Solving Turán's Tetrahedron Problem for the ℓ_2-Norm

József Balogh * Felix Christian Clemen † Bernard Lidický ‡

August 29, 2021

Abstract

Turán's famous tetrahedron problem is to compute the Turán density of the tetrahedron K_4^3. This is equivalent to determining the maximum ℓ_1-norm of the codegree vector of a K_4^3-free n-vertex 3-uniform hypergraph. We will introduce a new way for measuring extremality of hypergraphs and determine asymptotically the extremal function of the tetrahedron in our notion.

The codegree squared sum, $\mathrm{co}_2(G)$, of a 3-uniform hypergraph G is the sum of codegrees squared $d(x,y)^2$ over all pairs of vertices xy, or in other words, the square of the ℓ_2-norm of the codegree vector of the pairs of vertices. Define $\mathrm{exco}_2(n,H)$ to be the maximum $\mathrm{co}_2(G)$ over all H-free n-vertex 3-uniform hypergraphs G. We use flag algebra computations to determine asymptotically the codegree squared extremal number for K_4^3 and K_5^3 and additionally prove stability results. In particular, we prove that the extremal function for K_4^3 in ℓ_2-norm is asymptotically the same as the one obtained from one of the conjectured extremal K_4^3-free hypergraphs for the ℓ_1-norm. Further, we prove several general properties about $\mathrm{exco}_2(n,H)$ including the existence of a scaled limit, blow-up invariance and a supersaturation result.

1 Introduction

For a k-uniform hypergraph H (shortly k-graph), the Turán function (or extremal number) $\mathrm{ex}(n,H)$ is the maximum number of edges in an H-free n-vertex k-uniform hypergraph. The graph case, $k=2$, is reasonably well-understood. The classical Erdős-Stone-Simonovits theorem [15, 17] determines asymptotically the extremal number of graphs with chromatic number at least three. However, for general k, the problem of determining the extremal function is much harder and widely open. Despite enormous efforts, our understanding of Turán functions is still limited. Even the extremal function of the *tetrahedron* K_4^3, the 3-graph on 4 vertices with 4 edges, is unknown. There are exponentially (in the number of vertices) many conjectured extremal examples which is believed to be the root of the difficulty of this problem. Brown [10], Kostochka [35], Fon-der-Flaass [23] and Frohmader [25] constructed families of K_4^3-free 3-graphs which they conjectured to be extremal. For an excellent survey on Turán functions of cliques see [53] by Sidorenko.

*Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA, and Moscow Institute of Physics and Technology, Russian Federation. E-mail: `jobal@illinois.edu`. Research is partially supported by NSF Grant DMS-1764123, Arnold O. Beckman Research Award (UIUC Campus Research Board RB 18132), the Langan Scholar Fund (UIUC), and the Simons Fellowship.

†Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA, E-mail: `fclemen2@illinois.edu`.

‡Iowa State University, Department of Mathematics, Iowa State University, Ames, IA., E-mail: `lidicky@iastate.edu`. Research of this author is partially supported by NSF grant DMS-1855653.

Successively, the upper bound for extremal number of the tetrahedron has been improved by de Caen [13], Giraud (unpublished, see [11]), Chung and Lu [11], and finally Razborov [46] and Baber [2], both making use of Razborov's flag algebra approach [45] (see also Baber and Talbot [3]). Another relevant result towards solving Turán's tetrahedron problem is by Pikhurko [43]. Building on a result by Razborov [46], Pikhurko [43] determined the exact extremal hypergraph when the induced 4-vertex graph with one edge is forbidden in addition to the tetrahedron.

In this paper we study a different notion of extremality and solve the tetrahedron problem asymptotically for this notion. It is interesting that the extremal function for K_4^3 in our notion is asymptotically the same as one of the conjectured one for the Turán density. For an integer n, denote by $[n]$ the set of the first n integers. Given a set A and an integer k, we write $\binom{A}{k}$ for the set of all subsets of A of size k. Let G be an n-vertex k-uniform hypergraph. For $T \subset V(G)$ with $|T| = k-1$ we denote by $d_G(T)$ the *codegree* of T, i.e., the number of edges in G containing T. If the choice of G is obvious, we will drop the index and just write $d(T)$. The *codegree vector* of G is the vector

$$X \in \mathbb{Z}^{\binom{V(G)}{k-1}}, \text{ where } X(v_1, v_2, \dots, v_{k-1}) = d(v_1, v_2, \dots, v_{k-1})$$

for all $\{v_1, v_2, \dots, v_{k-1}\} \in \binom{V(G)}{k-1}$. The ℓ_1-norm of the codegree vector, or to put it in other words, the sum of codegrees, is k times the number of edges. Thus, Turán's problem for k-graphs is equivalent to the question of finding the maximum ℓ_1-norm for the codegree vector of H-free k-graphs. We propose to study this maximum with respect to other norms. A particular interesting case seems to be the ℓ_2-norm of the codegree vector. We will refer to the square of the ℓ_2-norm of the codegree vector as the *codegree squared sum* denoted by $\text{co}_2(G)$,

$$\text{co}_2(G) = \sum_{\substack{T \subset \binom{[n]}{k-1} \\ |T| = k-1}} d_G^2(T).$$

Question 1.1. *Given a k-uniform hypergraph H, what is the maximum codegree squared sum a k-uniform H-free n-vertex hypergraph G can have?*

Many different types of extremality in hypergraphs have been studied. The most related one is the minimum codegree-threshold. For a given k-graph, the minimum codegree-threshold is the largest minimum codegree an n-vertex k-graph can have without containing a copy of H. This problem has not even been solved for H being the tetrahedron. For a collection of results on the minimum codegree-threshold see [18–20, 38–42, 54]. Reiher, Rödl and Schacht [49, 50] introduced new variants of the Turán density, which ask for the maximum density d for which H-free hypergraph with certain quasirandomness properties of density d exists. Roughly speaking, a quasirandomness property is a property the random hypergraph has with probability close to 1. Reiher, Rödl and Schacht [49] determined such a variant of the Turán density of the tetrahedron.

In this paper we solve asymptotically Question 1.1 for the tetrahedron. For a family $\mathcal{F}$ of k-uniform hypergraphs, we define $\text{exco}_2(n, \mathcal{F})$ to be the maximum codegree squared sum a k-uniform n-vertex $\mathcal{F}$-free hypergraph can have, and the *codegree squared density* $\sigma(F)$ to be its scaled limit, i.e.,

$$\text{exco}_2(n, \mathcal{F}) = \max_{\substack{G \text{ is an } n\text{-vertex} \\ \mathcal{F}\text{-free} \\ k\text{-uniform hypergraph}}} \text{co}_2(G) \qquad \text{and} \qquad \sigma(\mathcal{F}) = \lim_{n\to\infty} \frac{\text{exco}_2(n, \mathcal{F})}{\binom{n}{k-1}(n-k+1)^2}. \quad (1)$$

We will observe in Proposition 1.8 that the limit in (1) exists. Denote by K_ℓ^3 the complete 3-uniform hypergraph on ℓ vertices. Our main result is that we determine the codegree squared density asymptotically for K_4^3 and K_5^3, respectively.

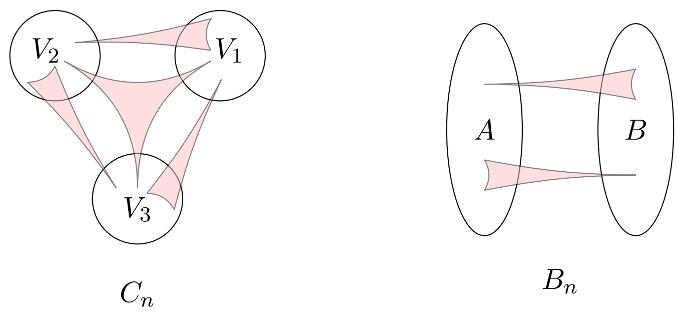


Figure 1: Illustration of C_n and B_n.

Theorem 1.2. *We have*

$$\sigma(K_4^3) = \frac{1}{3} \qquad \textit{and} \qquad \sigma(K_5^3) = \frac{5}{8}.$$

Denote C_n the 3-uniform hypergraph[1] on n vertices with vertex set $V(C_n) = V_1 \cup V_2 \cup V_3$ such that $||V_i| - |V_j|| \leq 1$ for $i \neq j$ and edge set

$$\begin{aligned} E(C_n) = \{abc : a \in V_1, b \in V_2, c \in V_3\} \cup \{abc : a, b \in V_1, c \in V_2\} \\ \cup \{abc : a, b \in V_2, c \in V_3\} \cup \{abc : a, b \in V_3, c \in V_1\}. \end{aligned}$$

Further, denote by B_n the balanced, complete, bipartite 3-uniform hypergraph on n vertices, that is the hypergraph where the vertex set is partitioned into two sets A, B such that $||A|-|B|| \leq 1$ and the edge set is the set of triples intersecting both A and B. See Figure 1 for an illustration of C_n and B_n. The 3-graphs C_n and B_n are one of the asymptotically extremal examples in ℓ_1-norm for K_4^3 and K_5^3 respectively. We conjecture that C_n and B_n are the unique extremal hypergraphs in ℓ_2-norm.

Conjecture 1.3. *There exists n_0 such that for all $n \geq n_0$*

$$\mathrm{exco}_2(n, K_4^3) = \mathrm{co}_2(C_n)$$

and C_n is the unique K_4^3-free n-vertex 3-uniform hypergraph with codegree squared sum equal to $\mathrm{exco}_2(n, K_4^3)$.

Note that Kostochka's [35] result suggests that in the ℓ_1-norm there are exponentially many extremal graphs, C_n is one of them.

Conjecture 1.4. *There exists n_0 such that for all $n \geq n_0$*

$$\mathrm{exco}_2(n, K_5^3) = \mathrm{co}_2(B_n)$$

and B_n is the unique K_5^3-free n-vertex 3-uniform hypergraph with codegree squared sum equal to $\mathrm{exco}_2(n, K_5^3)$.

We believe that existing methods could prove these conjectures, though the potential proofs might be long and technical.

In Section 3.3 we observe that giving upper bounds on $\sigma(H)$ for some 3-graph H is equivalent to giving upper bounds on a certain linear combination of densities of 4-vertex subgraphs in

[1] This hypergraph is often referred to as Turán's construction.

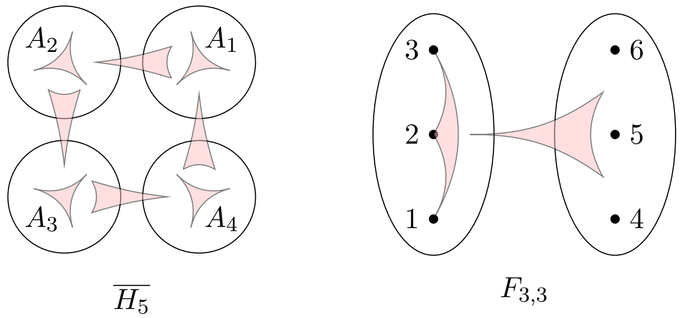


Figure 2: Left: The complement of H_5. Right: A sketch of $F_{3,3}$, which has 6 vertices and edge set $\{123, 145, 146, 156, 245, 246, 256, 345, 346, 356\}$.

large H-free graphs, see (2). By now it is a standard technique in the field to use the computer-assisted method of flag algebras to prove such bounds. If one gets an asymptotically tight upper bound from a flag algebra computation, it is typically the case that there is an essentially unique stable extremal example and that one can extract a stability result from the flag algebra proof. This also happens for K_4^3 and K_5^3. For $\varepsilon > 0$, we say a given n-vertex 3-graph H is *ε-near* to an n-vertex 3-graph G if there exists a bijection $\phi : V(G) \to V(H)$ such that the number of 3-sets $\{x, y, z\}$ satisfying $xyz \in E(G), \phi(x)\phi(y)\phi(z) \notin E(H)$ or $xyz \notin E(G), \phi(x)\phi(y)\phi(z) \in E(H)$ is at most $\varepsilon|V(H)|^3$.

Theorem 1.5. *For every $\varepsilon > 0$ there exists $\delta > 0$ and n_0 such that for every $n > n_0$, if G is a K_4^3-free 3-uniform hypergraph on n vertices with*

$$\mathrm{co}_2(G) \geq \left(\frac{1}{3} - \delta\right)\frac{n^4}{2},$$

then G is ε-near to C_n.

Theorem 1.6. *For every $\varepsilon > 0$ there exists $\delta > 0$ and n_0 such that for every $n > n_0$, if G is a K_5^3-free 3-uniform hypergraph on n vertices with*

$$\mathrm{co}_2(G) \geq \left(\frac{5}{8} - \delta\right)\frac{n^4}{2},$$

then G is ε-near to B_n.

There is a K_5^3-free 3-graph [52] with the same edge density as B_n, namely H_5. The vertex set of H_5 is divided into 4 parts A_1, A_2, A_3, A_4 with $||A_j| - |A_i|| \leq 1$ for all $1 \leq i \leq j \leq 4$ and say a triple e is not an edge of H_5 iff there is some j $(1 \leq j \leq 4)$ such that $|e \cap A_j| \geq 2$ and $|e \cap A_j| + |e \cap A_{j+1}| = 3$, where $A_5 = A_1$, see Figure 2 for an illustration of the complement of H_5. While H_5 is conjectured to be one of the asymptotically extremal examples in ℓ_1-norm, it is not an extremal example in ℓ_2-norm, because B_n has an asymptotically higher codegree squared sum.

Besides giving asymptotic result for cliques, we prove an exact result for $F_{3,3}$. Denote by $F_{3,3}$ the 3-graph on 6 vertices with edge set $\{123, 145, 146, 156, 245, 246, 256, 345, 346, 356\}$, see Figure 2. We prove that the codegree squared extremal example of $F_{3,3}$ is the balanced, complete, bipartite hypergraph B_n. Keevash and Mubayi [33] and independently Goldwasser and Hansen [27] proved that B_n is also extremal for ℓ_1-norm.

Theorem 1.7. *There exists n_0 such that for all $n \geq n_0$*

$$\text{exco}_2(n, F_{3,3}) = \text{co}_2(B_n).$$

Furthermore, B_n is the unique $F_{3,3}$-free 3-uniform hypergraph G on n vertices satisfying

$$\text{co}_2(G) = \text{exco}_2(n, F_{3,3}).$$

We also prove some general results on σ. First, we prove that the limit in (1) exists.

Proposition 1.8. *Let $\mathcal{F}$ be a family of k-graphs. Then, $\frac{\text{exco}_2(n,\mathcal{F})}{\binom{n}{k-1}(n-k+1)^2}$ is non-increasing as n increases. In particular, it tends to a limit $\sigma(\mathcal{F})$ as $n \to \infty$.*

A classical result in extremal combinatorics is the supersaturation phenomenon, discovered by Erdős and Simonovits [16]. For hypergraphs it states, that when the edge density of a hypergraph H exceeds the Turán density of a different hypergraph G, then H contains many copies of G. Proposition 1.9 shows that the same phenomenon holds for σ.

Proposition 1.9. *Let F be a k-graph on f vertices. For every $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon, f) > 0$ and n_0 such that every n-vertex k-uniform hypergraph G with $n > n_0$ and $\text{co}_2(G) > (\sigma(F) + \varepsilon)\binom{n}{k-1}n^2$ contains at least $\delta\binom{n}{f}$ copies of F.*

Supersaturation has been used to show that blowing up a k-graph does not change its Turán density [16]. We will use our Supersaturation result, Proposition 1.9, to show the same conclusion holds for σ: Blowing up a k-graph does also not change the codegree squared density. For a k-graph H and $t \in \mathbb{N}$, the *blow-up* $H(t)$ of H is defined by replacing each vertex $x \in V(H)$ by t vertices $x^1, \dots, x^t$ and each edge $x_1 \cdots x_k \in E(H)$ by the t^k edges $x_1^{a_1} \cdots x_k^{a_k}$ with $1 \leq a_1, \dots, a_k \leq t$.

Corollary 1.10. *Let H be a k-uniform hypergraph and $t \in \mathbb{N}$. Then,*

$$\sigma(H) = \sigma(H(t)).$$

Similarly to the Turán density [14], the codegree squared density has a jump at 0. Note that this phenomenon is not happening for the minimum codegree threshold [38].

Proposition 1.11. *Let H be a k-uniform hypergraph. Then*

(i) $(\pi(H))^2 \leq \sigma(H) \leq \pi(H)$,

(ii) $\sigma(H) = 0$ *or* $\sigma(H) \geq \frac{(k-1)!}{k^k}$.

Our paper is organised as follows. In Section 2 we calculate the extremal ℓ_2-norm for a classical, but easy, example in ℓ_1-norm as a warm-up. Next, in Section 3 we introduce terminology and give an overview of the tools we will be using. In Section 4 we present our general results on maximal codegree squared sums. Section 5 is dedicated to proving our main results on cliques, meaning proving Theorems 1.5 and 1.6. In Section 6 we present the proof of our exact result, Theorem 1.7.

In a follow-up paper [4] we systematically study the codegree squared densities of several hypergraphs. Also we discuss further open problems there.

2 A Toy Example: Forbidding F_4 and F_5

In this section we will provide an example of how a classical Turán-type result on the ℓ_1-norm can imply a result for the codegree squared density, ℓ_2-norm. Denote by F_4 the 4-vertex 3-graph[2] with edge set $\{123, 124, 234\}$ and F_5 the 5-vertex 3-graph with edge set $\{123, 124, 345\}$,

[2]This hypergraph is also knows as K_4^{3-}.

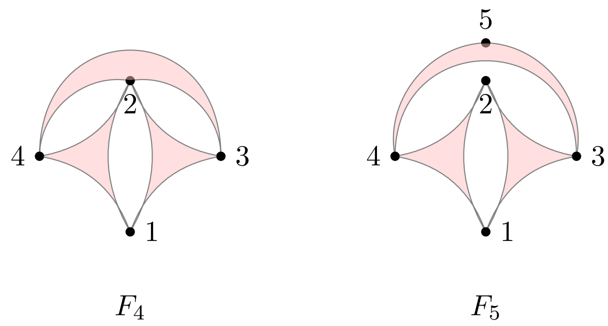


Figure 3: Hypergraphs F_4 and F_5.

see Figure 3. The 3-graphs which are F_4- and F_5-free are called *cancellative hypergraphs*. Denote by S_n the complete balanced 3-partite 3-graph on n vertices. This is the 3-graph with vertex partition $A \cup B \cup C$ with part sizes $|A| = \lfloor n/3 \rfloor$, $|B| = \lfloor (n+1)/3 \rfloor$ and $|C| = \lfloor (n+2)/3 \rfloor$, where triples abc are edges iff a, b and c are each from a different class. Bollobás [8] proved that the n-vertex cancellative hypergraph with the most edges is S_n. Using his result and a double counting argument we show that S_n is also the largest cancellative hypergraph in the ℓ_2-norm.

Theorem 2.1. *We have*

$$\text{exco}_2(n, \{F_4, F_5\}) = \text{co}_2(S_n),$$

and therefore also

$$\sigma(\{F_4, F_5\}) = \frac{2}{27}.$$

The unique extremal hypergraph is S_n*.*

Proof. Let G be an F_4- and F_5-free hypergraph with n vertices. For an edge $e = xyz \in E(G)$, we define its weight $w(e) = d(x,y) + d(x,z) + d(y,z)$. Then, $w(e) \leq n$; otherwise G contains an F_4. Bollobás [8] proved that $|E(G)| \leq |E(S_n)|$ with equality iff $G = S_n$. This allows us to conclude

$$\text{co}_2(G) = \sum_{xy \in \binom{[n]}{2}} d(x,y)^2 = \sum_{e \in E(G)} w(e) \leq n|E(G)| \leq n|E(S_n)| = \text{co}_2(S_n). \qquad \blacksquare$$

Frankl and Füredi [24] proved that for just F_5-free 3-graphs, S_n is also the extremal example in ℓ_1-norm when $n \geq 3000$. In a follow-up paper [4] we prove that for F_5-free 3-graphs, S_n is also the extremal example in the ℓ_2-norm provided n is sufficiently large. However, this requires more work than the proof of Theorem 2.1 and it is not derived by just applying the corresponding Turán result.

3 Preliminaries

3.1 Terminology and notation

Let H be a 3-uniform hypergraph, $x \in V(H)$ and $A, B \subseteq V(H)$ be disjoint sets.

1. $L(x)$ denotes the link graph of x, i.e., the graph on $V(H) \setminus \{x\}$ with $ab \in E(L(x))$ iff $abx \in E(H)$.

2. $L_A(x) = L(x)[A]$ denotes the induced link graph on A.
3. $L_{A,B}(x)$ denotes the subgraph of the link graph of x containing only edges between A and B. This means $V(L_{A,B}(x)) = V(H) \setminus \{x\}$ and $ab \in E(L_{A,B}(x))$ iff $a \in A, b \in B$ and $abx \in E(H)$.
4. $L^c_{A,B}(x)$ denotes the subgraph of the link graph of x containing only non-edges between A and B. This means $V(L_{A,B}(x)) = V(H) \setminus \{x\}$ and $ab \in E(L^c_{A,B}(x))$ iff $a \in A, b \in B$ and $abx \notin E(H)$.
5. $e(A, B)$ denotes the number of cross edges between A and B, this means $e(A, B) := |\{xyz \in E(H) : x, y \in A, z \in B\}| + |\{xyz \in E(H) : x, y \in B, z \in A\}|$.
6. $e^c(A, B)$ denotes the number of missing cross edges between A and B, this means $e^c(A, B) := \binom{|A|}{2}|B| + \binom{|B|}{2}|A| - e(A, B)$.
7. For an edge $e = xyz \in E(H)$, we define its *weight* as

$$w_H(e) = d(x, y) + d(x, z) + d(y, z).$$

3.2 Tool 1: Induced hypergraph removal Lemma

We will use the induced hypergraph removal lemma of Rödl and Schacht [51].

Definition 3.1. Let $\mathcal{F}, \mathcal{P}$ be families of k-graphs.

- $\text{Forb}_{ind}(\mathcal{F})$ denotes the family of all k-graphs H which contain no induced copy of any member of $\mathcal{F}$.
- For a constant $\mu \geq 0$ we say a given k-graph H is *μ-far* from $\mathcal{P}$ if every k-graph G on the same vertex set $V(H)$ with $|G \triangle H| \leq \mu |V(H)|^k$ satisfies $G \notin \mathcal{P}$, where $G \triangle H$ denotes the symmetric difference of the edge sets of G and H. Otherwise we call H *μ-near* to $\mathcal{P}$.

Theorem 3.2 (Rödl, Schacht [51])**.** *For every (possibly infinite) family $\mathcal{F}$ of k-graphs and every $\mu > 0$ there exist constants $c > 0, C > 0$, and $n_0 \in \mathbb{N}$ such that the following holds. Suppose H is a k-graph on $n \geq n_0$ vertices. If for every $\ell = 1, \dots, C$ and every $F \in \mathcal{F}$ on ℓ vertices, H contains at most cn^ℓ induced copies of F, then H is μ-near to $\text{Forb}_{ind}(\mathcal{F})$.*

3.3 Tool 2: Flag Algebras

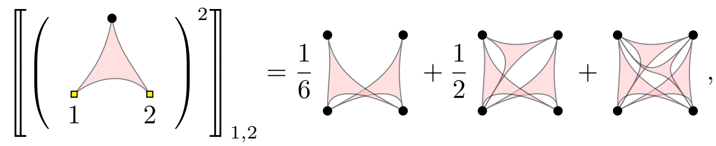
In this section we give an insight on how we apply Razborov's flag algebra machinery [45] for calculating the codegree squared density. The main power comes from the possibility of formulating a problem as a semidefinite program and using a computer to solve it.

The method can be applied in various settings such as graphs [28, 44], hypergraphs [3, 19], oriented graphs [29, 37], edge-coloured graphs [5, 12], permutations [6, 55], discrete geometry [7, 36], or phylogenetic trees [1]. For a detailed explanation of the flag algebra method in the setting of 3-uniform hypergraphs see [22]. Further, we recommend looking at the survey [47] and the expository note [48], both by Razborov. Here, we will focus on the problem formulation rather than a formal explanation of the general method.

Let F be a fixed 3-graph. Let $\mathcal{F}$ denote the set of all F-free 3-graphs up to isomorphism. Denote by $\mathcal{F}_\ell$ all 3-graphs in $\mathcal{F}$ on ℓ vertices. For two 3-graphs F_1 and F_2, denote by $P(F_1, F_2)$ the probability that $|V(F_1)|$ vertices chosen uniformly at random from $V(F_2)$ induce a copy of F_1. A sequence of 3-graphs $(G_n)_{n\geq 1}$ of increasing orders is *convergent*, if $\lim_{n\to\infty} P(H, G_n)$ exists for every $H \in \mathcal{F}$. Notice that if this limit exists, it is in $[0, 1]$.

For readers familiar with flag algebras and its usual notation, for a convergent sequence $(G_n)_{n\geq 1}$ of n-vertex 3-graphs G_n, we get

$$\lim_{n\to\infty} \frac{\mathrm{co}_2(G_n)}{\binom{n}{2}(n-2)^2} =$$



(2)

where $[\![\cdot]\!]$ denotes the averaging operator and the terms on the right are interpreted as

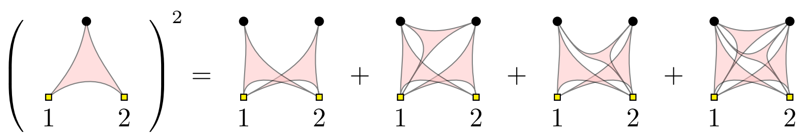
$$\lim_{n\to\infty} \frac{1}{6}P(K_4^{3=}, G_n) + \frac{1}{2}P(K_4^{3-}, G_n) + P(K_4^3, G_n),$$

where $K_4^{3=}$ is the 3-graph with 4 vertices and 2 edges and K_4^{3-} the 3-graph with 4 vertices and 3 edges, also known as F_4. It is a routine application of flag algebras to find an upper bound on the right-hand side of (2).

For readers less familiar with flag algebras, the following paragraphs give a slightly less formal explanation of the problem formulation. Let G be a 3-graph. Let θ be an injective function $\{1,2\} \to V(G)$. In other words, θ labels two distinct vertices in G. We call the pair (G,θ) a *labelled 3-graph* although only two vertices in G are labelled by θ.

Let (H,θ') and (G,θ) be two labelled 3-graphs. Let X be a subset of $V(G) \setminus \mathrm{Im}\,\theta$ of size $|V(H)|-2$ chosen uniformly at random. By $P((H,\theta'),(G,\theta))$ we denote the probability that the labelled subgraph of G induced by X and the two labelled vertices, i.e., $(G[X \cup \mathrm{Im}\,\theta],\theta)$, is isomorphic to (H,θ'), where the isomorphism maps $\theta(i)$ to $\theta'(i)$ for $i \in \{1,2\}$.

Let E be a labelled 3-graph consisting of three vertices, two of them labelled, and one edge containing all three vertices. Notice that $P(E,(G,\theta))(n-2)$ is the codegree of $\theta(1)$ and $\theta(2)$ in a 3-graph G. The square of the codegree of $\theta(1)$ and $\theta(2)$ is $(P(E,(G,\theta))(n-2))^2$. One of the tricks in flag algebras is that calculating $P(E,(G,\theta))^2$ in G of order n can be done with error $O(1/n)$ by selecting two distinct vertices in addition to $\theta(1)$ and $\theta(2)$ and examining subgraphs on four vertices instead. In our case, it looks like the following, where $P(H,(G,\theta))$ is depicted simply as H.



$$+\,o(1) \qquad (3)$$

The next step is to sum over all possible choices for θ, there are $n(n-1)$ of them, and divide by 2 since the codegree squared sum is over unordered pairs of vertices, unlike θ. When summing over all possible θ, one could look at all subsets of vertices of size 4 of G and see what the probability is that randomly labelling two vertices among these four by θ gives one of the labelled 3-graphs from the right hand side of (3). This gives the coefficients on the right-hand side of (2).

We use flag algebras to prove Lemmas 5.1, 6.1, and 5.3. The calculations are computer assisted. We use CSDP [9] for finding numerical solutions of semidefinite programs and SageMath [56] for rounding the numerical solutions to exact ones. The files needed to perform the corresponding calculations are available at http://lidicky.name/pub/co2/.

4 General results: Proofs of Propositions 1.8, 1.9 and 1.10

4.1 The limit exists

Proof of Proposition 1.8. Let $n \geq k$ be a positive integer and let G be an $\mathcal{F}$-free k-graph on vertex set $[n]$ satisfying $\mathrm{co}_2(G) = \mathrm{exco}_2(n,\mathcal{F})$. Take S to be a randomly chosen $(n-1)$-subset of $V(G)$. Now, we calculate the expectation of $\mathrm{co}_2(G[S])$,

$$\begin{aligned}
\mathbb{E}[\mathrm{co}_2(G[S])] &= \sum_{T\in\binom{[n]}{k-1}} \mathbb{E}[\mathbf{1}_{\{T\subset S\}} d^2_{G[S]}(T)] = \sum_{T\in\binom{[n]}{k-1}} \mathbb{P}(T\subset S)\mathbb{E}[d^2_{G[S]}(T)|T\subset S] \\
&= \sum_{T\in\binom{[n]}{k-1}} \frac{\binom{n-1}{k-1}}{\binom{n}{k-1}} \mathbb{E}[d^2_{G[S]}(T)|T\subset S] \geq \sum_{T\in\binom{[n]}{k-1}} \frac{\binom{n-1}{k-1}}{\binom{n}{k-1}} \mathbb{E}[d_{G[S]}(T)|T\subset S]^2 \\
&= \sum_{T\in\binom{[n]}{k-1}} \frac{\binom{n-1}{k-1}}{\binom{n}{k-1}} \left(d_G(T)\frac{n-k}{n-k+1}\right)^2 = \frac{\binom{n-1}{k-1}}{\binom{n}{k-1}} \left(\frac{n-k}{n-k+1}\right)^2 \mathrm{co}_2(G).
\end{aligned}$$

We used that $d_{G[S]}(T)$ conditioned on $T\subset S$ has hypergeometric distribution. By averaging, we conclude that there exists an $(n-1)$-vertex subset $S'\subset V(G)$ with $\mathrm{co}_2(G[S'])\geq \mathbb{E}[\mathrm{co}_2(G[S])]$. Thus, we conclude that $G[S']$ is an $(n-1)$-vertex k-graph satisfying

$$\mathrm{co}_2(G[S']) \geq \frac{\binom{n-1}{k-1}}{\binom{n}{k-1}} \left(\frac{n-k}{n-k+1}\right)^2 \mathrm{co}_2(G).$$

Therefore, since $G[S']$ is $\mathcal{F}$-free,

$$\frac{\mathrm{exco}_2(n-1,\mathcal{F})}{\binom{n-1}{k-1}(n-k)^2} \geq \frac{\mathrm{co}_2(G[S'])}{\binom{n-1}{k-1}(n-k)^2} \geq \frac{\mathrm{co}_2(G)}{\binom{n}{k-1}(n-k+1)^2} = \frac{\mathrm{exco}_2(n,\mathcal{F})}{\binom{n}{k-1}(n-k+1)^2}. \qquad \blacksquare$$

4.2 Supersaturation

In this section we prove Proposition 1.9. We will make use of the following tail bound on the hypergeometric distribution.

Lemma 4.1 (e.g. [30] p.29)**.** *Let $\beta,\lambda>0$ with $\beta+\lambda<1$. Suppose that $X\subseteq[n]$ and $|X|\geq(\beta+\lambda)n$. Then*

$$\left|\left\{S\in\binom{[n]}{m} : |S\cap X|\leq\beta m\right\}\right| \leq \binom{n}{m}e^{-\frac{\lambda^2 m}{3(\beta+\lambda)}} \leq \binom{n}{m}e^{-\lambda^2 m/3}.$$

Mubayi and Zhao [41] used Lemma 4.1 to prove a supersaturation result for the minimum codegree threshold. We adapt their proof to our setting.

Lemma 4.2. *Let $\alpha>0$, $\varepsilon>0$ and $k\geq3$. Then there exists m_0 such that the following holds. If $n\geq m\geq m_0$ and G is a k-graph on $[n]$ with $\mathrm{co}_2(G)\geq(\alpha+\varepsilon)\binom{n}{k-1}(n-k+1)^2$, then the number of m-sets S satisfying $\mathrm{co}_2(G[S])>\alpha\binom{m}{k-1}(m-k+1)^2$ is at least $\frac{\varepsilon}{4}\binom{n}{m}$.*

Proof. Given a $(k-1)$-element set $T\subset[n]$, we call an m-set S with $T\subset S\subset[n]$ *bad for T* if $|d(T)\cap S|\leq\left(\frac{d(T)}{n-k+1}-\frac{\varepsilon}{6}\right)(m-k+1)$. An m-set is *bad* if it is bad for some T. Otherwise, it is *good*. We will show that there are few bad sets. Denote by Φ the number of bad m-sets, and let Φ_T be the number of m-sets that are bad for T. Then, by applying Lemma 4.1 with $\beta=\frac{d(T)}{n-k+1}-\frac{\varepsilon}{6}$ and $\lambda=\varepsilon/7$, we get

$$\begin{aligned}
\Phi &\leq \sum_{T\in\binom{[n]}{k-1}} \Phi_T = \sum_{T\in\binom{[n]}{k-1}} \left|\left\{S'\in\binom{[n]\setminus T}{m-k+1} : |d(T)\cap S'|\leq\left(\frac{d(T)}{n-k+1}-\frac{\varepsilon}{6}\right)(m-k+1)\right\}\right| \\
&\leq \sum_{T\in\binom{[n]}{k-1}} \binom{n-k+1}{m-k+1}\exp\left(-\frac{\varepsilon^2(m-k+1)}{147}\right) \leq \binom{n}{k-1}\binom{n-k+1}{m-k+1}\exp\left(-\frac{\varepsilon^2(m-k+1)}{147}\right) \\
&= \binom{n}{m}\binom{m}{k-1}\exp\left(-\frac{\varepsilon^2(m-k+1)}{147}\right) \leq \frac{\varepsilon}{4}\binom{n}{m},
\end{aligned}$$

where the last inequality holds for m large enough. So the number of bad m-sets is at most $\frac{\varepsilon}{4}\binom{n}{m}$. Now let $\ell\binom{n}{m}$ be the number of m-sets S satisfying

$$\sum_{T\in\binom{S}{k-1}} d_G^2(T) \geq \left(\alpha+\frac{\varepsilon}{2}\right)\binom{m}{k-1}(n-k+1)^2. \tag{4}$$

On one side

$$\sum_{|S|=m}\sum_{T\in\binom{S}{k-1}} d_G^2(T) = \binom{n-k+1}{m-k+1}\mathrm{co}_2(G) = \binom{n-k+1}{m-k+1}\binom{n}{k-1}(n-k+1)^2(\alpha+\varepsilon).$$

On the other side,

$$\begin{aligned}\sum_{|S|=m}\sum_{T\in\binom{S}{k-1}} d_G^2(T) &\leq (\alpha+\frac{\varepsilon}{2})\binom{m}{k-1}(n-k+1)^2\binom{n}{m} + \ell\binom{m}{k-1}(n-k+1)^2\binom{n}{m}\\ &= (\alpha+\frac{\varepsilon}{2}+\ell)\binom{m}{k-1}(n-k+1)^2\binom{n}{m}.\end{aligned}$$

By this double counting argument, we conclude $\ell \geq \varepsilon/2$. Since the number of bad m-sets is at most $\frac{\varepsilon}{4}\binom{n}{m}$, there are at least $\frac{\varepsilon}{4}\binom{n}{m}$ good m-sets satisfying (4). All of these m-sets satisfy

$$\begin{aligned}\mathrm{co}_2(G[S]) = \sum_{T\in\binom{S}{k-1}} d_{G[S]}^2(T) &\geq \sum_{T\in\binom{S}{k-1}}\left(\left(\frac{d_G(T)}{n-k+1}-\frac{\varepsilon}{6}\right)(m-k+1)\right)^2\\ &= \frac{(m-k+1)^2}{(n-k+1)^2}\sum_{T\in\binom{S}{k-1}}\left(d_G(T)-\frac{\varepsilon}{6}(n-k+1)\right)^2\\ &\geq \frac{(m-k+1)^2}{(n-k+1)^2}\sum_{T\in\binom{S}{k-1}}\left(d_G^2(T)-\frac{\varepsilon}{3}(n-k+1)^2\right)\\ &\geq \frac{(m-k+1)^2}{(n-k+1)^2}\left(\left(\alpha+\frac{\varepsilon}{2}\right)\binom{m}{k-1}(n-k+1)^2-\binom{m}{k-1}\frac{\varepsilon}{3}(n-k+1)^2\right)\\ &> \alpha\binom{m}{k-1}(m-k+1)^2,\end{aligned}$$

proving the statement of this lemma. ■

Proof of Proposition 1.9. This proof follows Erdős and Simonovits' proof [16] of the supersaturation result for the Turán density.

Let F be a k-graph on f vertices, $\varepsilon > 0$ and G be an n-vertex k-graph satisfying $\mathrm{co}_2(G) > (\sigma(F)+\varepsilon)\binom{n}{k-1}n^2$ for n large enough. By Lemma 4.2, there exists an m_0 such that for $m \geq m_0$ the number of m-sets S satisfying $\mathrm{co}_2(G[S]) > (\sigma(F)+\varepsilon/2)\binom{m}{k-1}(m-k+1)^2$ is at least $\frac{\varepsilon}{8}\binom{n}{m}$. There exists some fixed $m_1 \geq m_0$ such that $\mathrm{exco}_2(m_1, F) \leq (\sigma(F)+\varepsilon/2)\binom{m_1}{k-1}(m_1-k+1)^2$. Thus, there are at least $\frac{\varepsilon}{8}\binom{n}{m_1}$ m_1-sets S such that $G[S]$ contains F. Each copy of F may be counted at most $\binom{n-f}{m_1-f}$ times. Therefore, the number of copies for F is at least

$$\frac{\frac{\varepsilon}{8}\binom{n}{m_1}}{\binom{n-f}{m_1-f}} = \delta\binom{n}{f},$$

for $\delta = \frac{\varepsilon}{8\binom{m_1}{f}}$. ■

4.3 Proof of Corollary 1.10 and Proposition 1.11

Now we use a standard argument to show that blowing-up a k-graph does not change the codegree squared density. We will follow the proof of the analogous Turán result given in [31].

Proof of Corollary 1.10. Since $H \subset H(t)$, $\text{exco}_2(n, H(t)) \leq \text{exco}_2(n, H)$ holds trivially. Thus, $\sigma(H(t)) \leq \sigma(H)$.
For the other direction, let $\varepsilon > 0$ and G be an n-vertex k-uniform hypergraph satisfying $\text{co}_2(G)/(\binom{n}{k-1}(n-k+1)^2) > \sigma(H) + \varepsilon$. Then, by Proposition 1.9, G contains at least $\delta\binom{n}{v(H)}$ copies of H for $\delta = \delta(\varepsilon, k) > 0$. We create an auxiliary $v(H)$-graph F on the vertex set $V(G)$. A $v(H)$-set $A \subset V(G)$ is an edge in F iff $G[A]$ contains a copy of H. The auxiliary hypergraph F has density at least $\delta/v(H)!$. Thus, as it is well-known [14], for any $t' > 0$ as long as n is large enough, F contains a copy of $K_{v(H)}^{v(H)}(t')$, the complete $v(H)$-partite $v(H)$-graph with t' vertices in each part.
We choose t' large enough such that the following is true. We colour each edge of $K_{v(H)}^{v(H)}(t')$ by one of $v(H)!$ colours, depending on which of the $v(H)!$ orders the vertices of H are mapped to in the corresponding copy of H in G. By a classical result in Ramsey theory (for a density version see [14]), there is a monochromatic copy of $K_{v(H)}^{v(H)}(t)$, which contains a copy of $H(t)$ in G. We conclude $\sigma(H(t)) \leq \sigma(H) + \varepsilon$ for all $\varepsilon > 0$. ■

Proof of Proposition 1.11. Let H be a k-graph. For any k-graph G, we have by the Cauchy-Schwarz inequality

$$\text{co}_2(G) = \sum_{T \in \binom{[n]}{k-1}} d_G(T)^2 \geq \frac{\left(\sum_{T \in \binom{[n]}{k-1}} d_G(T)\right)^2}{\binom{n}{k-1}} = \frac{(k|E(G)|)^2}{\binom{n}{k-1}}.$$

After scaling this implies $\sigma(H) \geq \pi(H)^2$. For the upper bound we have

$$\text{co}_2(G) = \sum_{T \in \binom{[n]}{k-1}} d_G(T)^2 = \sum_{e \in E(G)} w_G(e) \leq kn|E(G)|,$$

where $w_G(e) := \sum_{T \in \binom{e}{k-1}} d_G(T)$. After scaling this implies $\sigma(H) \leq \pi(H)$, completing the proof of part (i). Erdős [14] proved that the Turán density of a k-partite k-graph is 0. In this case, the codegree squared density is also 0 by part (i).

If H is not k-partite then the complete k-partite k-graph does not contain H and provides a construction for lower bounds. It gives that the Turán density of H is at least $k!/k^k$ and $\sigma(H) \geq (k-1)/k^k$. ■

5 Cliques

In this section we will prove Theorems 1.5 and 1.6.

5.1 Proof of Theorem 1.5

Flag algebras give us the following results for K_4^3.

Lemma 5.1. *For all $\varepsilon > 0$ there exists $\delta > 0$ and n_0 such that for all $n \geq n_0$: if G is a K_4^3-free 3-uniform graph on n vertices with $\text{co}_2(G) \geq (1-\delta)\frac{1}{3}n^4/2$, then the densities of all 3-graphs on 4, 5 and 6 vertices in G that are not contained in C_n are at most ε. Additionally,*

$$\sigma(K_4^3) = \frac{1}{3}.$$

The flag algebra calculation proving Lemma 5.1 is computer assisted and not practical to fit in the paper. The calculation is available at http://lidicky.name/pub/co2/. For proving Theorem 1.5 we will make use of the following stability result due to Pikhurko [43].

Theorem 5.2 (Pikhurko [43])**.** *For every $\varepsilon > 0$ there exists $\delta > 0$ and n_0 such that for every $n > n_0$, if G is a K_4^3-free 3-uniform hypergraph on n vertices not spanning exactly one edge on four vertices and with*

$$e(G) \geq \left(\frac{5}{9} - \delta\right)\binom{n}{3},$$

then G is ε-near to C_n.

Proof of Theorem 1.5. Let $\varepsilon > 0$ be fixed. We choose n_0 sufficiently large for the following proof to work. We will choose constants

$$1 \gg \varepsilon \gg \delta_3 \gg \delta_2 \gg \delta_1 \gg \delta \gg 0$$

in order from left to right where each constant is a sufficiently small positive number depending only on the previous ones. Let G be a K_4^3-free 3-uniform hypergraph on $n \geq n_0$ vertices with

$$\mathrm{co}_2(G) \geq \left(\frac{1}{3} - \delta\right)\frac{n^4}{2}.$$

By applying Lemma 5.1, we get that the density of the 4-vertex 3-graph with exactly one edge in G is at most δ_1. Now, we apply the induced hypergraph removal lemma, Theorem 3.2, to obtain G' where G' is δ_2-near to G, and G' is K_4^3-free and does not induce exactly one edge on four vertices. We have

$$\mathrm{co}_2(G') \geq \mathrm{co}_2(G) - 6\delta_2 n^4 \geq \left(\frac{1}{3} - \delta\right)\frac{n^4}{2} - 6\delta_2 n^4 \geq (1 - 37\delta_2)\frac{1}{6}n^4,$$

where the first inequality holds because when one edge is removed from a 3-uniform hypergraph, then the codegree squared sum can go down by at most $6n$. By a result of Falgas-Ravry and Vaughan [21, Theorem 4], $P(K_4^{3-}, G') \leq 16/27 + o(1)$. Let $x \in [0,1]$ such that $P(K_4^{3-}, G') = 16/27(1-x) + o(1)$. By (2) and the fact that G' is K_4^3-free, we have

$$\frac{1}{3}(1 - 37\delta_2) \leq \frac{\mathrm{co}_2(G')}{\binom{n}{2}(n-2)^2} = \frac{1}{6}P(K_4^{3=}, G') + \frac{1}{2}P(K_4^{3-}, G') \leq \frac{1}{6}P(K_4^{3=}, G') + \frac{8}{27}(1-x) + o(1).$$

Thus,

$$P(K_4^{3=}, G') \geq \frac{2 + 16x}{9} - 75\delta_2.$$

Since G' does not contain a 4-set spanning exactly 1 or 4 edges, a result of Razborov [46] says

$$\frac{|E(G')|}{\binom{n}{3}} \leq \frac{5}{9} + o(1). \tag{5}$$

Since

$$\frac{|E(G')|}{\binom{n}{3}} = \frac{1}{2}P(K_4^{3=}, G') + \frac{3}{4}P(K_4^{3-}, G') + o(1) \geq \frac{5 + 4x}{9} - 38\delta_2,$$

this implies that $x \leq 100\delta_2$. Thus, by Pikhurko's stability theorem (Theorem 5.2), G' is δ_3-near to C_n. Since G' is δ_2-near to G, we conclude that G is ε-near to C_n. ∎

5.2 Proof of Theorem 1.6

Flag algebras give us the following for K_5^3.

Lemma 5.3. *For all $\varepsilon > 0$ there exists $\delta > 0$ and n_0 such that for all $n \geq n_0$: if G is a K_5^3-free 3-uniform graph on n vertices with $\mathrm{co}_2(G) \geq (1-\delta)\frac{5}{8}n^4/2$, then the densities of all 3-graphs on 4, 5 and 6 vertices in G that are not contained in B_n are at most ε. In particular,*

$$\sigma(K_5^3) = \frac{5}{8}.$$

Again, the flag algebra calculation proving Lemma 5.3 is computer assisted and available at http://lidicky.name/pub/co2/. We use this result to prove Theorem 1.6.

Proof of Theorem 1.6. Let $\varepsilon > 0$. During the proof we will use the following constants:

$$1 \gg \varepsilon \gg \delta_2 \gg \delta_1 \gg \delta \gg 0.$$

The constants are chosen in this order and each constant is a sufficiently small positive number depending only on the previous ones. Apply Lemma 5.3 and get $\delta = \delta(\delta_1) > 0$ such that for all n large enough: If G is an K_5^3-free 3-uniform graph on n vertices with $\mathrm{co}_2(G) \geq (1-\delta)\frac{5}{8}n^4/2$, then the densities of all 3-graphs on 4, 5 and 6 vertices in G that are not contained in B_n are at most δ_1.
Now, apply the induced hypergraph removal lemma Theorem 3.2 to obtain G' where G' is δ_2-near to G, and G' contains only those induced subgraphs on 4, 5 or 6 vertices which appear as induced subgraphs in B_n. Note that

$$\mathrm{co}_2(G') \geq \mathrm{co}_2(G') - 6\delta_2 n^4 \geq (1-\delta)\frac{5}{8}\frac{n^4}{2} - 6\delta_2 n^4 \geq (1-20\delta_2)\frac{5}{8}\frac{n^4}{2},$$

because when one edge is removed the codegree squared sum can go down by at most $6n$. Next we show that G' has to have the same structure as B_n. We say that a 3-graph H is 2-*colourable*, if there is a partition of the vertex set $V(H) = V_1 \cup V_2$ such that V_1 and V_2 are independent sets in H.

Claim 5.4. *G' is 2-colourable.*

Proof. Take an arbitrary non-edge abc in G'. For $0 \leq i \leq 4$, define A_i to be the set of vertices $v \in V(G) \setminus \{a,b,c\}$ such that G' induces i edges on $\{a,b,c,v\}$. Then, $A_1 = A_2 = \emptyset$ because on 4 vertices there are either 0, 3 or 4 edges in B_n, hence in G' as well. Further $A_4 = \emptyset$, because abc is a non-edge. Clearly, A_0 is an independent set, because if there is an edge $v_1v_2v_3$ in $G'[A_0]$, then the induced graph of G' on $\{a,b,c,v_1,v_2,v_3\}$ spans a forbidden subgraph, i.e., a hypergraph which is not an induced subhypergraph of B_n. Similarly, A_3 is an independent set else G' contains a copy of $F_{3,3}$, which is not an induced subhypergraph of B_n. Let $A' = A_0 \cup \{a,b,c\}$. Then $V(G') = A_3 \cup A'$ and A' also forms an independent set. To observe the second statement, let v_1, v_2, v_3 be three vertices in A_0. The number of edges induced on v_1, v_2, v_3, a, b, c is at most nine, because every edge needs to be incident to exactly two vertices of $\{a,b,c\}$ by the definition of A_0. However, 6-vertex induced subgraphs of B_n have either 0, 10, 16, or 18 edges. We conclude that $\{v_1, v_2, v_3, a, b, c\}$ induces no edge in G'. Thus, A' is also an independent set in G' and therefore G' is 2-colourable. ■

Claim 5.5. *We have $|E(G')| \geq (1 - 2\sqrt{\delta_2})\frac{n^3}{8}$.*

Proof. By Claim 5.4, G' is 2-colourable and we can partition the vertex set $V(G') = A \cup B$ such that A and B are independent sets. Let $a \in [0,1]$ such that $|A| = an$ and $|B| = (1-a)n$. We have

$$(1-20\delta_2)\,\frac{5}{8}\frac{n^4}{2} \le \mathrm{co}_2(G') \le \left(\frac{a^2}{2}(1-a)^2 + \frac{(1-a)^2}{2}a^2 + a(1-a)\right)n^4 \le \frac{5}{4}a(1-a)n^4.$$

Thus, $4a(1-a) \ge 1-20\delta_2$. We conclude $1/2 - 3\sqrt{\delta_2} \le a \le 1/2 + 3\sqrt{\delta_2}$, otherwise

$$4a(1-a) < 4\left(\frac{1}{2} - 3\sqrt{\delta_2}\right)\left(\frac{1}{2} + 3\sqrt{\delta_2}\right) = 1 - 36\delta_2,$$

a contradiction. For every edge $e \in E(G')$, we have $w_{G'}(e) \le \left(5/2 + 3\sqrt{\delta_2}\right)n$. Therefore,

$$(1-20\delta_2)\,\frac{5}{8}\frac{n^4}{2} \le \mathrm{co}_2(G') = \sum_{e\in E(G')} w_{G'}(e) \le |E(G')|\left(\frac{5}{2} + 3\sqrt{\delta_2}\right)n.$$

Thus,

$$|E(G')| \ge \frac{(1-20\delta_2)}{\left(1+\frac{6}{5}\sqrt{\delta_2}\right)}\frac{n^3}{8} \ge (1-2\sqrt{\delta_2})\frac{n^3}{8}.$$

■

The 3-graph G is δ_2-near to G'. By Claims 5.4 and 5.5, G' is $\varepsilon/2$-near to B_n. Therefore we can conclude that G is $\delta_2 + \varepsilon/2 \le \varepsilon$-near to B_n. ■

5.3 Discussion on Cliques

Keevash and Mubayi [31] constructed the following family of 3-graphs obtaining the best-known lower bound for the Turán density of cliques. Denote $\mathcal{D}_k$ the family of directed graphs on $k-1$ vertices that are unions of vertex-disjoint directed cycles. Cycles of length two are allowed, but loops are not. Let $D \in \mathcal{D}_k$ and $V = [n] = V_1 \cup \ldots \cup V_{k-1}$ be a vertex partition with class sizes as balanced as possible, that is $||V_i| - |V_j|| \le 1$ for all $i \neq j$. Denote $G(D)$ the 3-graph on V where a triple is a non-edge iff it is contained in some V_i or if it has two vertices in V_i and one vertex in V_j where (i,j) is an arc of D. The 3-graph $G(D)$ is K_k^3-free and has edge density $1-(2/t)^2 + o(1)$. While all 3-graphs $D \in \mathcal{D}_k$ give the same edge density for $G(D)$, up to isomorphism there is only one where $G(D)$ is maximising the codegree squared sum. Let $D_k^* \in \mathcal{D}_k$ be the directed graph on $k-1$ vertices $v_1, \ldots, v_{k-1}$ such that if k odd, then

$$(v_iv_{i+1}), (v_{i+1}v_i) \in E(D_k^*) \quad \text{for all odd } i,$$

and if k even, then

$$\begin{aligned} &(v_iv_{i+1}), (v_{i+1}v_i) \in E(D_k^*) \quad \text{for all odd } i \le k-5 \\ \text{and} \quad &(v_{k-3}v_{k-2}), (v_{k-2}v_{k-1}), (v_{k-1}v_{k-3}) \in E(D_k^*). \end{aligned}$$

Note that D_k^* is maximising the number of directed cycles. The 3-graph $G(D_4^*)$ is isomorphic to C_n and $G(D_5^*)$ is isomorphic to B_n. See Figure 4 for a drawing of D_7^*, D_8^* and the complements $\overline{G(D_7^*)}$ and $\overline{G(D_8^*)}$ of $G(D_7^*)$ and $G(D_8^*)$, respectively. Next, we observe that among all directed graphs $D \in \mathcal{D}_k$, D_k^* maximises the codegree squared sum of $G(D)$.

For a function $f: X \to \mathbb{R}$, and $S \subseteq X$, define

$$\arg\max_{x\in S} f(x) := \{x \in S : f(s) \le f(x) \text{ for all } s \in S\}.$$

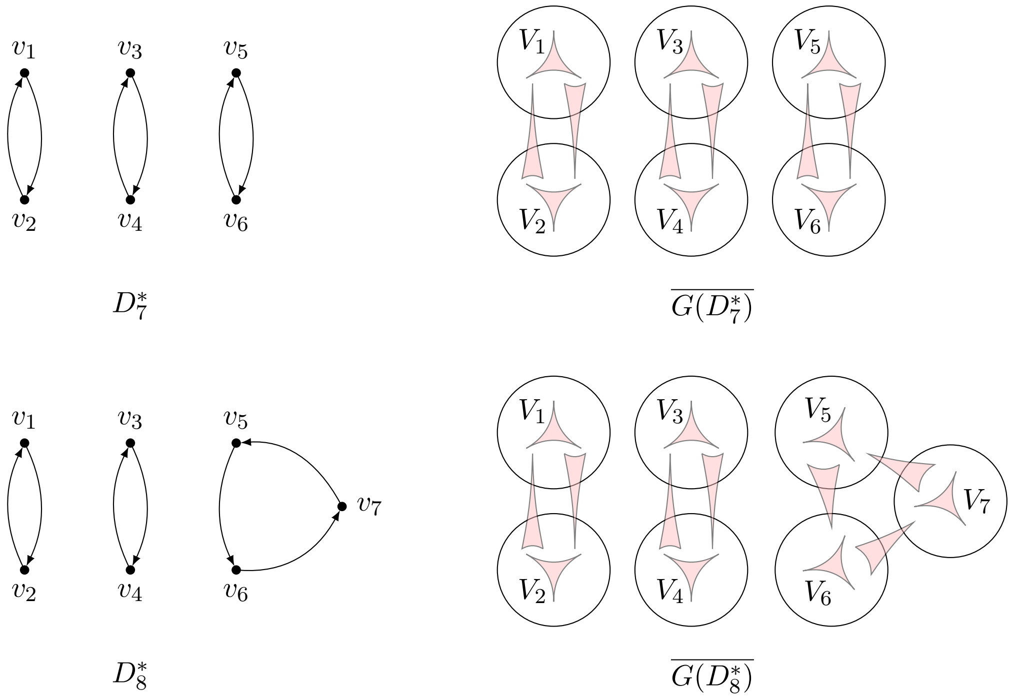


Figure 4: Representations of D_7^*, D_8^* and the complements $\overline{G(D_7^*)}$ and $\overline{G(D_8^*)}$ of $G(D_7^*)$ and $G(D_8^*)$, respectively.

Lemma 5.6. *Let $k \geq 4$. For n sufficiently large, D_k^* is isomorphic to any directed graph in*

$$\underset{D \in \mathcal{D}_k}{\arg\max}\, \mathrm{co}_2(G(D)).$$

Proof. Let $D \in \arg\max_{D \in \mathcal{D}} co_2(G(D))$. Suppose for contradiction that D contains a directed cycle $v_1, v_2, \ldots, v_\ell$ of length $\ell \geq 4$. Construct a directed graph D' by replacing that ℓ-cycle with an $(\ell-2)-$ cycle $v_1, v_4, \ldots, v_{\ell-2}$ and a 2-cycle v_2, v_3. Let $V_1, V_2, \ldots, V_\ell$ be the corresponding classes in G. The only pairs of vertices x, y for which the codegree changes by more than $O(1)$ are described in the following.

- For $x \in V_1, y \in V_2$, $d(x,y)$ increased from $n - n/(k-1) + O(1)$ to $n + O(1)$.
- For $x \in V_3, y \in V_4$, $d(x,y)$ increased from $n - n/(k-1) + O(1)$ to $n + O(1)$.
- For $x \in V_2, y \in V_3$, $d(x,y)$ decreased from $n - n/(k-1) + O(1)$ to $n - 2n/(k-1) + O(1)$.
- For $x \in V_1, y \in V_4$, $d(x,y)$ decreased from $n - n/(k-1) + O(1)$ to $n - 2n/(k-1) + O(1)$ if $\ell = 4$ or from $n + O(1)$ to $n - n/(k-1) + O(1)$ if $\ell > 4$.

Thus,

$$\mathrm{co}_2(G(D')) - \mathrm{co}_2(G(D)) \geq O(1) + \frac{n^2}{(k-1)^2}\left(n^2 - \left(n - \frac{2n}{k-1}\right)^2\right) > 0,$$

a contradiction. Thus, D contains no cycle of length at least 4. Next, towards contradiction, suppose that D contains at least two cycles of length 3. Denote v_1, v_2, v_3 and v_4, v_5, v_6 those two 3-cycles. Let D' be the directed graph constructed from D by replacing those two 3-cycles with three 2-cycles v_1, v_2 and v_3, v_4 and v_5, v_6. The pairs of vertices x, y for which the codegree changed by more than $O(1)$ are among those pairs where $x, y \in V_1 \cup \ldots \cup V_6$ and where x and y were in different classes. It follows that

$$\mathrm{co}_2(G(D')) - \mathrm{co}_2(G(D)) = O(1) + \frac{n^2}{(k-1)^2} n^2 \left(3 + 3\left(1 - \frac{2}{k-1}\right)^2 - 6\left(1 - \frac{1}{k-1}\right)^2\right) > 0,$$

a contradiction. Thus, we can conclude that D contains at most one 3-cycle. Hence, D is isomorphic to D_k^*. ∎

Depending on the parity of k, D_k^* either contains a 3-cycle or not. In the case k is odd, D_k^* contains no 3-cycles and based on Lemma 5.6 it seems reasonable to conjecture that in this case $G(D_k^*)$ could be an asymptotical extremal hypergraph in the ℓ_2-norm.

Question 5.7. *Let $k \geq 7$ odd and $\ell = (k-1)/2$. Is*

$$\sigma(K_k^3) = \lim_{n\to\infty} \frac{\mathrm{co}_2(G(D_k^*))}{\binom{n}{2}(n-2)^2} = 1 - \frac{2}{\ell^2} + \frac{1}{\ell^3} \ ?$$

The situation is slightly different for odd k. It is better to consider an unbalanced version of $G(D_k^*)$. Denote $G^*(D_k^*)$ the 3-graph with the largest codegree squared sum among the following 3-graphs G. Partition the vertex set of G into $[n] = V_1 \cup \ldots \cup V_{k-1}$, where the class sizes are balanced as follow

- $||V_i| - |V_j|| \leq 1$ for all $i \neq j$ with $i, j \leq k-4$ and
- $||V_i| - |V_j|| \leq 1$ for all $i \neq j$ with $k-3 \leq i, j \leq k-1$.

Again, a triple is a non-edge in $G^*(D_k^*)$ iff it is contained in some V_i or if it has two vertices in V_i and one vertex in V_j where (i,j) is an arc of D_k^*.

Question 5.8. *Let $k \geq 6$ even. Is*

$$\sigma(K_k^3) = \lim_{n\to\infty} \frac{\mathrm{co}_2(G^*(D_k^*))}{\binom{n}{2}(n-2)^2} \ ?$$

6 Proof of Theorem 1.7

In this section we prove Theorem 1.7, i.e., we determine the codegree squared extremal number of $F_{3,3}$. Flag algebras give us the following corresponding asymptotical result and also a weak stability version.

Lemma 6.1. *For all $\varepsilon > 0$ there exists $\delta > 0$ and n_0 such that for all $n \geq n_0$: if G is an $F_{3,3}$-free 3-uniform graph on n vertices with $\mathrm{co}_2(G) \geq (1-\delta)\frac{5}{8}n^4/2$, then the densities of all 3-graphs on $4, 5$ and 6 vertices in G that are not contained in B_n are at most ε. Additionally,*

$$\sigma(F_{3,3}) = \frac{5}{8}.$$

This result implies the following stability theorem.

Theorem 6.2. *For every $\varepsilon > 0$ there is $\delta > 0$ and n_0 such that if G is an $F_{3,3}$-free 3-uniform hypergraph on $n \geq n_0$ vertices with $\mathrm{co}_2(G) \geq (1-\delta)\frac{5}{8}\frac{n^4}{2}$, then we can partition $V(G)$ as $A \cup B$ such that $e(A) + e(B) \leq \varepsilon n^3$ and $e(A, B) \geq \frac{1}{8}n^3 - \varepsilon n^3$.*

Proof. The proof is the same as the proof of Theorem 1.6, except instead of applying Lemma 5.3 we apply Lemma 6.1. ∎

We now determine the exact extremal number by using the stability result, Theorem 6.2, and a standard cleaning technique, see for example [26, 32, 34, 43]. To do so we will first prove the statement under an additional universal minimum-degree-type assumption.

Theorem 6.3. *There exists n_0 such that for all $n \geq n_0$ the following holds. Let G be an $F_{3,3}$-free n-vertex 3-graph such that*

$$q(x) := \sum_{y \in V, y \neq x} d(x,y)^2 + 2 \sum_{\{v,w\} \in E(L(x))} d(v,w) \geq \frac{5}{4}n^3 - 6n^2 =: d(n) \tag{6}$$

for all $x \in V(G)$. Then,

$$\mathrm{co}_2(G) \leq \mathrm{co}_2(B_n) = \binom{\lceil \frac{n}{2} \rceil}{2} \left\lfloor \frac{n}{2} \right\rfloor^2 + \binom{\lfloor \frac{n}{2} \rfloor}{2} \left\lceil \frac{n}{2} \right\rceil^2 + \left\lceil \frac{n}{2} \right\rceil \left\lfloor \frac{n}{2} \right\rfloor (n-2)^2.$$

Furthermore, B_n is the unique such 3-graph G satisfying $\mathrm{co}_2(G) = \mathrm{exco}_2(n, F_{3,3})$.

Proof. Let G be a 3-uniform $F_{3,3}$-free hypergraph which has a codegree squared sum of at least $\mathrm{co}_2(G) \geq \mathrm{co}_2(B_n)$ and satisfies (6). Choose $\varepsilon = 10^{-10}$ and apply Theorem 6.2. We get a vertex partition $A \cup B$ with $e(A) + e(B) \leq \varepsilon n^3$ and $e^c(A,B) \leq \varepsilon n^3$. Among all such partitions choose one which minimises $e(A) + e(B)$. We can assume that $|L_B(x)| \geq |L_A(x)|$ for all $x \in A$ and $|L_A(x)| \geq |L_B(x)|$ for all $x \in B$, as otherwise we could switch a vertex from one class to the other class and strictly decrease both $e(A) + e(B)$ and $e^c(A,B)$, a contradiction. This is not possible, because we chose A and B minimising $e(A)+e(B)$. We start by making an observation about the class sizes.

Claim 6.4. *We have*

$$\left| |A| - \frac{n}{2} \right| \leq 2\sqrt{\varepsilon} n \quad \textit{and} \quad \left| |B| - \frac{n}{2} \right| \leq 2\sqrt{\varepsilon} n.$$

Proof. Assume that $|A| < n/2 - 2\sqrt{\varepsilon}n$. Then, we have

$$\begin{aligned} e(A,B) &\leq \binom{|A|}{2}|B| + |A|\binom{|B|}{2} \leq \frac{1}{2}|A|(n-|A|)n \\ &< \frac{1}{2}\left(\frac{n}{2} - 2\sqrt{\varepsilon}n\right)\left(\frac{n}{2} + 2\sqrt{\varepsilon}n\right) n < \frac{1}{8}n^3 - \varepsilon n^3, \end{aligned}$$

a contradiction. Thus, $|A| \geq n/2 - 2\sqrt{\varepsilon}n$. Similarly, we get $|B| \geq n/2 - 2\sqrt{\varepsilon}n$. ■

Define *junk* sets J_A, J_B to be the sets of vertices which are not typical, i.e.,

$$\begin{aligned} J_A &:= \left\{x \in A : |L^c_{A,B}(x)| \geq \sqrt{\varepsilon}n^2\right\} \cup \left\{x \in A : |L_A(x)| \geq \sqrt{\varepsilon}n^2\right\}, \text{ and} \\ J_B &:= \left\{x \in B : |L^c_{A,B}(x)| \geq \sqrt{\varepsilon}n^2\right\} \cup \left\{x \in B : |L_B(x)| \geq \sqrt{\varepsilon}n^2\right\}. \end{aligned}$$

These junk sets need to be small.

Claim 6.5. *We have $|J_A|, |J_B| \leq 5\sqrt{\varepsilon}n$.*

Proof. Towards contradiction assume that $|J_A| > 5\sqrt{\varepsilon}n$. Then the number of vertices $x \in J_A$ satisfying $|L^c_{A,B}(x)| \geq \sqrt{\varepsilon}n^2$ is at least $2\sqrt{\varepsilon}n$ or the number of vertices $x \in J_A$ satisfying $|L_A(x)| \geq \sqrt{\varepsilon}n^2$ is at least $3\sqrt{\varepsilon}n$. If the first case holds, then we get $e^c(A,B) > \varepsilon n^3$. In the second case we have $e(A) > \varepsilon n^3$. Both are in contradiction with the choice of the partition $A \cup B$. Thus, $|J_A| \leq 5\sqrt{\varepsilon}n$. The second statement of this claim, $|J_B| \leq 5\sqrt{\varepsilon}n$, follows by a similar argument. ■

Claim 6.6. *$A \setminus J_A$ and $B \setminus J_B$ are independent sets.*

Proof. If there is an edge $a_1a_2a_3$ with $a_1, a_2, a_3 \in A \setminus J_A$, since all its vertices satisfy $|L_B^c(a_i)| \leq \sqrt{\varepsilon}n^2$, we can find a triangle in $L_B(a_1) \cap L_B(a_2) \cap L_B(a_3)$, call its vertices b_1, b_2, b_3. However, now $\{b_1, b_2, b_3, a_1, a_2, a_3\}$ spans an $F_{3,3}$ in G, a contradiction. A similar proof gives that $B \setminus J_B$ is an independent set. ∎

Claim 6.7. *There is no edge $a_1a_2a_3$ with $a_1 \in J_A$, $a_2, a_3 \in A \setminus J_A$ or with $a_1 \in J_B$, $a_2, a_3 \in B \setminus J_B$.*

Proof. Let $a_1a_2a_3$ be an edge with $a_1 \in J_A$, $a_2, a_3 \in A \setminus J_A$. We show that $q(a_1) < d(n)$, to get a contradiction with (6). Let M_i, for $i = 2, 3$, be the set of non-edges in $L_B(a_i)$ and $L_{A,B}(a_i)$. Set $K = L(a_1) - M_2 - M_3$. Since $|M_2|, |M_3| \leq 2\sqrt{\varepsilon}n^2$, we have $|E(K)| \geq |L(a_1)| - 4\sqrt{\varepsilon}n^2$. Let

$$\Delta = \frac{\max_{x \in A \setminus \{a_1,a_2,a_3\}} |N_K(x) \cap B|}{n},$$

be the maximum size of a neighbourhood in the graph K in B of a vertex in A, scaled by n. We have $0 \leq \Delta \leq |B|/n \leq 1/2 + \sqrt{\varepsilon}$. Let $z \in A \setminus \{a_1, a_2, a_3\}$ such that $|N_K(z) \cap B)| = \Delta n$. Observe that $N_K(z) \cap B$ is an independent set in K, otherwise if $v, w \in N_K(z) \cap B$ with $vw \in E(K)$, then $\{v, w, z, a_1, a_2, a_3\}$ spans a $F_{3,3}$ in G. Now,

$$\sum_{x \in V \setminus \{a_1\}} d(a_1, x)^2 = \sum_{x \in V \setminus \{a_1\}} \deg_{L(a_1)}(x)^2 \leq 16\sqrt{\varepsilon}n^3 + \sum_{x \in V(K)} \deg_K(x)^2, \tag{7}$$

because for each edge removed from the linkgraph $L(a_1)$ the degree squared sum can go down by at most $4n$. Now, we bound the sum on the right hand side of (7) from above. For $x \in A$, $\deg_k(x) \leq |A| + \Delta n$ and for $x \in N(z) \cap B$, $\deg_k(x) \leq n - \Delta n$. Thus, we get

$$\begin{aligned}
&\sum_{x \in V \setminus \{a_1\}} d(a_1, x)^2 \leq 16\sqrt{\varepsilon}n^3 + |A|(|A| + \Delta n)^2 + \Delta n(n - \Delta n)^2 + (|B| - \Delta n)n^2 \\
&\leq \left(\frac{n}{2} + 2\sqrt{\varepsilon}n\right)\left(\frac{n}{2} + 2\sqrt{\varepsilon}n + \Delta n\right)^2 + \Delta n(n - \Delta n)^2 + \left(\frac{n}{2} + 2\sqrt{\varepsilon}n - \Delta n\right)n^2 + 16\sqrt{\varepsilon}n^3 \\
&\leq n^3\left(\frac{1}{2}\left(\frac{1}{2} + \Delta\right)^2 + \Delta(1 - \Delta)^2 + \left(\frac{1}{2} - \Delta\right) + 25\sqrt{\varepsilon}\right) = n^3\left(\frac{5}{8} + \frac{\Delta}{2} - \frac{3}{2}\Delta^2 + \Delta^3 + 25\sqrt{\varepsilon}\right).
\end{aligned} \tag{8}$$

Furthermore, we can give an upper bound for the second summand in $q(a_1)$:

$$2\sum_{\{x,y\} \in E(L(a_1))} d(x, y) \leq 8\sqrt{\varepsilon}n^3 + 2\sum_{\{x,y\} \in E(K)} d(x, y), \tag{9}$$

where we used that for each edge removed from G, the sum on the left hand side in (9) is lowered by at most n. Now, we will give an upper bound for the right hand side of (9). For edges $xy \in E(K[A])$ not incident to J_A we have $d_G(x, y) \leq |J_A| + |B|$ because by Claim 6.6 they have no neighbour in $A \setminus J_A$. Similarly, for edges $xy \in E(K[B])$ not incident to J_B we have $d_G(x, y) \leq |J_B| + |A|$. For all other edges $xy \in E(K)$, we will use the trivial bound $d_G(x, y) \leq n$. We have

$$\begin{aligned}
2\sum_{\{x,y\} \in E(L(a_1))} d(x, y) \leq 8\sqrt{\varepsilon}n^3 + 2\Big(&e(K[A, B])n + e(K[A])(|J_A| + |B|) + |J_A||A|n \\
&+ e(K[B])(|J_B| + |B|) + |J_B||B|n\Big).
\end{aligned} \tag{10}$$

By the choice of our partition we have $|L_A(x_1)| \leq |L_B(x_1)|$ and thus $e(K[A]) \leq e(K[B]) + 4\sqrt{\varepsilon}n^2$. Therefore, by upper bounding the right hand side in (10) we get

$$\begin{aligned}
2 \sum_{\{x,y\} \in E(L(a_1))} d(x,y) &\leq 2\left(\Delta n^2|A| + 2e(K[B])\left(7\sqrt{\varepsilon}n + \frac{n}{2}\right) + 18\sqrt{\varepsilon}n^3\right) \\
&\leq 2n^3\left(\frac{\Delta}{2} + \frac{e(G[B])}{n^2} + 30\sqrt{\varepsilon}\right) \\
&\leq 2n^3\left(\frac{\Delta}{2} + \Delta\left(\frac{|B|}{n} - \Delta\right) + \frac{1}{4}\left(\frac{|B|}{n} - \Delta\right)^2 + 30\sqrt{\varepsilon}\right) \\
&\leq 2n^3\left(\frac{\Delta}{2} + \Delta\left(\frac{1}{2} - \Delta\right) + \frac{1}{4}\left(\frac{1}{2} - \Delta\right)^2 + 40\sqrt{\varepsilon}\right) \\
&\leq n^3\left(-\frac{3}{2}\Delta^2 + \frac{3}{2}\Delta + \frac{1}{8} + 80\sqrt{\varepsilon}\right), \qquad (11)
\end{aligned}$$

where we used that $e(K[B]) \leq \Delta n(|B| - \Delta n) + \frac{(|B|-\Delta n)^2}{4}$, because $K[B]$ contains an independent set of size Δn and is triangle-free. Now, we can combine (8) and (11) to upper bound $q(a_1)$.

$$\begin{aligned}
q(a_1) &\leq n^3\left(\frac{5}{8} + \frac{\Delta}{2} - \frac{3}{2}\Delta^2 + \Delta^3 + 25\sqrt{\varepsilon}\right) + n^3\left(-\frac{3}{2}\Delta^2 + \frac{3}{2}\Delta + \frac{1}{8} + 80\sqrt{\varepsilon}\right) \\
&= n^3\left(\Delta^3 - 3\Delta^2 + 2\Delta + \frac{3}{4} + 105\sqrt{\varepsilon}\right) \leq \left(\frac{2}{3\sqrt{3}} + \frac{3}{4} + 105\sqrt{\varepsilon}\right)n^3 < \frac{5}{4}n^3 - 6n^2,
\end{aligned}$$

contradicting (6). In the second-to-last inequality we used that the polynomial $\Delta^3 - 3\Delta^2 + 2\Delta$ obtains its maximum in $[0,1]$ at $\Delta = 1 - \frac{1}{\sqrt{3}}$. ∎

Now, we can make use of Claim 6.7 to show that there is no edge inside A, respectively inside B.

Claim 6.8. *A and B are independent sets.*

Proof. Let $\{a_1, a_2, a_3\} \subset A$ span an edge. Again, $L_B(a_1) \cap L_B(a_2) \cap L_B(a_3)$ is triangle-free. Thus, $|L_B(a_1) \cap L_B(a_2) \cap L_B(a_3)| \leq |B|^2/4$. By the pigeon-hole principle, we may assume without loss of generality that $|L_B(a_1)| \leq 5|B|^2/12$. Furthermore, by Claims 6.6 and 6.7, $|L_A(a_1)| \leq |J_A||A| \leq 5\sqrt{\varepsilon}n^2$. Again, our strategy will be to give an upper bound on $q(a_1)$. Let L be the graph obtained from $L(a_1)$ by removing all edges inside A.

$$\begin{aligned}
\sum_{x \in V \setminus \{a_1\}} d(a_1, x)^2 &= \sum_{x \in V \setminus \{a_1\}} \deg_{L(a_1)}(x)^2 \leq 20\sqrt{\varepsilon}n^3 + \sum_{x \in V(L)} \deg_L(x)^2 \\
&\leq 20\sqrt{\varepsilon}n^3 + |B|n^2 + |A||B|^2 \leq n^3\left(\frac{5}{8} + 30\sqrt{\varepsilon}\right). \qquad (12)
\end{aligned}$$

Furthermore,

$$\begin{aligned}
2 \sum_{\{x,y\} \in E(L(a_1))} d(x,y) &\leq 10\sqrt{\varepsilon}n^3 + 2\sum_{xy \in E(L)} d(x,y) \\
&\leq 2\left(\frac{5}{12}|B|^2\left(|A| + |J_B|\right) + 5\sqrt{\varepsilon}n^3 + |A||B|n\right) \\
&\leq 2n^3\left(\frac{5}{96} + 20\sqrt{\varepsilon} + \frac{1}{4}\right) = n^3\left(\frac{29}{48} + 40\sqrt{\varepsilon}\right). \qquad (13)
\end{aligned}$$

Thus, by combining (12) and (13), we can give an upper bound on $q(a_1)$,

$$q(a_1) \le \left(\frac{5}{8} + 30\sqrt{\varepsilon}\right) n^3 + n^3 \left(\frac{29}{48} + 40\sqrt{\varepsilon}\right) = n^3 \left(\frac{59}{48} + 70\sqrt{\varepsilon}\right) < \frac{5}{4}n^3 - 6n^2,$$

contradicting (6). Therefore A is an independent set. By a similar argument B is also an independent set.

∎

By Claim 6.8, G is 2-colourable. Since among all 2-colourable 3-graphs B_n has the largest codegree squared sum, we conclude $\mathrm{co}_2(G) \le \mathrm{co}_2(B_n)$. This completes the proof of Theorem 6.3.

∎

We now complete the proof of Theorem 6.3 by showing that imposing the additional assumption (6) is not more restrictive.

Proof of Theorem 1.7. Let G be an n-vertex 3-uniform $F_{3,3}$-free hypergraph which has a codegree squared sum of at least $\mathrm{co}_2(G) \ge \mathrm{co}_2(B_n)$. Set $d(n) = 5/4n^3 - 6n^2$ and note that $\mathrm{co}_2(B_n) - \mathrm{co}_2(B_{n-1}) > d(n) + 1$. We claim that we can assume that every vertex $x \in V(G)$ satisfies (6). Otherwise, we can remove a vertex x with $q(x) < d(n)$ to get G_{n-1} with $\mathrm{co}_2(G_{n-1}) \ge \mathrm{co}_2(B_n) - d(n) \ge \mathrm{co}_2(B_{n-1}) + 1$. By repeating this process as long as possible, we obtain a sequence of hypergraphs G_m on m vertices with $\mathrm{co}_2(G_m) \ge \mathrm{co}_2(B_m) + n - m$, where G_m is the hypergraph obtained from G_{m+1} by deleting a vertex x with $q(x) \le d(m+1)$. We cannot continue until we reach a hypergraph on $n_0 = n^{1/4}$ vertices, as then $\mathrm{co}_2(G_{n_0}) > n - n_0 > \binom{n_0}{2}(n_0 - 2)^2$ which is impossible. Therefore, the process stops at some n' where $n \ge n' \ge n_0$ and we obtain the corresponding hypergraph $G_{n'}$ satisfying $q(x) \ge d(n')$ for all $x \in V(G_{n'})$ and $\mathrm{co}_2(G_{n'}) \ge \mathrm{co}_2(B_{n'})$ (with strict inequality if $n > n'$). Hence, we can assume that G satisfies $q(x) \ge d(n')$ for all $x \in V(G_{n'})$. Applying Theorem 6.3 finishes the proof. ∎

Acknowledgements

We thank an anonymous referee for many useful comments and suggestions, in particular for pointing out a shorter proof of Theorem 1.5.

References

[1] N. Alon, H. Naves, and B. Sudakov. On the maximum quartet distance between phylogenetic trees. *SIAM J. Discrete Math.*, 30(2):718–735, 2016. doi:10.1137/15M1041754.

[2] R. Baber. Turán densities of hypercubes. *arXiv preprint*, 2012. arXiv:1201.3587.

[3] R. Baber and J. Talbot. Hypergraphs do jump. *Combin. Probab. Comput.*, 20(2):161–171, 2011. doi:10.1017/S0963548310000222.

[4] J. Balogh, F. C. Clemen, and B. Lidický. The codegree squared density. *In preparation*, 2020.

[5] J. Balogh, P. Hu, B. Lidický, F. Pfender, J. Volec, and M. Young. Rainbow triangles in three-colored graphs. *J. Combin. Theory Ser. B*, 126:83–113, 2017. doi:10.1016/j.jctb.2017.04.002.

[6] J. Balogh, P. Hu, B. Lidický, O. Pikhurko, B. Udvari, and J. Volec. Minimum number of monotone subsequences of length 4 in permutations. *Combin. Probab. Comput.*, 24(4):658–679, 2015. doi:10.1017/S0963548314000820.

[7] J. Balogh, B. Lidický, and G. Salazar. Closing in on Hill's conjecture. *SIAM J. Discrete Math.*, 33(3):1261–1276, 2019. doi:10.1137/17M1158859.

[8] B. Bollobás. Three-graphs without two triples whose symmetric difference is contained in a third. *Discrete Math.*, 8:21–24, 1974. doi:10.1016/0012-365X(74)90105-8.

[9] B. Borchers. CSDP, a C library for semidefinite programming. volume 11/12, pages 613–623. 1999. Interior point methods. doi:10.1080/10556789908805765.

[10] W. G. Brown. On an open problem of Paul Turán concerning 3-graphs. In *Studies in pure mathematics*, pages 91–93. Birkhäuser, Basel, 1983.

[11] F. Chung and L. Lu. An upper bound for the Turán number $t_3(n, 4)$. *J. Combin. Theory Ser. A*, 87(2):381–389, 1999. doi:10.1006/jcta.1998.2961.

[12] J. Cummings, D. Král', F. Pfender, K. Sperfeld, A. Treglown, and M. Young. Monochromatic triangles in three-coloured graphs. *J. Combin. Theory Ser. B*, 103(4):489–503, 2013. doi:10.1016/j.jctb.2013.05.002.

[13] D. de Caen. On upper bounds for 3-graphs without tetrahedra. *Congr. Numer.*, 62:193–202, 1988. Seventeenth Manitoba Conference on Numerical Mathematics and Computing (Winnipeg, MB, 1987).

[14] P. Erdős. On extremal problems of graphs and generalized graphs. *Israel J. of Math.*, 2(3):183–190, 1964.

[15] P. Erdős and M. Simonovits. A limit theorem in graph theory. *Studia Sci. Math. Hungar.*, 1:51–57, 1966.

[16] P. Erdős and M. Simonovits. Supersaturated graphs and hypergraphs. *Combinatorica*, 3(2):181–192, 1983. doi:10.1007/BF02579292.

[17] P. Erdős and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946. doi:10.1090/S0002-9904-1946-08715-7.

[18] V. Falgas-Ravry. On the codegree density of complete 3-graphs and related problems. *Electron. J. Combin.*, 20(4):Paper 28, 14, 2013.

[19] V. Falgas-Ravry, E. Marchant, O. Pikhurko, and E. R. Vaughan. The codegree threshold for 3-graphs with independent neighborhoods. *SIAM J. Discrete Math.*, 29(3):1504–1539, 2015. doi:10.1137/130926997.

[20] V. Falgas-Ravry, O. Pikhurko, E. Vaughan, and J. Volec. The codegree threshold of K_4^-. *Electronic Notes in Discrete Mathematics*, 61:407–413, 2017.

[21] V. Falgas-Ravry and E. R. Vaughan. Turán H-densities for 3-graphs. *Electron. J. Combin.*, 19(3):Paper 40, 26, 2012. doi:10.37236/2733.

[22] 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(1):21–54, 2013. doi:10.1017/S0963548312000508.

[23] D. G. Fon-Der-Flaass. A method for constructing (3, 4)-graphs. *Mat. Zametki*, 44(4):546–550, 559, 1988. doi:10.1007/BF01158925.

[24] P. Frankl and Z. Füredi. A new generalization of the Erdős-Ko-Rado theorem. *Combinatorica*, 3(3–4):341–349, 1983. doi:10.1007/BF02579190.

[25] A. Frohmader. More constructions for Turán's (3, 4)-conjecture. *Electron. J. Combin.*, 15(1):Research Paper 137, 23, 2008. doi:10.37236/861.

[26] Z. Füredi and M. Simonovits. Triple systems not containing a Fano configuration. *Combin. Probab. Comput.*, 14(4):467–484, 2005. doi:10.1017/S0963548305006784.

[27] J. Goldwasser and R. Hansen. The exact Turán number of $F(3, 3)$ and all extremal configurations. *SIAM J. Discrete Math.*, 27(2):910–917, 2013. doi:10.1137/110841837.

[28] A. Grzesik, P. Hu, and J. Volec. Minimum number of edges that occur in odd cycles. *J. Combin. Theory Ser. B*, 137:65–103, 2019. doi:10.1016/j.jctb.2018.12.003.

[29] J. Hladký, D. Král', and S. Norin. Counting flags in triangle-free digraphs. *Combinatorica*, 37(1):49–76, 2017. doi:10.1007/s00493-015-2662-5.

[30] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs*, volume 45. John Wiley & Sons, 2011.

[31] P. Keevash. Hypergraph Turán problems. In *Surveys in combinatorics 2011*, volume 392 of *London Math. Soc. Lecture Note Ser.*, pages 83–139. Cambridge Univ. Press, Cambridge, 2011.

[32] P. Keevash and D. Mubayi. Stability theorems for cancellative hypergraphs. *J. Combin. Theory Ser. B*, 92(1):163–175, 2004. doi:10.1016/j.jctb.2004.05.003.

[33] P. Keevash and D. Mubayi. The Turán number of $F_{3,3}$. *Combin. Probab. Comput.*, 21(3):451–456, 2012. doi:10.1017/S0963548311000678.

[34] P. Keevash and B. Sudakov. The Turán number of the Fano plane. *Combinatorica*, 25(5):561–574, 2005. doi:10.1007/s00493-005-0034-2.

[35] A. V. Kostochka. A class of constructions for Turán's (3, 4)-problem. *Combinatorica*, 2(2):187–192, 1982. doi:10.1007/BF02579317.

[36] D. Krǎl̆, L. Mach, and J.-S. Sereni. A new lower bound based on Gromov's method of selecting heavily covered points. *Discrete Comput. Geom.*, 48(2):487–498, 2012. doi:10.1007/s00454-012-9419-3.

[37] N. Linial and A. Morgenstern. On the number of 4-cycles in a tournament. *J. Graph Theory*, 83(3):266–276, 2016. doi:10.1002/jgt.21996.

[38] A. Lo and K. Markström. ℓ-degree Turán density. *SIAM J. Discrete Math.*, 28(3):1214–1225, 2014. doi:10.1137/120895974.

[39] A. Lo and Y. Zhao. Codegree Turán density of complete r-uniform hypergraphs. *SIAM J. Discrete Math.*, 32(2):1154–1158, 2018. doi:10.1137/18M1163956.

[40] D. Mubayi. The co-degree density of the Fano plane. *J. Combin. Theory Ser. B*, 95(2):333–337, 2005. doi:10.1016/j.jctb.2005.06.001.

[41] D. Mubayi and Y. Zhao. Co-degree density of hypergraphs. *J. Combin. Theory Ser. A*, 114(6):1118–1132, 2007. doi:10.1016/j.jcta.2006.11.006.

[42] B. Nagle. Turán related problems for hypergraphs. In *Proceedings of the Thirtieth Southeastern International Conference on Combinatorics, Graph Theory, and Computing (Boca Raton, FL, 1999)*, volume 136, pages 119–127, 1999.

[43] O. Pikhurko. The minimum size of 3-graphs without a 4-set spanning no or exactly three edges. *European J. Combin.*, 32(7):1142–1155, 2011. doi:10.1016/j.ejc.2011.03.006.

[44] O. Pikhurko, J. Sliačan, and K. Tyros. Strong forms of stability from flag algebra calculations. *J. Combin. Theory Ser. B*, 135:129–178, 2019. doi:10.1016/j.jctb.2018.08.001.

[45] A. A. Razborov. Flag algebras. *J. Symbolic Logic*, 72(4):1239–1282, 2007. doi:10.2178/jsl/1203350785.

[46] A. A. Razborov. On 3-hypergraphs with forbidden 4-vertex configurations. *SIAM J. Discrete Math.*, 24(3):946–963, 2010. doi:10.1137/090747476.

[47] A. A. Razborov. Flag algebras: an interim report. In *The mathematics of Paul Erdős. II*, pages 207–232. Springer, New York, 2013. doi:10.1007/978-1-4614-7254-4_16.

[48] A. A. Razborov. What is...a flag algebra? *Notices Amer. Math. Soc.*, 60(10):1324–1327, 2013. doi:10.1090/noti1051.

[49] C. Reiher, V. Rödl, and M. Schacht. Embedding tetrahedra into quasirandom hypergraphs. *J. Combin. Theory Ser. B*, 121:229–247, 2016. doi:10.1016/j.jctb.2016.06.008.

[50] C. Reiher, V. Rödl, and M. Schacht. Hypergraphs with vanishing Turán density in uniformly dense hypergraphs. *J. Lond. Math. Soc. (2)*, 97(1):77–97, 2018. doi:10.1112/jlms.12095.

[51] V. Rödl and M. Schacht. Generalizations of the removal lemma. *Combinatorica*, 29(4):467–501, 2009. doi:10.1007/s00493-009-2320-x.

[52] A. Sidorenko. Systems of sets that have the T-property. *Moscow University Mathematics Bulletin 36*, 36:22–26, 1981.

[53] A. Sidorenko. What we know and what we do not know about Turán numbers. *Graphs Combin.*, 11(2):179–199, 1995. doi:10.1007/BF01929486.

[54] A. Sidorenko. Extremal problems on the hypercube and the codegree Turán density of complete r-graphs. *SIAM J. Discrete Math.*, 32(4):2667–2674, 2018. doi:10.1137/17M1151171.

[55] J. Sliačan and W. Stromquist. Improving bounds on packing densities of 4-point permutations. *Discrete Math. Theor. Comput. Sci.*, 19(2):Paper No. 3, 18, 2017. doi:10.1109/mcse.2017.21.

[56] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9)*, 2021. URL: https://www.sagemath.org.