremal 3-graphs are close to balanced $T_{A,B,C}$ constructions. We use this fact as our starting point and analyze an extremal example G via a series of lemmas to show that in fact G is not only close to a certain fixed, balanced $T_{A,B,C}$ construction but that it consists exactly of a subgraph of this $T_{A,B,C}$ construction together with a small number of "tripartite" 3-edges. As an immediate corollary, we have that for all n sufficiently large $\mathrm{coex}(n, F_{3,2}) \leq \lfloor n/3 \rfloor$.

At that point we separate into cases corresponding to the congruence class of n modulo 3 and determine both the codegree threshold and the extremal constructions for all n sufficiently large.

4.1. The structure of almost extremal configurations. In our argument, we shall frequently need to locate potential $F_{3,2}$-subgraphs inside larger 3-graphs, and it will be convenient just as in sections 2 and 3 to write $ab|cde$ to mean that abc, abd, abe, and cde are all 3-edges (and thus that $\{abcde\}$ spans a copy of $F_{3,2}$).

Let G be a 3-graph on n vertices with independent neighborhoods and minimal codegree $\delta_2(G) \geq n/3 + o(n)$. Pick a partition of its vertex set $V(G) = V_1 \cup V_2 \cup V_3$ such that $|E(G) \setminus E(T_{V_1,V_2,V_3})|$ is minimized.

Write T for T_{V_1,V_2,V_3}. Set $B = E(G) \setminus E(T)$ to be the set of *bad* 3-edges, i.e., 3-edges which are in G and not in T, and set $M = E(T) \setminus E(G)$ to be the set of *missing* 3-edges, i.e., 3-edges which are in T but not in G.

By Theorem 2, we know that G lies at edit distance $o(n^3)$ of a balanced $T_{A,B,C}$ construction. As an easy consequence of this fact, we have the following.

LEMMA 25.
(i) $|B| = o(n^3)$,
(ii) $|M| = o(n^3)$, *and*
(iii) $|V_i| = n/3 + o(n)$ *for* $i = 1, 2, 3$.

Proof. Since the edit distance between G and a balanced $T_{A,B,C}$ construction is $o(n^3)$, we have that $|B| = o(n^3)$ (since otherwise T would not be minimizing $|E(G) \setminus E(T)|$).

Let $\alpha_i = |V_i|/n$ for $i = 1, 2, 3$. The number of 3-edges in G with at least two vertices in V_i is at most the number of 3-edges in T with this property plus the total number of bad 3-edges $|B|$. In particular the average codegree in G of pairs of vertices in V_i is at most

$$\left({\alpha_i}^2\alpha_{i+1}n^3/2 + o(n^3)\right) / \left({\alpha_i}^2 n^2/2\right) = \alpha_{i+1}n + o(n).$$

Since $\delta_2(G) \geq n/3 + o(n)$, we must in particular have $\alpha_i = 1/3 + o(1)$ for $i = 1, 2, 3$. We have thus established parts (i) and (iii) of our lemma.

Finally, for part (ii) observe that the total number of 3-edges in G satisfies

$$e(G) = \sum_{x,y \in V(G)} \frac{d(x,y)}{3} \geq \binom{n}{2} \frac{\delta_2(G)}{3} = \frac{n^3}{18} + o(n^3).$$

It then follows from (iii) and (i) that $|M| = |E(T)| - |E(G)| + |B|$ is $o(n^3)$. $\square$

Now let us analyze the *link graphs* of vertices in G. Given $x \in V(G)$, let G_x be the 2-graph on $V(G)$ with 2-edges $\{uv : \; xuv \in E(G)\}$ and let $e(G_x) = |E(G_x)|$ be the number of edges it contains. Also let $G_x[V_i]$ denote the subgraph of G_x induced by the vertices in V_i,

$$G_x[V_i] = (V_i, \{uv \in E(G_x) : \; u, v \in V_i\}),$$

and let $G_x[V_i, V_j]$ denote the bipartite subgraph of G_x on $V_i \cup V_j$ with edges $\{uv \in E(G_x) :\ u \in V_i, v \in V_j\}$.

We shall also write V_{i+1} for the part coming after V_i in the cyclic order on $\{1, 2, 3\}$, so that $V_{3+1} = V_1$.

We first prove six lemmas which show that the link graphs of *all* vertices of G look like they ought to (up to some small error) if G were a $T_{A,B,C}$ construction.

LEMMA 26. *For every $x \in V(G)$, there is at most one $i \in \{1, 2, 3\}$ for which $e(G_x[V_i]) = \Omega(n^2)$.*

Proof. Pick $x \in V(G)$, and suppose that both V_1 and V_2 contain $\Omega(n^2)$ edges of G_x. Then there are $\Omega(n^4)$ choices of pairs $yz \in E(G_x[V_1])$ and $vw \in E(G_x[V_2])$. For each such choice, at least one of the triples yzv and yzw is missing from G and lies in M (for otherwise we would have $yz|vwx$, violating the assumption that G is $F_{3,2}$-free).

Now each such forbidden triple is counted in at most n quadruples $\{v, w, y, z\}$, implying that $|M| = \Omega(n^3)$ and contradicting part (ii) of Lemma 25. ∎

LEMMA 27. *For every $x \in V(G)$, there are at most $o(n^3)$ triples w, y, z such that $wz, yz \in E(G_x)$ and w, y come from two different parts V_i, $i \in \{1, 2, 3\}$.*

Proof. Pick $x \in V(G)$, and suppose for contradiction that $\Omega(n^3)$ such triples could be found. Then in particular we can find $\Omega(n^4)$ quadruples v, w, y, z such that vz, wz, and yz all lie in $E(G_x)$ and $y \in V_i$, $v, w \in V_{i-1}$ for some $i \in \{1, 2, 3\}$.

For each such quadruple, the triple vwy is missing from G and lies in M (for otherwise we would have $xz|vwy$). As before, each such triple is counted in at most n quadruples, giving $|M| = \Omega(n^3)$ missing edges and contradicting part (ii) of Lemma 25. ∎

LEMMA 28. *For every $x \in V(G)$, exactly one of V_1, V_2, V_3 contains $\Omega(n^2)$ 2-edges of G_x.*

Proof. Pick $x \in V(G)$. By Lemma 26, we know that at most one of $e(G_x[V_1])$, $e(G_x[V_2])$, and $e(G_x[V_3])$ may be of order $\Omega(n^2)$. Assume for contradiction that all three are of order $o(n^2)$. Then for every i all but $o(n)$ vertices in V_i have $o(n)$ neighbors in $G_x[V_i]$.

Lemma 27 implies that for all but $o(n)$ vertices $z \in V_i$ at least one of $\Gamma(x, z) \cap V_{i+1}$, $\Gamma(x, z) \cap V_{i-1}$ has size $o(n)$. Thus we can partition all but $o(n)$ vertices of V_i into two parts V_i' and V_i'' satisfying the following:

- for every $z \in V_i'$, there are at most $o(n)$ $y \in V_i \cup V_{i+1}$ such that $yz \in E(G_x)$;
- for every $z \in V_i''$, there are at most $o(n)$ $y \in V_{i-1} \cup V_i$ such that $yz \in E(G_x)$.

Since for every $z \in V(G)$ the codegree of x and z in G is at least $n/3 + o(n)$, since by Lemma 25 we have $|V_i| = n/3 + o(n)$ for $i = 1, 2, 3$, and since $e(G_x[V_i]) = o(n^2)$ by assumption, it follows that for every i the following hold:

- $G_x[V_{i-1}, V_i']$ is almost complete bipartite (contains all but $o(n^2)$ of the possible 2-edges);
- $G_x[V_i'', V_{i+1}]$ is almost complete bipartite (contains all but $o(n^2)$ of the possible 2-edges).

Now if V_1' contained $\Omega(n)$ vertices, then almost all vertices in V_3 would send $\Omega(n)$ edges to $V_1' \subseteq V_1$. If follows in particular that $|V_3'| = o(n)$. Similarly, if V_1'' contained $\Omega(n)$ vertices, then it would follow that $|V_2''| = o(n)$.

Thus if both V_1' and V_1'' contained $\Omega(n)$ vertices, then there would be only $o(n^2)$ edges of G_x between V_2 and V_3. Since we are also assuming that V_3 contains only $o(n^2)$ edges of G_x, it follows that the average degree in G_x of vertices in V_3 is at most $|V_1'| + o(n)$. But now since $|V_1| = n/3 + o(n)$, and since V_1' and V_1'' are disjoint subsets of V_1 both containing $\Omega(n)$ vertices, it follows that this average degree is at most $(1-c)n/3+o(n)$ for some strictly positive constant $c>0$. For n sufficiently large, this contradicts the fact that the minimal codegree in G is at least $n/3+o(n)$ (since the degree of a vertex in G_x is its codegree with x in G).

On the other hand, if we had, for example, $|V_1'|=|V_1|+o(n)$, then all but $o(n)$ vertices from V_3 would send $\Omega(n)$ edges to V_1 in G_x, so that $|V_3|=|V_3''|+o(n)$. But now by definition of V_1' and V_3'', there are only $o(n^2)$ edges of G_x from $V_1\cup V_3$ to V_2. Since we are assuming that $e(G_x[V_2])=o(n^2)$, this implies in particular that all but $o(n)$ vertices in V_2 have degree $o(n)$ in G_x, which again contradicts the fact that $\delta_2(G)\geq n/3+o(n)$. ∎

LEMMA 29. *For every $x\in V(G)$ and every $i\in\{1,2,3\}$, we have $e(G_x[V_i])=o(n^2)$ or $e(G_x[V_i,V_{i+1}])=o(n^2)$.*

Proof. Pick $x\in V(G)$, and suppose the claim of the lemma does not hold for some i. Then we have $\Omega(n^4)$ possible choices of a quadruple $\{v,w,y,z\}$ with $vw\in E(G_x[V_i])$ and $yz\in E(G_x[V_i,V_{i+1}])$. For each such choice, at least one of the triples vyz, wyz is missing from G and lies in M (for otherwise we would have $yz|vwx$).

Each such forbidden triple is counted in at most n quadruples, so, just as in Lemmas 26 and 27, this implies that $|M|=\Omega(n^3)$, contradicting Lemma 25, part (ii). ∎

With these lemmas in hand, we can now show that G has no vertex of high *bad* or *missing* degree, where the bad degree $d_B(x)$ is just the number of bad 3-edges incident with x while the missing degree $d_M(x)$ is the number of 3-edges from M incident with x.

LEMMA 30. *For every $x\in V(G)$, $d_B(x)=o(n^2)$.*

Proof. Pick $x\in V(G)$. By Lemma 28, we may assume without loss of generality that $e(G_x[V_1])$ and $e(G_x[V_2])$ are both $o(n^2)$, while $e(G_x[V_3])=\Omega(n^2)$, just as one would expect it to be if G were a subgraph of T_{V_1,V_2,V_3} and x were chosen from V_1.

By Lemma 29, we then know that $e(G_x[V_3,V_1])=o(n^2)$. Thus for $y\in V_1$ there are on average only $o(n)$ edges of G_x joining y to vertices in $V_1\cup V_3$. On the other hand, we know from the codegree condition on G that for every $y\in V_1$ the joint neighborhood of x and y has size at least $n/3+o(n)$. Since $|V_2|=n/3+o(n)$ (Lemma 25, part (iii)), it follows that for all but $o(n)$ vertices $y\in V_1$, y is adjacent in G_x to all but at most $o(n)$ vertices $z\in V_2$. In particular, $G_x[V_1,V_2]$ is almost complete: at most $o(n^2)$ of the possible edges between V_1 and V_2 are missing.

This and Lemma 27 imply that $e(G_x[V_2,V_3])=o(n^2)$. Thus all but $o(n^2)$ edges of G_x are internal to V_3 or lie between V_1 and V_2. If $x\in V_1$, then $d_B(x)=o(n^2)$, whereas if $x\in V_2\cup V_3$, we would have $d_B(x)=\Omega(n^2)$. Since our partition $V_1\cup V_2\cup V_3$ was chosen to minimize the number of bad 3-edges, it must be that x was assigned to V_1. The claim of the lemma thus holds for x. ∎

LEMMA 31. *For every $x\in V(G)$, $d_M(x)=o(n^2)$.*

Proof. Pick $x\in V(G)$, and write $d_T(x)$ for the number of 3-edges of $T=T_{V_1,V_2,V_3}$ containing x. Since by Lemma 25 we have $|V_i|=n/3+o(n)$ for $i=1,2,3$, it readily follows that $d_T(x)=n^2/6+o(n^2)$.

Now the codegree condition $\delta_2(G)\geq n/3+o(n)$ tells us that every $y\in V(G)\setminus\{x\}$ is incident with at least $n/3+o(n)$ edges in G_x. It follows in particular that

$$e(G_x)=\frac{1}{2}\sum_y d(x,y)\geq\frac{n^2}{6}+o(n^2).$$

Thus

$$d_M(x) = d_B(x) + d_T(x) - e(G_x) \leq d_B(x) + o(n^2),$$

which by Lemma 30 is $o(n^2)$, as desired. ∎

We can now show that in fact all bad edges are *tripartite*, i.e., meet each of V_1, V_2, and V_3 in one vertex.

LEMMA 32. *For every $i \in \{1, 2, 3\}$, V_i is an independent set in G.*

Proof. Suppose for contradiction that we had a 3-edge of G entirely contained within V_i for some i. Without loss of generality, we may assume that we have $\{x, y, z\} \in E(G)$ with all of x, y, z lying in V_1. Then for every pair u, v from V_3 we have that at least one of the triples uvx, uvy, uvz is missing from G, for otherwise $uv|xyz$. There are $n^2/18 + o(n)$ such pairs uv (since $|V_3| = n/3 + o(n)$). It follows that at least one of $\{x, y, z\}$ has missing degree at least $n^2/54 + o(n)$. This contradicts Lemma 31. ∎

LEMMA 33. *For every $i \in \{1, 2, 3\}$, there are no 3-edges with two vertices in V_i and one in V_{i-1}.*

Proof. Suppose we had such a bad 3 edge—without loss of generality $xyz \in E(G)$ with $x, y \in V_3$ and $z \in V_2$. Since $\delta_2(G) \geq n/3 + o(n)$, the joint neighborhood $\Gamma(x, y)$ contains at least $n/3 + o(n)$ vertices. We know from Lemma 32 that $\Gamma(x, y) \subseteq V_1 \cup V_2$.

Suppose $|\Gamma(x, y) \cap V_1| = \Omega(n)$. Then there are $\Omega(n^2)$ $a, a' \in V_1$ such that axy and $a'xy$ are both in $E(G)$. But for such pairs, the 3-edge $aa'z$ is missing from G, since otherwise we would have $xy|aa'z$. It follows that $d_M(z) = \Omega(n^2)$, contradicting Lemma 31.

We must therefore have $|\Gamma(x, y) \cap V_1| = o(n)$, and thus by the codegree condition $|\Gamma(x, y) \cap V_2| = n/3 + o(n)$. Now consider triples w, w', w'' from V_2. For all but $o(n^3)$ triples, xyw is in $E(G)$. Also, since $d_M(x) = o(n^2)$ by Lemma 31, for all but $o(n^3)$ of such triples, both of xww' and xww'' are in $E(G)$. But then $w'w''y$ is missing from G, as otherwise we would have $xw|yw'w''$. This implies that $d_M(y) = \Omega(n^2)$, contradicting Lemma 31.

It follows that we cannot have bad 3-edges taking one vertex in V_{i-1} and two vertices in V_i. ∎

COROLLARY 34.

$$\delta_2(G) \leq \lfloor n/3 \rfloor.$$

Proof. Suppose without loss of generality that V_1 is the smallest of the three parts V_1, V_2, and V_3. Then $|V_1| \leq \lfloor n/3 \rfloor$. Now consider a pair of vertices $x, y \in V_3$. By Lemmas 32 and 33, there is no bad edge of G containing both x and y. In particular the codegree of x and y in G is at most the codegree of x and y in T, which is exactly $|V_1|$. ∎

4.2. Divisibility and tripartite matchings. By Corollary 34, we know that for n large enough $\mathrm{coex}(n, F_{3,2}) \leq \lfloor n/3 \rfloor$. Construction 2 from the introduction shows that for all n we have $\mathrm{coex}(n, F_{3,2}) \geq \lfloor n/3 \rfloor - 1$. Continuing on the work in the previous section (and reusing the previous section's notation), we now determine for n large enough which of the two possible values is the actual codegree threshold. In addition, we seek to describe the set of extremal examples. As this set depends on some divisibility conditions—specifically, on the congruence class of n modulo 3—we separate out into three cases.

Before we do so, however, let us introduce some useful terminology. Let $V_1 \sqcup V_2 \sqcup V_3$ be a tripartition of a vertex set V. A *tripartite* 3-edge is a triple $x_1x_2x_3$ with $x_i \in V_i$

for $i = 1, 2, 3$. Let F be a set of tripartite 3-edges. A pair of vertices is *overused (by F)* if it is contained in at least two 3-edges of F. Next, F is a *tripartite pair matching*, or just a *tripartite matching*, if every two elements of F intersect in at most one vertex (that is, there are no overused pairs).

PROPOSITION 35. *Let V be a set of vertices with tripartition $V = V_1 \sqcup V_2 \sqcup V_3$. Then for any tripartite pair matching F the 3-graph G on V obtained by adding the 3-edges in F to T_{V_1,V_2,V_3} is $F_{3,2}$-free.*

Proof. This is a simple check. We know that T_{V_1,V_2,V_3} is $F_{3,2}$-free. By symmetry of the construction, it is sufficient to check that for every $a, a', a''' \in V_1$, $b, b' \in V_2$, and $c \in V_3$ neither of the 5-sets $\{a, a', b, b', c\}$ and $\{a, a', a'', b, c\}$ induces a copy of $F_{3,2}$ in G. Without loss of generality, the 3-edges contained in these two 5-sets are subsets of $\{aa'b, aa'b', bb'c, abc, a'b'c\}$ and $\{aa'b, aa''b, a'a''b, abc\}$, respectively, neither of which contains a copy of $F_{3,2}$. ∎

4.2.1. The case n congruent to 0 modulo 3. When n is congruent to 0 modulo 3 and sufficiently large, the upper bound in Corollary 34 is sharp, and moreover there is a simple description of all extremal configurations.

Before we give this construction, let us recall a basic fact from graph theory. A *proper edge coloring* of a 2-graph G with m colors is a map ϕ which assigns to each edge $\{a, b\} \in E(G)$ a *color* $\phi(a, b) \in [m]$ such that edges which meet at a vertex are assigned different colors. It is trivial to check that if G is the complete bipartite 2-graph $K_{m,m} = ([2m], \{ij : i \in [m], j \in [2m] \setminus [m]\})$, then there exists a proper edge coloring of G with m colors. (Consider, e.g., $\phi(i, j) = i + j \pmod m$.) Such edge colorings are in bijective correspondence with Latin squares. We do not have an explicit description of all such structures; in fact, even the counting problem is difficult (see, e.g., [28]).

CONSTRUCTION 3 (Family $\mathcal{T}(3m)$). *Let $n = 3m$. Take disjoint sets A, B, C, each of size m. Assume, for convenience, that $C = [m]$. Let ϕ be an edge coloring of the complete bipartite 2-graph with parts A and B with m colors. Take the 3-graph $T_{A,B,C}$ and all triples abc where $a \in A$, $b \in B$, and $\phi(ab) = c$.*

It follows from the definition of proper colorings that F is a tripartite pair matching on $A \sqcup B \sqcup C$. Thus every $H \in \mathcal{T}(n)$ is $F_{3,2}$-free by Proposition 35. Furthermore, all vertex pairs in H have codegree m. It follows from Corollary 34 that H is extremal for the codegree problem for all n sufficiently large.

COROLLARY 36. *For all n divisible by 3 and sufficiently large,* $\mathrm{coex}(n, F_{3,2}) = n/3$.

What is more, every extremal configuration belongs to $\mathcal{T}(n)$.

THEOREM 37. *Let $n = 3m$ be large. Let G be an $F_{3,2}$-free 3-graph such that $v(G) = n$ and $\delta_2(G) = m$. Then $G \in \mathcal{T}(n)$.*

Proof. Let V_1, V_2, and V_3 be as in section 4.1. Consider any pair of vertices from V_1. By Lemmas 32 and 33, their joint neighborhood is a subset of V_2, so that by the codegree condition we must have $|V_2| \geq m$. Similarly, we have $|V_3|$ and $|V_1|$ both at least m, so that in fact we must have $|V_i| = m$ for $i = 1, 2, 3$. Furthermore, observe that all 3-edges taking two vertices x, x' in V_i and one in V_{i+1} must be in $E(G)$ (otherwise the pair x, x' would have codegree at most $m - 1$). So there are no missing edges in G.

Write F for the set of tripartite 3-edges of G associated with the partition $V_1 \sqcup V_2 \sqcup V_3$. We claim that F contains no overused pair. Indeed, suppose this was not the case. Without loss of generality, we would then have vertices $a \in V_1$, $b \in V_2$, and c, c' in V_3 such that abc and abc' are both in F and hence in G. Now let a' be any vertex in $V_1 \setminus \{a\}$. By the observation in the previous paragraph, both of $cc'a'$ and $aa'b$ are in $E(G)$. But then we would have $ab|cc'a'$, a contradiction.

Now let $b \in V_2$ and $c \in V_3$. We know that $|\Gamma(b,c)| \geq m$ and that $\Gamma(b,c) \subseteq V_1 \cup V_2 \setminus \{b\}$ (Lemma 33). Thus there exists at least one vertex $a = \psi_c(b) \in V_1$ with $abc \in E(G)$, and this vertex is unique (else (b,c) would be an overused pair). What is more, if b' is an element of V_2 distinct from b, then we cannot have both of $ab'c$ and abc being 3-edges of G, for otherwise F would have an overused pair $\{a,c\}$. Since there are m distinct elements in each of V_1 and V_2, it follows that for any $c \in V_3$, ψ_c is a bijection from V_2 to V_1. Finally, observe that if c and c' are distinct elements of V_3, then for any $b \in V_2$, $\psi_c(b) \neq \psi_{c'}(b)$, since otherwise $\{b, \psi_c(b)\}$ would be an overused pair for F. In particular the map ϕ assigning color c to the 2-edge $(b, \psi_c(b))$ is an edge coloring of the complete bipartite 2-graph between V_1 and V_2 using m colors.

The 3-graph G thus belongs to $\mathcal{T}(n)$, as claimed. ∎

4.2.2. The case n congruent to 2 modulo 3. When n is congruent to 2 modulo 3 and sufficiently large, the upper bound in Corollary 34 is again sharp. Extremal constructions are very similar to those in the previous case. However, there are now some 3-edges in the extremal configuration which can be deleted without lowering the minimal codegree, so that a proof of an analogue of Theorem 37 becomes more delicate.

CONSTRUCTION 4 (Family $\mathcal{T}(3m+2)$). *Pick any H from the family $\mathcal{T}(3m+3)$ that was defined by Construction 3, and remove one vertex from H.*

Clearly, any obtained 3-graph is $F_{3,2}$-free and, as is easy to check, has minimum codegree m.

COROLLARY 38. *For all n that are congruent to 2 modulo 3 and sufficiently large,* $\mathrm{coex}(n, F_{3,2}) = \lfloor n/3 \rfloor$.

THEOREM 39. *Let $n = 3m+2$ be large. Let G be an $F_{3,2}$-free 3-graph with $v(G) = n$ and $\delta_2(G) = m$. Then G is a subgraph of some $H \in \mathcal{T}(n)$.*

Proof. Let V_1, V_2, V_3 be as in section 4.1. Consider any pair of vertices from V_1. By Lemmas 32 and 33, their joint neighborhood is a subset of V_2, so that by the codegree condition we must have $|V_2| \geq m$. Similarly, we have $|V_3|$ and $|V_1|$ both at least m.

Without loss of generality, we may therefore assume that $|V_3| = m$, and $m \leq |V_i| \leq m+2$ for $i = 1, 2$. We know (Lemmas 32 and 33) that for every $b, b' \in V_2$ their joint neighborhood is a subset of V_3. By the codegree condition $\delta_2(G) = m$, it follows that *all* 3-edges taking two vertices in V_2 and one vertex in V_3 must be in $E(G)$. We claim that in addition all 3-edges taking two vertices in V_3 and one in V_1 must be in $E(G)$.

LEMMA 40. *For all $c, c' \in V_3$ and all $a \in V_1$, $acc' \in E(G)$.*

Proof. Suppose for contradiction we had a triple $acc' \notin E(G)$ with $c, c' \in V_3$ and $a \in V_1$. Consider $\Gamma(a,c)$. We know from Lemma 33 that this is a subset of $V_3 \cup V_2 \setminus \{c, c'\}$, and it must have size at least m. Since $|V_3 \setminus \{c, c'\}| = m-2$, it follows that there must be at least two vertices $b, b' \in \Gamma(a,c) \cap V_2$.

Now we know that for all $c'' \in V_3$, $bb'c'' \in E(G)$. In particular, for all $c'' \in V_3 \setminus \{c, c'\}$ the triple acc'' must also be missing from $E(G)$, since otherwise we would have $ac|bb'c''$. Running through the argument again with c'' instead of c', it follows that axy is missing for *all* possible choices of distinct $x, y \in V_3$. But then $a \in V_1$ has missing degree $d_M(a) \geq \binom{m}{2} = \Omega(n^2)$, contradicting Lemma 31. Thus all triples taking two vertices in V_3 and one vertex in V_1 must be in G. ∎

Now let F be the set of tripartite 3-edges of G associated with the tripartition $V_1 \sqcup V_2 \sqcup V_3$.

LEMMA 41. *F contains no overused pairs.*

Proof. We consider each possible type of overused pairs in turn and show that they cannot occur in G:

(i) Suppose first of all that we had an overused pair ac with $a \in V_1$, $c \in V_3$. Then there exist $b, b' \in V_2$ such that abc and $ab'c$ are both in G. But then let c' be any element of $V_3 \setminus \{c\}$. We know that both of acc', $bb'c'$ are in G (by Lemma 40 and the preceding remark), so we have $ac|bb'c'$, a contradiction.

(ii) Now suppose that we had an overused pair bc with $b \in V_2$, $c \in V_3$. Then there exist $a, a' \in V_1$ with $abc, a'bc \in E(G)$. But we know that for any $b' \in V_2 \setminus \{b\}$ we have $bb'c \in E(G)$. In particular we cannot have $aa'b' \in E(G)$, since otherwise $bc|aa'b'$. But we know that $\Gamma(a, a') \subseteq V_2$ (Lemmas 32 and 33), so this would imply that a, a' have codegree at most 1, contradicting our minimum codegree condition (provided $n \geq 8$).

(iii) Finally, suppose that we had an overused pair ab with $a \in V_1$ and $b \in V_2$. Then there exist $c, c' \in V_3$ such that $abc, abc' \in E(G)$. For any $a' \in V_1 \setminus \{a\}$, we have $a'cc' \in E(G)$ (by Lemma 40). In particular we must have $aa'b \notin E(G)$, since otherwise $ab|a'cc'$.

It then follows from our codegree assumption that $\Gamma(a, b) = V_3$. Also, for all $a' \in V_1 \setminus \{a\}$, $\Gamma(a, a') \subseteq V_2 \setminus \{b\}$. By our codegree assumption again we deduce that $|V_2| \geq m + 1$, and hence $|V_1| \leq m + 1$.

Now for all $a' \in V_1 \setminus \{a\}$ we have $\Gamma(a', b) \subseteq (V_1 \setminus \{a, a'\}) \cup V_3$, so that by the codegree assumption again there is at least one $c'' \in V_3$ such that $a'bc'' \in E(G)$. The pair bc'' is then an overused pair (used by a, a') taking one vertex in each of V_2 and V_3, contradicting (ii). $\square$

LEMMA 42. $|V_1| = |V_2| = m + 1$.

Proof. We already know that $m \leq |V_1|$ and $|V_2| \leq m+2$. Suppose for contradiction that $|V_2| = m + 2$, and thus $|V_1| = m$. For every $(a, b) \in V_1 \times V_2$, we know $\Gamma(a, b) \subseteq (V_1 \setminus \{a\}) \cup V_3$. Since $|V_1 \setminus \{a\}| = m - 1$, there must be at least one tripartite 3-edge containing the pair (a, b). Thus there must be in total at least $|V_1| \cdot |V_2| = m(m + 2)$ distinct tripartite 3-edges. Averaging over the m^2 pairs $(a, c) \in V_1 \times V_3$, we deduce that at least one such pair must be contained in at least two tripartite 3-edges, contradicting Lemma 41.

By symmetry, it also cannot be the case that $|V_1| = m + 2$ and $|V_2| = |V_3| = m$, and we are done. $\square$

For every $a, c \in V_1 \times V_3$, we have $\Gamma(a, c) \subseteq V_2 \cup (V_3 \setminus \{c\})$. Since $\delta_2(G) = m$ and $|V_3| = m$, it follows that there is at least one $b \in V_2$ such that $abc \in E(G)$. Furthermore, we know this b is unique since the set of tripartite 3-edges of G contains no overused pair. Define $\phi(a, c) = b$.

Also, $\phi^{-1}(b)$ consists of vertex-disjoint pairs (again, as there are no overused pairs). Thus ϕ corresponds to some proper $(m + 1)$-edge coloring of $V_1 \times V_3$. It is easy to see that any $(m + 1)$-edge coloring of the complete bipartite graph $K_{m+1,m}$ extends to that of $K_{m+1,m+1}$ (in fact, in the unique way). We conclude that G is a subgraph of some 3-graph in $\mathcal{T}(n + 1)$ and thus of some $H \in \mathcal{T}(n)$. This finishes the proof of Theorem 39. $\square$

Remark 43. Note that an extremal G with $|V_3| = \frac{n-2}{3}$ can have some edges of the form $aa'b$ with $a, a' \in V_1$ and $b \in V_2$ missing. Namely, if there exist $c, c' \in V_3$ such that abc and $a'bc'$ are both 3-edges of G, then we may delete $aa'b$ without lowering the codegree of G. On the other hand, for each pair $a, a' \in V_1$ we have at most one $b \in V_2$ for which $aa'b$ is missing, and similarly for every pair $(a, b) \in V_1 \times V_2$ we have at most one a' for which $aa'b$ is missing.

4.2.3. The case n congruent to 1 modulo 3. In this section, let $n = 3m+1$ be congruent to 1 modulo 3 and sufficiently large. Unlike the two previous cases, the upper bound in Corollary 34 is not sharp.

PROPOSITION 44. *For all n congruent to* 1 *modulo* 3 *and sufficiently large,* $\mathrm{coex}(n, F_{3,2}) = \lfloor n/3 \rfloor - 1$.

Proof. Let $n = 3m + 1$ be large, and let G, V_1, V_2, V_3 be as in section 4.1. Suppose for contradiction that $\delta_2(G) = m$. Consider any pair of vertices from V_1. By Lemmas 32 and 33, their joint neighborhood is a subset of V_2, so that by the codegree condition we must have $|V_2| \geq m$. Similarly, we have $|V_3|$ and $|V_1|$ both at least m, so that in fact we must have two parts of size m and one part of size $m+1$. Assume without loss of generality that $|V_3| = m+1$ and that $|V_1| = |V_2| = m$.

By the codegree condition, all edges with two vertices in V_3 and one in V_1 or two vertices in V_1 and one vertex in V_2 must be in $E(G)$. In addition, for every pair $(b, c) \in V_2 \times V_3$ we know that $\Gamma(b, c) \subseteq V_1 \cup (V_2 \setminus \{b\})$. Since (b, c) has codegree at least m and $|V_2| = m$, it follows that there exists at least one $a \in V_1$ such that $abc \in E(G)$. Summing over all possible pairs (b, c), we see that there must be at least $m(m + 1)$ tripartite 3-edges in G. But there are only m^2 distinct pairs $(a, b) \in V_1 \times V_2$. Thus there is at least one such pair appearing in at least two tripartite 3-edges; i.e., there must be $a \in V_1$, $b \in V_2$, $c, c' \in V_3$ such that both abc and abc' are in $E(G)$.

But then let a' be any vertex in $V_1 \setminus \{a\}$. By our earlier observations, we know that $aa'b$ and $cc'a'$ are both 3-edges of G, so that $ab|cc'a'$, contradicting the fact that G is $F_{3,2}$-free. ∎

A consequence of this lower codegree threshold is that the extremal structures are considerably more complicated. We present three families $\mathcal{T}_1(n)$, $\mathcal{T}_2(n)$, and $\mathcal{T}_3(n)$ of extremal 3-graphs on $[n]$ and show that for every extremal G there is some $H \in \cup_{i=1}^{3} \mathcal{T}_i(n)$ containing G as a (spanning) subgraph. One could say more about the possible structure of $E(H) \setminus E(G)$ (along the lines of Remark 43), but we do not think that this description will be very illuminating. Let us define each family $\mathcal{T}_i(n)$.

CONSTRUCTION 5 (Family $\mathcal{T}_1(3m + 1)$). *Start with $T_{A,B,C}$, where $|A| = m$, $|B| = m + 2$, and $|C| = m - 1$. Add an arbitrary set of tripartite edges so that no overused pairs are created and for every $a \in A$ and $c \in C$ there is a tripartite edge containing $\{a, c\}$.*

CONSTRUCTION 6 (Family $\mathcal{T}_2(3m + 1)$). *Let $0 \leq k \leq m + 1$. Start with $T_{A,B,C}$, where $|A| = |B| = m + 1$ and $|C| = m - 1$. Let S consist of k vertex-disjoint pairs from $A \times B$.*

Remove all 3-edges of $T_{A,B,C}$ that contain a pair from S. Add all tripartite 3-edges that contain a pair from S. Thus S is precisely the set of overused pairs now. Add an arbitrary collection of tripartite 3-edges so that no new overused pair is created and for every $a \in A$ and $c \in C$ there is at least one tripartite edge containing $\{a, c\}$. (Note that if a belongs to a pair in S, then this condition is automatically satisfied.)

CONSTRUCTION 7 (Family $\mathcal{T}_3(3m+1)$). *Start with T_{V_1,V_2,V_3}, where $|V_1| = m+1$ and $|V_2| = |V_3| = m$.*

Let S consist of pairs of vertices containing at most one pair from $V_i \times V_{i+1}$ for each $i \in [3]$ so that if $i \in \{1,3\}$ and S contains both $(x, y) \in V_{i-1} \times V_i$ and $(y', z) \in V_i \times V_{i+1}$, then $y = y'$. (Thus $0 \leq |S| \leq 3$; for example, if $|S| = 3$, then the pairs in S form either a 3-cycle or a path ending and starting in V_2.)

Remove all 3-edges from T_{V_1,V_2,V_3} that contain a pair in S. Add an arbitrary collection of tripartite 3-edges so that

- *each pair of S is contained in at least $m-1$ added edges;*
- *there are no overused pairs other than those from S; and*
- *if $|V_i| = m$ (that is, $i \in \{2,3\}$) and $(x,y) \in V_i \times V_{i+1}$ is in S, then for every $x' \in V_i \setminus \{x\}$ the pair $\{x',y\}$ is contained in exactly one tripartite edge.*

We leave it to the reader to verify that each constructed 3-graph has minimum codegree $m-1$. The following result implies that all these 3-graphs are $F_{3,2}$-free.

PROPOSITION 45. *Let V be a set of vertices with tripartition $V = V_1 \sqcup V_2 \sqcup V_3$. Let G be obtained from T_{V_1,V_2,V_3} by adding some set F of tripartite 3-edges and removing all 3-edges of T_{V_1,V_2,V_3} that contain a pair overused by F. Then G is $F_{3,2}$-free.*

Proof. By Proposition 35 we need only check for copies of $F_{3,2}$ that contain two tripartite edges sharing an overused pair, say $abc, ab'c \in F$ with $a \in V_1$, $c \in V_3$ and $b, b' \in V_2$. Each such $F_{3,2}$ has to be of form $ac|bb'x$ for some x. Now, $bb'x \in E(G)$ implies $x \in V_3$. Since (a,c) is an overused pair, we have $acx \notin E(G)$ by the definition of G. Thus we cannot have $ac|bb'x$, as desired. ∎

Examples of 3-graphs in $\mathcal{T}_1(n)$, $\mathcal{T}_2(n)$, and $\mathcal{T}_3(n)$ can be obtained by taking a 3-graph in, respectively, $\mathcal{T}(n+5)$, $\mathcal{T}(n+2)$, and $\mathcal{T}(n+2)$ and deleting arbitrary vertices so that the parts have the desired sizes. However, note that, for example, not all 3-graphs in $\mathcal{T}_2(n) \cup \mathcal{T}_3(n)$ with $S = \emptyset$ come from $\mathcal{T}(n+2)$, as there are $(m+1)$-edge colorings of $K_{m+1,m-1}$ (for $m \geq 4$) and $K_{m,m}$ (for $m \geq 2$) that do not extend to an $(m+1)$-edge coloring of $K_{m+1,m+1}$.

We shall show that the 3-graphs in $\cup_{i=1}^{3} \mathcal{T}_i(n)$ contain (as spanning subgraphs) all possible extremal configurations of order n. We know from our analysis in section 4.1 that every extremal configuration G for the codegree problem consists of a subgraph of T_{V_1,V_2,V_3} together with a set of tripartite 3-edges. Thus the minimum codegree is at most $\min(|V_i| : i \in [3])$. As $\delta_2(G) = m-1$, we must have $|V_i| \geq m-1$ for every $i \in [3]$. We separate out into two cases according to whether or not we have equality for some i.

THEOREM 46. *Let G, V_1, V_2, V_3 be as in section 4.1, and suppose $n = 3m+1$ is large and $\delta_2(G) = m-1$. If $|V_i| = m-1$ for any $i = 1,2,3$, then G is isomorphic to a subgraph of some $H \in \mathcal{T}_1(n) \cup \mathcal{T}_2(n)$.*

Proof. Without loss of generality, assume that $|V_3| = m-1$. By Lemmas 32 and 33, we have that $\Gamma(x,x') \subseteq V_3$ for every $x, x' \in V_2$. The codegree condition $\delta_2(G) \geq m-1$ then implies that all 3-edges taking two vertices in V_2 and one in V_3 are in G. In addition, we have the following.

LEMMA 47. *All 3-edges taking two vertices in V_3 and one in V_1 are in G.*

Proof. Indeed, suppose that $acc' \notin E(G)$ for some $c, c' \in V_3$ and $a \in V_1$. Since $\Gamma(c,a)$ contains at least $m-1$ vertices and is contained in $V_2 \cup V_3 \setminus \{c,c'\}$ and since $V_3 \setminus \{c,c'\}$ has size $m-3$, it follows that there exist b, b' such that abc and $ab'c$ are both in $E(G)$. But then for all $x \in V_3 \setminus \{c\}$ the 3-edge acx cannot be in G, for otherwise $ac|bb'x$. Likewise, for every $y \in V_3 \setminus \{x\}$ we have that axy is missing from G. This implies $d_M(a) \geq \binom{m-1}{2} = \Omega(n^2)$, contradicting Lemma 31. ∎

With Lemma 47 in hand, we can now turn our attention to the tripartite 3-edges of G. Write F for the tripartite 3-edges associated with the tripartition $V_1 \sqcup V_2 \sqcup V_3$.

COROLLARY 48. *$V_1 \times V_3$ contains no overused pair.*

Proof. Suppose we had $a \in V_1$, $b, b' \in V_2$, and $c \in V_3$ with $abc, ab'c \in F$. Then for all $c' \in V_3 \setminus \{c\}$ we must have acc' missing from G to prevent $ac|bb'c'$, contradicting Lemma 47 (recall that $bb'c \in E(G)$, as observed just before Lemma 47). ∎

Next we show that $V_2 \times V_3$ does not contain overused pairs either.

Lemma 49. *$V_2 \times V_3$ contains no overused pairs*

Proof. Suppose we had $a, a' \in V_1$, $b \in V_2$, and $c \in V_3$ such that abc and $a'bc$ are both in F. We know that $\Gamma(a, a') \subseteq V_2$ (by Lemmas 32 and 33), so provided n is sufficiently large (which we are assuming) there is at least one $b' \in V_2 \setminus \{b\}$ such that $aa'b' \in E(G)$. But since we also have $bb'c \in E(G)$ (as observed just before Lemma 47), this means $bc|aa'b'$, a contradiction. ∎

In particular, all overused pairs from F come from $V_1 \times V_2$.

Lemma 50. *Let $(a, b) \in V_1 \times V_2$ be an overused pair from F. Then the following hold:*

(i) $\Gamma(a, b) = V_3$*;*

(ii) $\{f \in F : a \in f\} = \{f \in F : b \in f\}$.

Proof. Let $(a, b) \in V_1 \times V_2$ be such an overused pair. Then there exist $c, c' \in V_3$ such that abc and abc' are 3-edges of G.

By Lemma 33, we know $\Gamma(a, b) \subseteq V_1 \cup V_3$. Suppose $aa'b \in E(G)$ for some $a' \in V_1$. By Lemma 47, we know $a'cc' \in E(G)$, so that $ab|a'cc'$, a contradiction. Thus $\Gamma(a, b) \subseteq V_3$, and the codegree condition $d(a, b) \geq m - 1 = |V_3|$ tells us $\Gamma(a, b) = V_3$, proving part (i) of the lemma.

Part (ii) is then immediate from Corollary 48 and Lemma 49: if $ab'c'' \in E(G)$ for some $b' \in V_2 \setminus \{b\}$ and $c'' \in V_3$, then (a, c'') is an overused pair (used by b and b') from $V_1 \times V_3$, contradicting Corollary 48; similarly, if $a'bc'' \in E(G)$ for some $a' \in V_1 \setminus \{a\}$ and $c'' \in V_3$, then (b, c'') is an overused pair (used by a and a') from $V_2 \times V_3$, contradicting Lemma 49. ∎

Note that Lemma 50 implies that the overused pairs from F are vertex-disjoint pairs from $V_1 \times V_2$.

For every pair $(a, c) \in V_1 \times V_3$, the joint neighborhood $\Gamma(a, c)$ is a subset of $V_2 \cup (V_3 \setminus \{c\})$. By the codegree condition $\delta_2(G) \geq m-1$ and the fact that $|V_3| = m-1$, it follows that for every such pair there is at least one tripartite 3-edge $abc \in F$ with $b \in V_2$. Now there are exactly $(m-1)|V_1|$ distinct such pairs $(a, c) \in V_1 \times V_3$. On the other hand, since there are no overused $V_2 \times V_3$ pairs arising from F, there can be at most $(m-1)|V_2|$ such tripartite 3-edges, one for each pair $(b, c) \in V_2 \times V_3$. Thus $|V_2| \geq |V_1|$.

If $|V_2| = |V_1| = m + 1$, then by adding all missing $V_1V_1V_2$ 3-edges to G we obtain a member of $\mathcal{T}_2(n)$, as desired.

So let us suppose that $|V_1| \leq m$. We know from our codegree condition that $|V_1| \geq m - 1$, and the inequality $|V_1| \leq m$ implies $|V_2| \geq m + 2$.

We claim that F contains no overused pair. Indeed, suppose $(a, b) \in V_1 \times V_2$ is an overused pair. By Lemma 50, part (i), $aa'b \notin E(G)$ for all $a' \in V_1 \setminus \{a\}$. For each $a' \in V_1 \setminus \{a\}$, the codegree condition then tells us that $\Gamma(a', b)$ is a subset of $(V_1 \setminus \{a, a'\}) \cup V_3$ of size at least $m - 1$. In particular there must exist $c \in V_3$ with $a'bc \in E(G)$. But this is a tripartite 3-edge containing b and not a, contradicting part (ii) of Lemma 50. Thus F has no overused pair, as claimed.

Next, suppose that $|V_1| = m - 1$. Then for every $(a, b) \in V_1 \times V_2$, $\Gamma(a, b) \subseteq (V_1 \setminus \{a\}) \cup V_3$. By the codegree assumption $\delta_2(G) \geq m - 1$, we deduce that there must be at least one tripartite 3-edge involving the pair (a, b). Thus there must be at least $|V_1| \cdot |V_2| > |V_1| \cdot |V_3|$ tripartite 3-edges in G, implying the existence of an overused pair in $V_1 \times V_3$, contradicting Corollary 48. Thus $|V_1| = m$, and hence $|V_2| = m + 2$.

As observed after Lemma 50, every pair $(a, c) \in V_1 \times V_3$ is covered by at least one tripartite 3-edge (otherwise its codegree is at most $|V_3| - 1 < m - 1$); we have already shown that there are no overused pairs in F. By adding all missing 3-edges of the form $V_1V_1V_2$ to G we thus obtain a member of $\mathcal{T}_1(n)$, as required. □

THEOREM 51. *Let G, V_1, V_2, V_3 be as in section 4.1, and suppose $n = 3m+1$ is large and $\delta_2(G) = m-1$. If $|V_i| \geq m$ for all $i \in [3]$, then G is a subgraph of some $H \in \mathcal{T}_3(n)$.*

Proof. Assume without loss of generality that $|V_1| = m+1$ and $|V_2| = |V_3| = m$. Let us show first that overused pairs are contained in tripartite 3-edges only.

LEMMA 52. *If (x, y) is an overused pair in $V_i \times V_{i+1}$, then $\Gamma(x, y) \subseteq V_{i-1}$.*

Proof. Since (x, y) is an overused pair, there exist z, z' in V_{i-1} such that xyz, xyz' are 3-edges of G. Now $\Gamma(z, z') \subseteq V_i$ (by Lemmas 32 and 33) so that by the codegree condition $\Gamma(z, z')$ contains at least $m-2$ elements of $|V_i \setminus \{x\}|$. For any such element x', $xx'y \notin E(G)$, for otherwise we would have $xy|x'zz'$. Now the joint neighborhood of x and y is contained in $V_i \cup V_{i-1}$ (Lemma 33) and has size at least $m-1$, from which it follows that

$$\begin{aligned}|\Gamma(x,y) \cap V_{i-1}| &\geq m-1-(|V_i \setminus \{x\}| - (m-2)) \\ &= 2m-3-|V_i \setminus \{x\}| \\ &\geq m-3.\end{aligned}$$

Now suppose $xx'y \in E(G)$ for some $x' \in V_i$. Then for all $w, w' \in \Gamma(x, y) \cap V_{i-1}$ we would have $x'ww' \notin E(G)$, for otherwise $xy|x'ww'$. But then $d_M(x') \geq \binom{m-3}{2} = \Omega(n^2)$, contradicting Lemma 31. Thus if (x, y) is an overused pair from $V_i \times V_{i+1}$, then $\Gamma(x, y) \subseteq V_{i-1}$. □

We now turn our attention to showing that for each $i \in \{1, 2, 3\}$ the set $V_i \times V_{i+1}$ contains at most one overused pair.

LEMMA 53. *If $|V_{i+1}| = m$ and (a, b), (a', b') are overused pairs from $V_i \times V_{i+1}$, then $b = b'$.*

Proof. Suppose not. We know by Lemma 52 that for all $a'' \in V_i$ neither $aa''b$ nor $a'a''b'$ are 3-edges of G.

If $a = a'$, then we have for any $a'' \in V_i \setminus \{a\}$ that

$$|\Gamma(a, a'')| \leq |V_{i+1} \setminus \{b, b'\}| = m-2,$$

contradicting our codegree assumption $\delta_2(G) = m-1$. On the other hand, if $a \neq a'$, then

$$|\Gamma(a, a')| \leq |V_{i+1} \setminus \{b, b'\}| = m-2,$$

contradicting again the codegree assumption. □

LEMMA 54. *Suppose (a, b) and (a', b) are overused pairs from $V_i \times V_{i+1}$. Then $a = a'$.*

Proof. By Lemma 52, we know that $\Gamma(a, b)$ and $\Gamma(a', b)$ are both subsets of V_{i-1} of size at least $m-1$. In particular, since $|V_{i-1}| \leq m+1$, we have that $\Gamma(a, b) \cap \Gamma(a', b)$ is a subset of V_{i-1} of size at least $m-3$.

Now we know from Lemma 31 that $d_M(b) = o(n^2) = o(m^2)$. Thus for all but $o(m)$ vertices $b' \in V_{i+1} \setminus \{b\}$ we have that $bb'c \in E(G)$ for all but $o(m)$ vertices $c \in \Gamma(a, b) \cap \Gamma(a, b')$.

But for such b' and c, $aa'b' \notin E(G)$, for otherwise we would have $bc|aa'b'$. Thus $\Gamma(a, a')$ (which we know is a subset of V_{i+1}) can contain at most $o(m)$ vertices, contradicting our codegree assumption for n (and hence m) sufficiently large. □

Taken together, the last two lemmas imply the following.

COROLLARY 55. *$V_1 \times V_2$ and $V_2 \times V_3$ each contain at most one overused pair.*

We now prove analogues of Lemma 53 for $V_3 \times V_1$, to show that it also contains at most one overused pair.

LEMMA 56. *Suppose (c, a) and (c, a') are overused pairs from $V_3 \times V_1$. Then $a = a'$.*

Proof. Suppose not. Then by Lemma 52 we know that $\Gamma(a, c)$ and $\Gamma(a', c)$ are subsets of V_2 of size at least $\delta_2(G) = m - 1$. We also know (Lemmas 32 and 33) that $\Gamma(a, a')$ is a subset of V_2 of size at least $\delta_2(G) = m - 1$. Thus the intersection

$$I = \Gamma(a, c) \cap \Gamma(a', c) \cap \Gamma(a, a')$$

has size at least $3(m - 1) - 2|V_2| = m - 3$.

For every distinct $b, b' \in I$, we have that $bb'c \notin E(G)$ because otherwise we have $bc|aa'b'$. But then $d_M(c) \geq \binom{|I|}{2}$, contradicting Lemma 31. ◻

LEMMA 57. *Suppose (c, a) and (c', a') are overused pairs from $V_3 \times V_1$. Then $a = a'$ and $c = c'$. (In particular, $V_1 \times V_3$ contains at most one overused pair.)*

Proof. Suppose not. The only case left over from Lemmas 54 and 56 is the case when both $a \neq a'$ and $c \neq c'$, i.e., when we have vertex-disjoint overused pairs.

By Lemma 52, we know that $\Gamma(a, c)$ and $\Gamma(a', c')$ are both subsets of V_2. Now consider an arbitrary $c'' \in V_3 \setminus \{c, c'\}$. Since $acc'' \notin E(G)$ and $|V_3 \setminus \{c, c''\}| = m - 2$, there must exist $b = b(c'') \in V_2$ such that $abc'' \in E(G)$. Similarly, there must exist $b' = b'(c'') \in V_2$ such that $a'b'c'' \in E(G)$.

Now note that if $b \in \Gamma(a, c)$, then (a, b) is overused (since both abc and abc'' are in G). Similarly, if $b' \in \Gamma(a', c')$, then (a', b') is overused.

Also, V_2 has size m while $\Gamma(a, c)$ and $\Gamma(a', c')$ both have size at least $m - 1$. So there is at most one vertex $b_\star \in V_2 \setminus \Gamma(a, c)$ and at most one vertex $b'_\star \in V_2 \setminus \Gamma(a', c')$.

We now apply the pigeonhole principle to get a contradiction for m large enough (at least 4):

- if $b(c'') = b_\star$ for at least two distinct $c'' \in V_3 \setminus \{c, c'\}$, then $(a, b_\star)$ is as overused pair;
- if $b(c'') \neq b_\star$ for at least one $c'' \in V_3 \setminus \{c, c'\}$, then $(a, b(c''))$ is an overused pair;
- if $b'(c'') = b'_\star$ for at least two distinct $c'' \in V_3 \setminus \{c, c'\}$, then $(a', b'_\star)$ is an overused pair;
- if $b'(c'') \neq b'_\star$ for at least one $c'' \in V_3 \setminus \{c, c'\}$, then $(a', b'(c''))$ is an overused pair.

Thus, provided $|V_3 \setminus \{c, c'\}| \geq 2$, we have at least two distinct overused pairs from $V_1 \times V_2$, one involving a and the other a'. This contradicts Corollary 55. ◻

We have thus shown that for every $i \in [3]$, $V_i \times V_{i+1}$ contains at most one overused pair.

LEMMA 58. *If $(x, y) \in V_i \times V_{i+1}$ is an overused pair and $|V_i| = m$, then for every $x' \in V_i \setminus \{x\}$ there is exactly one $z \in V_{i-1}$ with $\{x', y, z\} \in E(G)$.*

Proof. The joint neighborhood of x', y lies inside $V_{i-1} \cup V_i \setminus \{x, x'\}$. Since $\delta_2(G) \geq m - 1$, there must exist at least one z as required. Since $\{x', y\}$ is not an overused pair, this z is unique. ◻

LEMMA 59. *Suppose (a, c) and (b', c') are overused pairs from $V_1 \times V_3$ and $V_2 \times V_3$, respectively. Then $c = c'$.*

Proof. Suppose not. For $b'' \in V_2 \setminus \{b'\}$, let $z(b'')$ be the vertex in V_1 with $\{b'', c', z(b'')\} \in E(G)$ given by Lemma 58.

If $a' = z(b''_1) = z(b''_2)$ for some distinct $b''_1, b''_2 \in V_2 \setminus \{b'\}$, then we have that (a', c') is an overused pair from $V_1 \times V_3$ distinct from (a, c) (since $c \neq c'$), contradicting Lemma 57. Thus the map $z : V_2 \setminus \{b_1\} \to V_1$ is injective.

By Lemma 52, $\Gamma(b', c')$ is a subset of V_1 of size at least $m - 1$. As n is large, $\Gamma(b', c')$ must contain some $a' = z(b'')$. But then $a'c'b', a'c'b'' \in E(G)$, so $a'c'$ is an overused pair from $V_1 \times V_3$ distinct from (a, c) (since $c \neq c'$), again contradicting Lemma 57. □

Similarly, we have the following.

LEMMA 60. *Suppose (a, c) and (a', b') are overused pairs from $V_1 \times V_3$ and $V_1 \times V_2$, respectively. Then $a = a'$.*

Proof. The proof is identical to that of Lemma 59, with V_i playing the role of V_{i-1}. □

The above lemmas show that if we add all edges from T_{V_1,V_2,V_3} to G, we obtain an element of $\mathcal{T}_3(n)$, as claimed. □

5. Turán density subject to a codegree constraint. A natural variation of the Turán density and codegree density problems is the following.

DEFINITION 6. *Let $\mathcal{F}$ be a family of nonempty 3-graphs, and let $(c_n)_{n\in\mathbb{N}}$ be a sequence of real numbers with $c_n \in [0, \frac{\mathrm{coex}(n,\mathcal{F})}{n-2}]$ for each $n \in \mathbb{N}$. The* Turán number *of $\mathcal{F}$ subject to the codegree constraint $(c_n)_{n\in\mathbb{N}}$ is the function $\mathrm{ex}_{c_n}(\cdot, \mathcal{F})$ sending $n \in \mathbb{N}$ to the maximum number of 3-edges in an $\mathcal{F}$-free n-vertex 3-graph with minimum codegree at least $c_n(n-2)$.*

PROBLEM 5. *Let $\mathcal{F}$ be a family of nonempty 3-graphs, and let $c \in [0, \gamma(\mathcal{F}))$. Determine $\mathrm{ex}_c(n, \mathcal{F})$.*

To the best of our knowledge, Lo and Markström [25] were the first to pose a question of the kind considered in Problem 5. They asked for the behavior of $\mathrm{ex}_c(n, \mathcal{F})$ when $\mathcal{F}$ is the 3-graph K_4^-.

Problem 5 can be thought of as a way of viewing Problems 1 and 3 together within a common framework. In addition, codegree constraints are natural in the context of 3-graphs, so that Problem 5 is appealing from an extremal hypergraph perspective.

For the Fano plane F_7, Problem 5 is trivial from the work of Keevash and Sudakov [23], Füredi and Simonovits [16], and Keevash [21]: the extremal configurations for the Turán number and for the codegree threshold are identical for all n sufficiently large, so that $\mathrm{ex}_c(n, F_7) = \mathrm{ex}(n, F_7)$ for all $c \in [0, 1/2]$ and all but finitely many n.

The situation is very different for $F_{3,2}$, where codegree-extremal configurations have $n^3/18 + o(n^3)$ 3-edges, as we have shown, while the extremal configurations have $2n^3/27 + o(n^3)$ 3-edges, i.e., about one and a third times as many. A first step towards the resolution of Problem 5 for $F_{3,2}$ would be to identify the asymptotic behavior of $\mathrm{ex}_c(n, F_{3,2})$ for $c \in [0, 1/3]$.

A lower bound can be obtained by shifting weight in a continuous fashion from part A to part C in a $T_{A,B,C}$ construction, and so one can move from Construction 1 (where $|A| = \frac{2n}{3} + O(1)$, $|B| = \frac{n}{3} + O(1)$, and $|C| = 0$) to Construction 2 (where all three parts have size $\frac{n}{3} + O(1)$). For $c \in [0, 1/3]$, this gives the following:

$$\mathrm{ex}_c(n, F_{3,2}) \geq \left(\frac{1}{3} + 3\left(\frac{1}{3} - c\right)^3\right)\binom{n}{3} + o(n^3).$$

QUESTION 2. *Is this lower bound asymptotically best possible?*

Acknowledgment. We are grateful to an anonymous referee for a careful reading of a long paper.

REFERENCES

[1] R. Baber and J. Talbot, *Hypergraphs do jump*, Combin. Probab. Comput., 20 (2011), pp. 161–171.

[2] R. Baber and J. Talbot, *New Turán densities for* 3*-graphs*, Electron. J. Combin., 19 (2012), pp. 1–21.

[3] T. Bohman, A. Frieze, D. Mubayi, and O. Pikhurko, *Hypergraphs with independent neighborhoods*, Combinatorica, 30 (2010), pp. 277–293.

[4] B. Bollobás, *Three-graphs without two triples whose symmetric difference is contained in a third*, Discrete Math., 8 (1974), pp. 21–24.

[5] A. Czygrinow and B. Nagle, *A note on codegree problems for hypergraphs*, Bull. Inst. Combin. Appl., 32 (2001), pp. 63–69.

[6] 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), pp. 274–276.

[7] L. DeBiasio and T. Jiang, *On the co-degree threshold for the Fano plane*, European J. Combin., 36 (2014), pp. 151–158.

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

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

[10] V. Falgas-Ravry, *On the codegree density of complete* 3*-graphs and related problems*, Electron. J. Combin., 20 (2013), pp. 1–14.

[11] V. Falgas-Ravry, E. Marchant, O. Pikhurko, and E. R. Vaughan, *The Codegree Threshold for* 3*-Graphs with Independent Neighbourhoods*, preprint, arXiv:1307.0075, 2013.

[12] V. Falgas-Ravry, O. Pikhurko, and E.R. Vaughan, *The Codegree Density of* K_4^-, in preparation, 2015.

[13] V. Falgas-Ravry and E.R. Vaughan, *Applications of the semi-definite method to the Turán density problem for* 3*-graphs*, Combin. Probab. Comput., 22 (2013), pp. 21–54.

[14] Z. Füredi, O. Pikhurko, and M. Simonovits, *The Turán density of the hypergraph* $\{abc, ade, bde, cde\}$, Electron. J. Combin., 10 (2003), pp. 1–7.

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

[16] Z. Füredi and M. Simonovits, *Triple systems not containing a Fano configuration*, Combin. Probab. Comput., 14 (2005), pp. 467–484.

[17] R. Glebov, D. Kráľ, and J. Volec, *An application of flag algebras to a problem of Erdős and Sós*, Electron. Notes Discrete Math., 43 (2013), pp. 171–177.

[18] R. Häggkvist and A. Thomason, *Oriented Hamilton cycles in digraphs*, J. Graph Theory, 19 (1995), pp. 471–479.

[19] J. Hirst, *The inducibility of graphs on four vertices*, J. Graph Theory, 75 (2014), pp. 231–243.

[20] P. Keevash, *The Turán problem for projective geometries*, J. Combin. Theory Ser. A, 111 (2005), pp. 289–309.

[21] P. Keevash, *A hypergraph regularity method for generalized Turán problems*, Random Structures Algorithms, 34 (2009), pp. 123–164.

[22] P. Keevash, *Hypergraph Turán problems*, in Surveys in Combinatorics, Cambridge University Press, Cambridge, UK, 2011, pp. 83–140.

[23] P. Keevash and B. Sudakov, *The Turán number of the Fano plane*, Combinatorica, 25 (2005), pp. 561–574.

[24] P. Keevash and Y. Zhao, *Codegree problems for projective geometries*, J. Combin. Theory Ser. B, 97 (2007), pp. 919–928.

[25] A. Lo and K. Markström, *l-degree Turán density*, SIAM J. Discrete Math., 28 (2014), pp. 1214–1225.

[26] E. Marchant, *Graphs with Weighted Colours and Hypergraphs*, Ph.D. thesis, University of Cambridge, Cambridge, UK, 2011.

[27] U. Matthias, *Hypergraphen ohne vollständige r-partite Teilgraphen*, Ph.D. thesis, Heidelberg University, Heidelberg, Germany, 1994.

[28] B. D. McKay and I. M. Wanless, *On the number of Latin squares*, Ann. Comb., 9 (2005), pp. 335–344.

[29] D. Mubayi, *On hypergraphs with every four points spanning at most two triples*, Electron. J.

Combin., 10 (2003), pp. 1–4.
[30] D. Mubayi, *The co-degree density of the Fano plane*, J. Combin. Theory Ser. B, 95 (2005), pp. 333–337.
[31] D. Mubayi and Y. Zhao, *Co-degree density of hypergraphs*, J. Combin. Theory Ser. A, 114 (2007), pp. 1118–1132.
[32] B. Nagle, *Turán-related problems for hypergraphs*, Congr. Numer., 136 (1999), pp. 119–127.
[33] O. Pikhurko and E.R. Vaughan, *Minimum number of k-cliques in graphs with bounded independence number*, Combin. Probab. Comput., 22 (2013), pp. 910–934.
[34] A.A. Razborov, *Flag algebras*, J. Symbolic Logic, 72 (2007), pp. 1239–1282.
[35] A.A. Razborov, *On 3-hypergraphs with forbidden 4-vertex configurations*, SIAM J. Discrete Math., 24 (2010), pp. 946–963.
[36] V. Rödl and M. Schacht, *Generalizations of the Removal Lemma*, Combinatorica, 29 (2009), pp. 467–501.
[37] P. Turán, *On an extremal problem in graph theory*, Mat. Fiz. Lapok, 48 (1941), pp. 436–452.
[38] E.R. Vaughan, *Flagmatic User's Guide (Version 2.0)*, http://flagmatic.org/, 2013.
[39] E.R. Vaughan, *Flagmatic: A Tool for Researchers in Extremal Graph Theory (Version 2.0)*, http://flagmatic.org/, 2013.