Stability theorems for cancellative hypergraphs

Peter Keevash * Dhruv Mubayi †

Dec 1, 2002

Abstract

A *cancellative* hypergraph has no three edges A, B, C with $A \Delta B \subset C$. We give a new short proof of an old result of Bollobás, which states that the maximum size of a cancellative triple system is achieved by the balanced complete tripartite 3-graph.

One of the two forbidden subhypergraphs in a cancellative 3-graph is $F_5 = \{abc, abd, cde\}$. For $n \geq 33$ we show that the maximum number of triples on n vertices containing no copy of F_5 is also achieved by the balanced complete tripartite 3-graph. This strengthens a theorem of Frankl and Füredi, who proved it for $n \geq 3000$.

For both extremal results, we show that a 3-graph with almost as many edges as the extremal example is approximately tripartite. These stability theorems are analogous to the Simonovits stability theorem for graphs.

1 Introduction

Given an k-uniform hypergraph $\mathcal{F}$, the Turán number $ex(n, \mathcal{F})$ of $\mathcal{F}$ is the maximum number of edges in an k-uniform hypergraph on n vertices that does not contain a copy of $\mathcal{F}$. Determining these numbers is one of the central problems in extremal combinatorics. Much is known for ordinary graphs (the case $k = 2$), going back to Turán [14], who solved the problem for complete graphs. The *Turán graph* $T_r(n)$ is the complete r-partite graph on n vertices, which is 'balanced', i.e. no two part sizes differ by more than one. Write $t_r(n)$ for the number of edges in $T_r(n)$. Then Turán's theorem states that a graph G on n vertices containing no copy of K_{r+1} can have at most $t_r(n)$ edges, and equality holds only when $G = T_r(n)$.

In contrast, for $k > 2$, the problem of finding the numbers $ex(n, \mathcal{F})$ is notoriously difficult. Exact results on hypergraph Turán numbers are very rare (see [5, 12] for surveys). It is even difficult to determine the *Turán density* of a k-graph $\mathcal{F}$, i.e. $\pi(\mathcal{F}) = \lim_{n\to\infty} \text{ex}(n, \mathcal{F})/\binom{n}{k}$. Recently, a new technique was developed by de Caen and Füredi [3] to show that, when $\mathcal{F}$ is the Fano plane, $\pi(\mathcal{F}) = 3/4$. Since then there has been considerable progress on Turán problems for 3-uniform

*Department of Mathematics, Princeton, NJ 08540. Email address: keevash@math.princeton.edu

†Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, 851 S. Morgan Street, Chicago, IL 60607. Email address: mubayi@math.uic.edu. Research supported in part by NSF grant DMS-9970325

hypergraphs. The second author and Rödl [9] developed this method to determine many new Turán densities, and the Turán number of the Fano plane was determined by the first author and Sudakov [7], and independently by Füredi and Simonovits [6], thus answering a conjecture of Sós.

In the 1960's, Katona posed the problem of determining the maximum number of triples in an n vertex 3-graph H such that no three distinct triples A, B, C have the property that $B \triangle C \subset A$. Here $\triangle$ is the symmetric difference, so $B \triangle C = (B - C) \cup (C - B)$. Equivalently, H has the property that whenever $A \cup B = A \cup C$, we have $B = C$. Such 3-graphs are called *cancellative*.

For graphs, the condition $B \Delta C \subset A$ can only occur if ABC form a triangle. Then the special case of Turán's theorem due to Mantel tells us that $e(H) \leq t_2(n) = \lfloor n^2/4 \rfloor$. This was the motivation for Katona's problem. Moving on to 3-graphs, we see that if $B \neq C$ then $B \Delta C \subset A$ can only occur when B and C have two points in common. This leads us to identify the two non-isomorphic configurations that are forbidden in a cancellative 3-graph: $F_4 = \{abc, abd, bcd\}$ and $F_5 = \{abc, abd, cde\}$.

Definition 1.1. *$S(n)$ is the complete 3-partite 3-graph on n vertices with part sizes $\lfloor n/3 \rfloor, \lfloor (n+1)/3 \rfloor \lfloor (n+2)/3 \rfloor$. Let $s(n)$ be the number of edges in $S(n)$.*

Katona conjectured that the unique cancellative 3-graph on n vertices with the maximum number of triples is $S(n)$, and this was proved by Bollobás [1].

Theorem 1.2. (Bollobás) *A cancellative 3-graph on n vertices has at most $s(n)$ edges, with equality only for $S(n)$.*

Bollobás conjectured that a similar result holds for k-graphs when $k \geq 4$, namely that the unique cancellative k-graph on n vertices with the most edges is the complete k-partite k-graph with almost equal part sizes. Sidorenko [11] proved Bollobás' conjecture for $k = 4$, and Kleitman and Sidorenko for $k = 5$ (unpublished). Shearer [10] gave an example showing that Bollobás' conjecture is false for $k \geq 10$.

As observed above, the case $k = 3$ of Katona's problem asks for $\text{ex}(n, \mathcal{F})$ where $\mathcal{F} = \{F_4, F_5\}$. Frankl and Füredi [4] sharpened Bollobás' theorem by proving the following:

Theorem 1.3. (Frankl-Füredi) *Let $n \geq 3000$. Then $ex(n, F_5) = s(n)$.*

In [9] a new proof of an asymptotic version of Theorem 1.3 is given. In this paper, we refine those ideas to give a new proof of Bollobás' result, Theorem 1.2 (see Section 3). We also extend the ideas of that proof to improve Theorem 1.3. Our approach has the advantages that it significantly lowers the minimum n for which Theorem 1.3 holds, and characterizes the extremal example.

Theorem 1.4. (Section 4) *Let $n \geq 33$. Then $ex(n, F_5) = s(n)$, with equality only for $S(n)$.*

Using our method, we prove stability versions of both extremal results, showing that any 3-graph that is cancellative, or even just F_5-free, and has nearly $s(n)$ edges, must be approximately tripartite. These results are analogous to the Simonovits stability theorem for graphs, which says that a K_{r+1}-free graph with nearly $t_r(n)$ edges is approximately r-partite (we will state this precisely later).

Theorem 1.5. (Section 5) *For any* $\epsilon > 0$*, there exists* $\delta > 0$ *and* n_0 *such that the following holds: if* H *is a cancellative* 3*-graph on* $n > n_0$ *vertices with at least* $(1-\delta)s(n)$ *edges, then there is a partition of the vertex set of* H *as* $V(H) = U_1 \cup U_2 \cup U_3$ *so that all but at most* ϵn^3 *edges of* H *have one point in each* U_i*.*

Theorem 1.6. (Section 5) *For any* $\epsilon > 0$*, there exists* $\delta > 0$ *and* n_0 *such that the following holds: if* H *is an* F_5*-free* 3*-graph on* $n > n_0$ *vertices with at least* $(1-\delta)s(n)$ *edges, then there is a partition of the vertex set of* H *as* $V(H) = U_1 \cup U_2 \cup U_3$ *so that all but at most* ϵn^3 *edges of* H *have one point in each* U_i*.*

As far as we know, these are the first stability theorems for hypergraphs. As well as being of interest in their own right, it seems that stability theorems may play a valuable role in the development of Turán theory for hypergraphs. The program, established in [7, 8], consists of first proving a stability result for the problem in question, and then refining it to prove a Turán result. It would be interesting to see if any other problems are susceptible to this line of attack.

2 Preliminaries

In a 3-graph H there are two natural analogies to the notion of 'neighborhood' in 2-graphs. The *link* of a vertex $x \in V(H)$ is $L(x) = \{(a,b) : abx \in E(H)\}$. We can think of the link as a graph on $V(H)$. The *degree of* x is $d(x) = |L(x)|$, the number of edges containing x. Given a pair of vertices x, y, their *neighborhood* is $N(x,y) = \{a : axy \in E(H)\}$. We record some simple properties of these definitions in the following useful lemma. Call an edge-colored graph *rainbow* if no two edges have the same color.

Lemma 2.1. *Let* H *be a* 3*-graph, and suppose that* $xyz \in E(H)$*. Let* $W = V(H) - \{x,y,z\}$*.*
(i) If H *contains no* F_4*, then* $L(x)$ *contains no triangle.*
(ii) If H *contains no* F_5*, then* $L(x), L(y)$ *and* $L(z)$ *are edge disjoint graphs when restricted to* W*. If* H *is cancellative they are edge disjoint as graphs on* $V(H)$*.*

When H *contains no* F_5*, we think of* $G = L(x) \cup L(y) \cup L(z)$ *as a* 3*-coloring of a graph on* W*.*
(iii) If H *contains no* F_5*, then any triangle in* G *is either monochromatic or rainbow. If* H *is cancellative then all triangles in* G *are rainbow.*
(iv) If H *is cancellative and* G *contains a* K_4*, then it must be colored so that there are two disjoint edges of each color.*

Proof. (i) If abc is a triangle in $L(x)$, then xab, xac, xbc forms a copy of F_4.
(ii) If a, b are in W with $ab \in L(x) \cap L(y)$, then xab, yab, xyz forms a copy of F_5. If $az \in L(x) \cap L(y)$, then xaz, yaz, xyz forms a copy of F_4.
(iii) If abc is a triangle with ab and bc in $L(x)$ and ac in $L(y)$, then xab, xbc, yac forms a copy of F_5. If H is also F_4-free there are no monochromatic triangles by (1).
(iv) Every triangle is rainbow, so no two incident edges have the same color. In a copy of K_4, this

implies that every color appears at most twice. There are three colors and six edges, so equality holds. □

We conclude this section by deriving some recurrences for the number of edges in $T_3(n)$ and $S(n)$. First we give the approximate formulae,

$$t_3(n) = \left(\frac{2}{3} + o(1)\right)\binom{n}{2}, \qquad s(n) = \left(\frac{2}{9} + o(1)\right)\binom{n}{3}.$$

Now consider both as being defined on the same vertex set, with the same partition X, Y, Z. Let xyz be an edge of $S(n)$. By inclusion/exclusion the number of edges meeting the set $\{x, y, z\}$ is

$$|L(x)| + |L(y)| + |L(z)| - |N(x,y)| - |N(y,z)| - |N(z,x)| + 1. \tag{1}$$

Note that the links $L(x), L(y), L(z)$ partition the edges of $T_3(n)$ so $|L(x)| + |L(y)| + |L(z)| = t_3(n)$. The sets $N(x,y), N(y,z), N(x,z)$ partition the vertex set. Consequently, (1) is $t_3(n) - n + 1$ and we get the identity

$$s(n) = s(n-3) + t_3(n) - n + 1. \tag{2}$$

Finally, every vertex is adjacent to exactly two of x, y, z in $T_3(n)$, so

$$t_3(n) = t_3(n-3) + 2n - 3. \tag{3}$$

3 A new proof of Bollobás' theorem

In this section we will illustrate the application of the ideas of de Caen and Füredi to cancellative hypergraphs by giving a new proof of Theorem 1.2.

Proof of Theorem 1.2: Suppose that H is a cancellative 3-graph with $s(n)$ edges. We will prove by induction on n that H is isomorphic to $S(n)$. Since $S(n)$ is a maximal cancellative 3-graph, this will prove the theorem. The result is obvious for $n \leq 4$, so assume that $n \geq 5$.

If an edge of H meets at most $t_3(n) - n$ edges (including itself), then delete its vertices to form H'. By (2), H' has more than $s(n-3)$ edges. Since H' is cancellative, the induction hypothesis applies to give a contradiction. We may therefore assume that every edge of H meets at least $t_3(n) - n + 1$ edges of H.

Consider an edge $e = x_1x_2x_3$. By Lemma 2.1 (ii) the links $L(x_i)$ form a 3-colored graph G. Suppose G contains a K_4 disjoint from e, which we denote K. By Lemma 2.1 (iv) each color meets K in a matching, so each vertex of K is incident to edges of all three colors. Then every pair from $K \cup e$ belongs to some edge of H, so by Lemma 2.1 (ii), the seven link graphs of the vertices in $K \cup e$ are pairwise disjoint. For each triple the total size of their links is at least the number of edges they are incident to, so averaging over all triples we see that the 7 links contain at least $\lceil \frac{7}{3}(t_3(n) - n + 1) \rceil$ edges. This is greater than $\binom{n}{2}$, a contradiction.

We deduce that G contains no copy of K_4. Applying Turán's theorem, we see that at most $t_3(n-3)$ edges can meet e in exactly one point. The three sets $N(x_i, x_j)$ are pairwise disjoint (or we find F_4) so at most $n-3$ edges meet e in exactly two points. Therefore the number of edges meeting e (including e itself) is at most $t_3(n-3)+(n-3)+1 = t_3(n)-n+1$ by (3). As argued before, we must have equality in all these computations. It follows that $H' = H - \{x_1, x_2, x_3\}$ has $s(n-3)$ edges, so is $S(n-3)$ by induction hypothesis. Also, G restricted to $V(H')$ has $t_3(n-3)$ edges and is K_4-free, so it must be $T_3(n-3)$. Let (X_1, X_2, X_3) be its tripartition. By Lemma 2.1 (iii), all of its triangles are rainbow, and this is only possible when all edges between X_i and X_j have the same color, which by relabelling we can take to be k, where $\{1,2,3\} = \{i,j,k\}$. Finally note that $N(x_i, x_j) \subset X_k \cup x_k$ as H is cancellative, and so $N(x_i, x_j) = X_k \cup x_k$ by the above computations. Therefore $\{X_i \cup x_i\}$ defines a tripartition of H, i.e. H is S_n. □

4 An improvement of the Frankl-Füredi theorem

In this section we prove Theorem 1.4. First we prove an easy lemma on edge colorings of K_8.

Lemma 4.1. *Suppose that K_8 is 4-edge-colored with no monochromatic K_4 and no 2-colored triangle. Then each color class consists of two disjoint triangles and a single isolated edge.*

Proof. The hypothesis implies that each color class yields an equivalence relation on the vertex set. Thus each color class is a union of complete subgraphs of size at most three, and therefore has at most seven edges. Since K_8 has 28 edges, each color class has exactly seven edges, and the result follows. □

It will be useful to consider the following 5 vertex hypergraph, which has the property that each pair of its vertices belongs to an edge. Let $F_{1,4}$ denote the 3-graph with vertex set $v \cup U$, $|U| = 4$, and edge set $\{vuu' : u, u' \in U\}$.

Proof of theorem 1.4: Suppose that $H = (V, E)$ is an n vertex 3-graph that contains no F_5. We will prove by induction on n that $e(H) \le s(n) + h(n)$, where $h(n)$ is a function that is equal to zero for $n \ge 32$ (this technique is sometimes called progressive induction). Consequently for $n \ge 32$, we have $e(H) \le s(n)$. For $n > 32$, we will show that either H is cancellative, or $e(H) < s(n)$. By Theorem 1.2, this implies that if $e(H) = s(n)$ for $n > 32$, then H is $S(n)$.

To define $h(n)$ we first define a function $g(n)$ by $g(0) = g(1) = g(2) = g(3) = 0$, $g(4) = 2$, and

$$\begin{aligned} g(n) = \max\{ & g(n-4) + t_7(n-4) + 2(n-4) + 3 - [s(n) - s(n-4)], \\ & g(n-5) + \binom{n-5}{2} + 4(n-5) + 6 - [s(n) - s(n-5)], \\ & g(n-12) + \binom{n-12}{2} + 14(n-12) + 37 - [s(n) - s(n-12)]\}. \end{aligned}$$

Asymptotically, we have

$$g(n) = \max\left\{ g(n-4) - \frac{3n^2}{189} + O(n),\ g(n-5) - \frac{n^2}{18} + O(n),\ g(n-12) - \frac{5n^2}{3} + O(n) \right\}.$$

Consequently, $g(n)$ is both decreasing and negative for large n. Direct computations yield $g(32) = 0$, $g(n) < 0$ for $n > 32$ and $g(n) - g(m) < 0$ when $n > 32$ and $n \geq m + 4$. Let $h(n) = \max\{g(n), 0\}$. Then
(a) $h(n) = 0$ for $n \geq 32$, and
(b) $h(n) - h(m) \geq g(n) - g(m)$ for $n \geq m + 4$ with strict inequality for $n > 32$.
This implies that

$$h(n) \geq \max\{h(n-4) + t_7(n-4) + 2(n-4) + 3 - [s(n) - s(n-4)], \tag{4}$$

$$h(n-5) + \binom{n-5}{2} + 4(n-5) + 6 - [s(n) - s(n-5)], \tag{5}$$

$$h(n-12) + \binom{n-12}{2} + 14(n-12) + 37 - [s(n) - s(n-12)]\}, \tag{6}$$

and the inequality is strict for $n > 32$.

For $n \leq 4$ any 3-graph has at most $s(n)+h(n)$ edges, so the basis for the induction holds trivially, and we can assume that $n \geq 5$. If H is cancellative, then by Theorem 1.2, $e(H) \leq s(n) \leq s(n)+h(n)$, with equality only for $S(n)$. We may therefore assume that H is not cancellative.

First consider the case that H contains a copy of $F_{1,4}$, denoted $v \cup U$ as before. For $0 \leq i \leq 3$ let f_i be the number of edges in H with i points in $v \cup U$. By the induction hypothesis, $f_0 \leq s(n-5) + h(n-5)$. Note that by definition of $F_{1,4}$ we have $f_3 = 6$. Also, there cannot be an edge of the form xuu' with $u, u' \in U$, or for any other $u'' \in U$ we have a copy of F_5 with edges $xuu', vuu'', vu'u''$. So any edge with 2 points in $v \cup U$ must contain v, giving $f_2 \leq 4(n-5)$. Now every pair of vertices in $F_{1,4}$ belongs to an edge, so Lemma 2.1 (ii) implies that the links of its vertices are edge disjoint, i.e. $f_1 \leq \binom{n-5}{2}$. Therefore by (5) we have

$$e(H) = \sum_0^3 f_i \leq s(n-5) + h(n-5) + \binom{n-5}{2} + 4(n-5) + 6 \leq s(n) + h(n),$$

with strict inequality for $n > 32$, and we are done in this case.

Therefore we can assume that H contains no copy of $F_{1,4}$. Since H is not cancellative, it contains a copy of F_4. Let $S = \{s_1, s_2, s_3, s_4\}$ be its vertex set and $\{s_1s_2s_3, s_1s_2s_4, s_1s_3s_4\}$ its edges. Note that every pair of vertices in S belong to an edge, so by Lemma 2.1 (ii) $G = \cup L(s_i)$ is a simple 4-colored graph on $V - S$.

Case 1: Suppose that G contains no copy of K_8.

Let f_i denote the number of edges of H with i points in S. By induction, we have $f_0 \leq s(n-4) + h(n-4)$. Also $f_1 \leq t_7(n-4)$ by Turán's theorem, and $f_3 = 3$ by definition. To bound f_2, note that for any $t \in V - S$ at most 2 pairs s_is_j can belong to $L(t)$. For suppose there are 3 such pairs. They cannot all contain s_1, as then $S \cup t$ would be a copy of $F_{1,4}$, which gives a case we have already excluded. Any other choice of 3 pairs must contain 2 of the form s_is_j,s_is_k with $i \neq 1$. Writing $\{1, 2, 3, 4\} = \{i, j, k, \ell\}$ we see that $s_js_ks_\ell$ is an edge, and so $ts_is_j, ts_is_k, s_js_ks_\ell$ is a copy of F_5. This contradiction shows that $L(t)$ contains at most 2 pairs from S, and so $f_2 \leq 2(n-5)$.

Consequently, (4) yields

$$e(H) = \sum_0^3 f_i \le s(n-4) + h(n-4) + t_7(n-4) + 2(n-5) + 3 \le s(n) + h(n),$$

with strict inequality for $n > 32$, and we are done in this case.

Case 2: Suppose that G contains a copy of K_8.

Let T denote the vertex set of this copy of K_8, and let f_i denote the number of edges of H with exactly i elements in $S \cup T$. By induction, we have $f_0 \le s(n-12) + h(n-12)$. Consider the 4-edge coloring on T given by the links $L(s_i)$. There can be no monochromatic K_4 (as this gives a copy of $F_{1,4}$) and no 2-colored triangle (this gives a copy of F_5) so the coloring satisfies the hypothesis of Lemma 4.1. It follows that every vertex of T is incident to an edge of each of the four colors, so $S \cup T$ induces a subhypergraph in which every pair of vertices belongs to an edge. Then Lemma 2.1 (ii) implies that the links $L(v)$ are edge-disjoint for $v \in S \cup T$, and so $f_1 \le \binom{n-12}{2}$.

To estimate f_2, we fix $z \notin S \cup T$ and count pairs xy in $S \cup T$ such that xyz is an edge. By the same argument as in Case 1, there are at most 2 such pairs with both x and y in S. For each y in T there is at most one x in S such that xyz is an edge, for if x' is another then yz is in $L(x)$ and $L(x')$, contrary to Lemma 2.1 (ii). Thus there are at most 8 pairs with x in S and y in T. Next note that $L(z)$ restricted to T forms a matching. Otherwise there are edges xyz, $x'yz$ and $s_i xx'$ (where xx' has color i) and this is a copy of F_5. This gives at most 4 pairs with x, y in T. In total there at most 14 pairs xy for each z, so $f_2 \le 14(n-12)$.

Now we bound f_3, the number of edges xyz contained in $S \cup T$. There are 3 edges with x, y, z all in S. There are no edges with x, y in S and z in T. Otherwise Lemma 4.1 shows that there is $w \in T$ with $wz \in L(x)$ and then $w' \in T$ with $ww' \in L(y)$. This gives $xyz, xwz, ww'y$ forming a copy of F_5. Consequently, there are no edges of this form. The edges with two elements of T and one of S we bound simply by $\binom{8}{2} = 28$. Finally we claim that if x, y, z are all in T then xyz is a monochromatic triangle. For suppose it is rainbow, and say xy has color 1. There is some $w \in T - xy$ such that wz has color 1, and then xyz, xys_1, wzs_1 forms F_5. Thus xyz is monochromatic, and by Lemma 4.1, there are at most 6 such edges. In total we have $f_3 \le 37$.

Putting everything together and recalling equation (6) we have

$$e(H) = \sum_0^3 f_i \le s(n-12) + h(n-12) + \binom{n-12}{2} + 14(n-12) + 37 \le s(n) + h(n),$$

with strict inequality for $n > 32$, and the theorem is proved. □

5 Stability theorems

In this section we will show that a cancellative 3-graph with almost as many edges as $S(n)$ looks approximately like $S(n)$. Moreover, we will also show that a F_5-free 3-graph with almost as many edges as $S(n)$ looks approximately like $S(n)$. Such theorems can be thought of as *stability* theorems, after the classical result of Simonovits [13] for graphs. To the best of our knowledge, these are the first such theorems for hypergraphs.

We will need the following slight variation on a case of the Simonovits stability theorem. It differs from the usual formulation, in that the conclusion refers to deleting vertices rather than edges, but this version follows easily from the proof given in [2] pp. 340-342. See also [8] for a version with explicit constants.

Proposition 5.1. *For any $\epsilon' > 0$ there exists $\delta' > 0$ and n_0 such that the following holds: if G is a K_4-free graph on $n > n_0$ vertices with at least $(1-\delta')t_3(n)$ edges, then one can delete $\epsilon' n$ vertices from G so that the remaining graph is tripartite.*

The following theorem is a stability version of Bollobás' theorem.

Theorem 1.5 For any $\epsilon > 0$ there exists $\delta > 0$ and n_0 such that the following holds: if H is a cancellative 3-graph with at least $(1-\delta)s(n)$ edges, then there is a partition of the vertex set of H as $V(H) = U_1 \cup U_2 \cup U_3$ so that all but at most ϵn^3 edges of H have one point in each U_i.

Proof. Note that we may and will assume that ϵ is sufficiently small, as this only strengthens the conclusion we reach. Our constants satisfy the hierarchy $\delta \ll \delta' \ll \epsilon' \ll \epsilon$. More precisely, given $\epsilon > 0$, choose $\epsilon' < (1/2)10^{-7}\epsilon^2$. Fix $\delta' < (1/2)10^{-7}\epsilon^2$ so that Proposition 5.1 applies with ϵ', and let $\delta < 27 \cdot 10^{-10}(\epsilon\delta')^2$.

In the course of the proof we will form a set V_0 of bad vertices, which will always have size at most $\frac{1}{2}\epsilon n$. Our arguments will always apply to H restricted to $V - V_0$, and we will show that we can delete at most $\frac{1}{2}\epsilon n^3$ edges from $V - V_0$ to make it tripartite. Since V_0 is incident to less than $\frac{1}{2}\epsilon n^3$ edges we can delete them and then extend to a tripartition of V by distributing the vertices of V_0 arbitrarily.

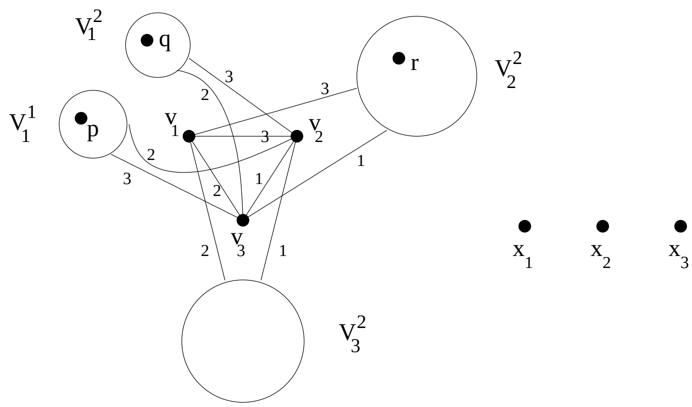
Suppose there are $10^{-5}\epsilon\delta' n$ vertices of degree at most $(1-10^{-4}\epsilon\delta')n^2/9$. Deleting them we arrive at a hypergraph H' on $(1-10^{-5}\epsilon\delta')n$ vertices with at least $(1-\delta-3\cdot 10^{-5}\epsilon\delta'(1-10^{-4}\epsilon\delta'))n^3/27$ edges. The choice of δ implies that

$$1-\delta-3\cdot 10^{-5}\epsilon\delta'(1-10^{-4}\epsilon\delta') > 1-3\cdot 10^{-5}\epsilon\delta' + 3(10^{-5}\epsilon\delta')^2 > (1-10^{-5}\epsilon\delta')^3.$$

Consequently, $e(H') > \frac{1}{27}[(1-10^{-5}\epsilon\delta')n]^3$, which contradicts Bollobás' theorem. We deduce that there are at most $10^{-5}\epsilon\delta' n$ vertices of degree at most $(1-10^{-4}\epsilon\delta')n^2/9$, and we will put them in the bad set V_0.

Now fix some edge $x_1x_2x_3$ in $V - V_0$. By Lemma 2.1 (ii), the 3 links $L(x_i)$ are edge disjoint graphs, and by the previous remark each has at least $(1-10^{-4}\epsilon\delta')n^2/9$ edges. We think of their union as a graph J with at least $(1-10^{-4}\epsilon\delta')n^2/3$ edges, where $L(x_i)$ is colored with color i. As in our proof of Theorem 1.2, we see that J cannot contain a K_4. For together with $x_1x_2x_3$ this would give a 7 vertex subhypergraph in which every pair of vertices is contained in an edge. This gives 7 pairwise disjoint link graphs each with at least $(1-10^{-4}\epsilon\delta')n^2/9$ edges, so their total number of edges is more than $\binom{n}{2}$, which is impossible.

Suppose that J has $10^{-1}\delta' n$ vertices of degree at most $(1-10^{-3}\epsilon)2n/3$. Deleting them we arrive at a graph J' on $(1-10^{-1}\delta')n$ vertices with at least $(1-10^{-4}\epsilon\delta'-2\cdot 10^{-1}\delta'(1-10^{-3}\epsilon))n^2/3$



Figure

edges. Since $\delta' < 10^{-2}\epsilon$ we have $e(J') > \frac{1}{3}[(1-10^{-1}\delta')n]^2$. But $J' \subset J$ is K_4-free, so this contradicts Turán's theorem. We deduce that there are at most $10^{-1}\delta' n$ vertices of degree at most $(1-10^{-3}\epsilon)2n/3$ in J, and we will put them in the bad set V_0.

Since J has at least $(1-10^{-4}\epsilon\delta' - 2\cdot 10^{-1}\delta'(1-10^{-3}\epsilon))n^2/3$ remaining edges, and this is trivially at least $(1-\delta')n^2/3$, we can apply Proposition 5.1. This shows that we can add at most $\epsilon' n$ vertices to V_0, and then partition $V - V_0$ into 3 sets V_1,V_2,V_3 each containing no edges of J. Then J restricted to $V_1 \cup V_2 \cup V_3$ has at least $(1-\delta'-\epsilon')n^2/3 > (1-10^{-7}\epsilon^2)n^2/3$ edges (by the choice of δ', ϵ').

Note that $|V_i - n/3| < 10^{-3}\epsilon n$ for each i, or J would have at most

$$|V_i|(n-|V_i|) + (n-|V_i|)^2/4 = \frac{1}{3}n^2 - \frac{1}{12}(3|V_i|-n)^2 < \frac{1}{3}n^2 - \frac{3}{4}10^{-6}\epsilon^2 n^2 < (1-10^{-7}\epsilon^2)n^2/3$$

edges which is impossible. Since each vertex in $V - V_0$ has degree at least $(1-10^{-3}\epsilon)2n/3$, we see that each vertex v_i in V_i has degree at least $(1-10^{-3}\epsilon)2n/3 - (1/3+10^{-3}\epsilon)n > n/3 - 10^{-2}\epsilon n$ in both V_j, $j \neq i$.

Let $v_1v_2v_3$ be a triangle in J with v_i in V_i. For each V_i add any vertex that is not adjacent to both v_j, $j \neq i$ to V_0. There are at most $6\cdot 10^{-2}\epsilon n$ such vertices. By Lemma 2.1 (iii) all triangles of J are multicolored, so we can suppose v_iv_j has color k, where $\{i,j,k\} = \{1,2,3\}$. Then each vertex of V_k is joined to the vertices v_i, v_j by one edge of color i and one of color j. Let V_k^1 consist of those vertices v in V_k for which vv_i has color i and vv_j has color j (so the color matches the subscript), and $V_k^2 = V_k - V_k^1$.

All edges from v_1 to $V_2^1 \cup V_3^1$ have color 1. Therefore there are no edges between V_2^1 and V_3^1, and the same holds between V_i^1 and V_j^1 for any two distinct $i,j \in \{1,2,3\}$. If both V_i^1 and V_j^1 have size at least $10^{-2}\epsilon n$, then J has at most $n^2/3 - (10^{-2}\epsilon n)^2 < (1-10^{-7}\epsilon^2)n^2/3$ edges, which is impossible. It follows that there is at most one l for which $|V_l^1| \geq 10^{-2}\epsilon n$. Without loss of generality we assume that $l = 1$. Thus both V_2^1 and V_3^1 have size at most $10^{-2}\epsilon n$, and we add their vertices to V_0.

Now take any edge pqr of H in $V - V_0$. Note that it cannot have 2 of its vertices in one of the sets V_1^1, V_1^2, V_2^2 or V_3^2. For example, if p and q both belonged to V_1^1 then x_2v_2 belongs to both the links $L(p)$ and $L(q)$, which contradicts Lemma 2.1 (ii). The other cases are similar. Next suppose that we have $p \in V_1^1$, $q \in V_1^2$ and $r \in V_2^2$. This situation is illustrated in the Figure above (the numbers on the edges are their colors). Note that qv_3 has color 2 and rv_3 has color 1, so if qr is an edge of J it must have color 3. But pv_3 has color 3, i.e. x_3pv_3 is an edge, and qr belongs to $L(p)$ and $L(x_3)$, which contradicts Lemma 2.1 (ii). We deduce that qr is not an edge of J. Since J has at least $(1-10^{-7}\epsilon^2)n^2/3$ edges respecting the partition (V_1, V_2, V_3) out of at most $n^2/3$ possible edges, there are at most $10^{-7}\epsilon^2 n^2/3$ choices for qr, so at most $10^{-7}\epsilon^2 n^3/3$ such edges pqr.

Similarly there are at most $10^{-7}\epsilon^2 n^3/3$ edges pqr with $p \in V_1^1$, $q \in V_1^2$ and $r \in V_3^2$. These edges are exceptional; all others have one point in each of $V_1^1 \cup V_1^2$, V_2^2 and V_3^2. Define a tripartition $V = U_1 \cup U_2 \cup U_3$ so that $V_1^1 \cup V_1^2 \subset U_1$, $V_2^2 \subset U_2$, $V_3^2 \subset U_3$ and the bad vertices V_0 are distributed arbitrarily. Since $|V_0| < \frac{1}{2}\epsilon n$ and there are less than $\frac{1}{2}\epsilon n^3$ exceptional edges we see that all but at most ϵn^3 edges of H have one point in each U_i, so the theorem is proved. □

Remark: It is possible that Theorem 1.5 can also be proved by extending the ideas of Bollobás' original proof of Theorem 1.2. Our framework seems more general however, since the same approach gives a proof of Theorem 1.6 below. As far as we can tell, there is no straightforward modification of Bollobás' original proof that yields this stronger result.

Now we use the preceding theorem to prove a stability version of the Frankl-Füredi theorem.

Theorem 1.6 For any $\epsilon > 0$ there is a $\delta > 0$ and an n_0 such that the following holds: if H is an F_5-free 3-graph on $n > n_0$ vertices with at least $(1-\delta)s(n)$ edges, then there is a partition of the vertex set of H as $V(H) = U_1 \cup U_2 \cup U_3$ so that all but at most ϵn^3 edges of H have one point in each U_i.

Proof. Again we can assume ϵ sufficiently small. By Theorem 1.5 we can choose $\delta < \epsilon/100$ so that any cancellative 3-graph on n' vertices with at least $(1-\delta)s(n')$ edges has a tripartition with at most $\frac{1}{2}\epsilon n'^3$ wrong edges.

Suppose that there are $\frac{1}{2}\epsilon n$ vertices of degree at most $(1-10^{-2})n^2/9$. Deleting them we arrive at a hypergraph H' on $(1-\epsilon/2)n$ vertices with at least $(1-\delta-3\cdot\epsilon/2\cdot(1-10^{-2}))n^3/27 > \frac{1}{27}[(1-\epsilon/2)n]^3$ edges, which contradicts Theorem 1.4. Consequently, there are at most $\frac{1}{2}\epsilon n$ vertices of degree at most $(1-10^{-2})n^2/9$, and we denote them V_0.

Set $n' = |V - V_0|$. Then H restricted to $V - V_0$ has at least $(1-\delta)s(n')$ edges. If H contains no copy of F_4, then it is cancellative, and by choice of δ there is a partition $V(H) = U_1 \cup U_2 \cup U_3$ so that all but at most $\frac{1}{2}\epsilon n^3$ edges of H have one point in each U_i. Since $|V_0| < \frac{1}{2}\epsilon n$, we can distribute the vertices of V_0 arbitrarily to obtain the required partition. Therefore we can assume that H contains a copy of F_4.

Now we argue as in our proof of Theorem 1.4. Let $S = s_1s_2s_3s_4$ be the vertex set of this F_4. Then $\cup L(s_i)$ is a simple 4-colored graph on $V - V_0 - S$ with at least $(1-10^{-2})4n^2/9$ edges, so it contains a copy of K_8. Again, by Lemma 4.1 this K_8 either contains a monochromatic K_4 or each of its vertices is incident to all 4 colors. Thus we find a set X with $|X| = 5$ or $|X| = 12$ that induces a subhypergraph in which every pair of vertices belongs to an edge. Then the links of its vertices form edge disjoint graphs on $V - V_0 - S$ each with at least $(1 - 10^{-2})n^2/9$ edges. This gives a total of more than $\binom{n}{2}$, which is a contradiction, so the theorem is proved. $\square$

References

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

[2] B. Bollobás, **Extremal graph theory**, London Mathematical Society Monographs, 11., Academic Press, Inc., London-New York, 1978.

[3] D. de Caen, Z. Füredi, The maximum size of 3-uniform hypergraphs not containing a Fano plane, *J. Combin. Theory Ser. B* 78 (2000), no. 2, 274–276.

[4] P. Frankl, Z. Füredi, A new generalization of the Erdős-Ko-Rado theorem, *Combinatorica* 3 (1983), no. 3-4, 341–349.

[5] Z. Füredi, Turán type problems, in: *Surveys in combinatorics*, London Math. Soc. Lecture Note Ser. 166, Cambridge Univ. Press, Cambridge, 1991, 253–300

[6] Z. Füredi, M. Simonovits, manuscript.

[7] P. Keevash, B. Sudakov, The exact Turán number of the Fano plane, submitted.

[8] P. Keevash, B. Sudakov, On a hypergraph Turán problem of Frankl, submitted.

[9] D. Mubayi, V. Rödl, On the Turán number of triple systems, *J. Combin. Theory Ser. A* 100, (2002), 136–152.

[10] J. B. Shearer, A new construction for cancellative families of sets, *Electron. J. Combin.* 3 (1996), no. 1, Research Paper 15, approx. 3 pp. (electronic).

[11] A. F. Sidorenko, On the maximal number of edges in a homogeneous hypergraph that does not contain prohibited subgraphs. *Mat. Zametki* 41 (1987), no. 3, 433–455.

[12] A. Sidorenko, What we know and what we do not know about Turán numbers, *Graphs and Combinatorics* 11 (1995), 179–199.

[13] M. Simonovits, A method for solving extremal problems in graph theory, stability problems. 1968 Theory of Graphs (Proc. Colloq., Tihany, 1966) pp. 279–319, Academic Press, New York

[14] P. Turán, Eine Extremalaufgabe aus der Graphentheorie, *Mat. Fiz. Lapok* 48 (1941) 436–452.