Set systems without a strong simplex

Tao Jiang [*] Oleg Pikhurko [†] Zelealem Yilma [‡]

May 31, 2009

Abstract

A *d-simplex* is a collection of $d+1$ sets such that every d of them have non-empty intersection and the intersection of all of them is empty. A *strong d-simplex* is a collection of $d+2$ sets $A, A_1, \ldots, A_{d+1}$ such that $\{A_1, \ldots, A_{d+1}\}$ is a d-simplex, while A contains an element of $\cap_{j \neq i} A_j$ for each i, $1 \leq i \leq d+1$.

Mubayi and Ramadurai [12] conjectured that if $k \geq d+1 \geq 3$, $n > k(d+1)/d$, and $\mathcal{F}$ a family of k-element subsets of an n-element set that contains no strong d-simplex, then $|\mathcal{F}| \leq \binom{n-1}{k-1}$ with equality only when $\mathcal{F}$ is a star. We prove their conjecture when $k \geq d+2$ and n is large. This also yields a shorter proof of a result of Frankl and Füredi [8] when $k \geq d+2$ and n is large. We also present some results for the case $k = d+1$, including completely proving the conjecture for $k = 3$ and all n.

1 Introduction

For any integer $k \geq 2$, we denote the family of all k-element subsets of $[n] := \{1, \ldots, n\}$ by $\binom{[n]}{k}$. A family $\mathcal{F}$ of sets is a *star* if there exists an element x that lies in all the members in $\mathcal{F}$. We say $\mathcal{F}$ is an *intersecting family* if every two of its members have nonempty intersection. We use $|\mathcal{F}|$ to denote the cardinality of $\mathcal{F}$, i.e. the number of members in $\mathcal{F}$.

The following is one of the most important results in extremal combinatorics.

Theorem 1.1 (Erdős, Ko and Rado [5]) *Let $n \geq 2k$ and $\mathcal{F} \subseteq \binom{[n]}{k}$ be an intersecting family. Then $|\mathcal{F}| \leq \binom{n-1}{k-1}$. If $n > 2k$ and equality holds, then $\mathcal{F}$ is a star.*

The forbidden configuration in Theorem 1.1 consists of a pair of disjoint sets. A generalization of this configuration, with geometric motivation, is as follows.

---

[*] Research partially supported by the National Security Agency under Grant Number H98230-07-1-027. Department of Mathematics and Statistics, Miami University, Oxford, OH 45056, USA. Email: `jiangt@muohio.edu`.

[†] Research partially supported by the National Science Foundation under Grant Number DMS-0758057. Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA. URL: `http://www.math.cmu.edu/~pikhurko`.

[‡] Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA. E-mail: `zyilma@andrew.cmu.edu`.

Definition 1.2 *Fix $d \geq 1$. A family of sets is d-*intersecting *if every d of its members have non-empty intersection. A collection of $d+1$ sets $A_1, A_2, \ldots, A_{d+1}$ is a d-*dimensional simplex *(or a d-*simplex*) if it is d-intersecting but not $(d+1)$-intersecting (that is, $\cap_{i=1}^{d+1} A_i = \emptyset$).*

Note that a 1-simplex is pair of disjoint edges, and Theorem 1.1 states that if $\mathcal{F} \subseteq \binom{[n]}{k}$ with $n \geq 2k$ and $|\mathcal{F}| > \binom{n-1}{k-1}$, then $\mathcal{F}$ contains a 1-simplex. In general, it is conjectured that the same threshold for $\mathcal{F}$ guarantees a d-simplex for every d, $1 \leq d \leq k-1$. For $d = 2$ this was a question of Erdős [4] while the general conjecture was formulated by Chvátal:

Conjecture 1.3 (Chvátal [2]) *Suppose that $k \geq d+1 \geq 2$ and $n \geq k(d+1)/d$. If $\mathcal{F} \subseteq \binom{[n]}{k}$ contains no d-simplex, then $|\mathcal{F}| \leq \binom{n-1}{k-1}$. Equality holds only if $\mathcal{F}$ is a star.*

Another motivation, see [2, Page 358], is that when we formally let $d = k$, then we obtain the famous open problem of finding the Turán function of the hypergraph $\binom{[k+1]}{k}$, posed by Turán [15] in 1941.

Various partial results on the case $d = 2$ of the conjecture have been obtained in [1, 2, 3, 6, 7] until this case was completely settled by Mubayi and Verstraëte [13]. Conjecture 1.3 has been proved by Frankl and Füredi [8] for every fixed k, d if n is sufficiently large. Keevash and Mubayi [10] have also proved the conjecture when k/n and $n/2 - k$ are both bounded away from zero.

Mubayi [11] proved a stability result for the $d = 2$ case of Conjecture 1.3 and conjectured that a similar result holds for larger d.

Conjecture 1.4 (Mubayi [11]) *Fix $k \geq d+1 \geq 3$. For every $\delta > 0$, there exist $\epsilon > 0$ and $n_0 = n_0(\epsilon, k)$ such that the following holds for all $n > n_0$. If $\mathcal{F} \subseteq \binom{[n]}{k}$ contains no d-simplex and $|\mathcal{F}| > (1-\epsilon)\binom{n-1}{k-1}$, then there exists an $S \subseteq [n]$ with $|S| = n-1$ such that $|\mathcal{F} \cap \binom{S}{k}| < \delta\binom{n-1}{k-1}$.*

Subsequently, Mubayi and Ramadurai [12] proved Conjecture 1.4 in a stronger form except in the case $k = d+1$, as follows.

Definition 1.5 *Fix $d \geq 1$. A collection of $d+2$ sets $A, A_1, A_2, \ldots, A_{d+1}$ is a* strong d-simplex *if $\{A_1, A_2, \ldots, A_{d+1}\}$ is a d-simplex, and A contains an element of $\cap_{j \neq i} A_j$ for each $i \in [d+1]$.*

Note that a strong 1-simplex is a collection of three sets A, B, C such that $A \cap B$ and $B \cap C$ are non-empty, and $A \cap C$ is empty. Note also that if a family $\mathcal{F}$ contains no strong d-simplex, then certainly it contains no d-simplex (but not vice versa). The main result of Mubayi and Ramadurai [12] can be formulated using asymptotic notation as follows, where $o(1) \to 0$ as $n \to \infty$.

Theorem 1.6 (Mubayi and Ramadurai [12]) *Fix $k \geq d+2 \geq 3$. Let $\mathcal{F} \subseteq \binom{[n]}{k}$ contain no strong d-simplex. If $|\mathcal{F}| \geq (1-o(1))\binom{n-1}{k-1}$, then there exists an element $x \in [n]$ such that the number of sets of $\mathcal{F}$ omitting x is $o(n^{k-1})$.*

Corollary 1.7 (Mubayi and Ramadurai [12]) *Fix* $k \geq d+2 \geq 3$. *Let* $\mathcal{F} \subseteq \binom{[n]}{k}$ *contain no strong* d*-simplex. Then* $|\mathcal{F}| \leq (1+o(1))\binom{n-1}{k-1}$ *as* $n \to \infty$.

In [10], a similar stability result was proved when k/n and $n/2-k$ are both bounded away from 0 and the result was used to settle Conjecture 1.3 in this range of n.

Let us describe our contribution. First, we observe that Theorem 1.6 does not hold when $k = d+1$.

Proposition 1.8 *Let* $k = d+1 \geq 3$. *For every* $\epsilon > 0$ *there is* n_0 *such that for all* $n \geq n_0$ *there is a* k*-graph* $\mathcal{F}$ *with* n *vertices and at least* $(1-\epsilon)\binom{n-1}{k-1}$ *edges without a strong* d*-simplex such that every vertex contains at most* ϵn^{k-1} *edges of* $\mathcal{F}$.

The authors of [12] pointed that that they were unable to use Theorem 1.6 to prove the corresponding exact result for large n (which would give a new proof of the result of Frankl and Füredi [8]). They subsequently made the following conjecture, which is a strengthening of Chvátal's conjecture.

Conjecture 1.9 (Mubayi and Ramadurai [12]) *Let* $k \geq d+1 \geq 3$, $n > k(d+1)/d$ *and* $\mathcal{F} \subseteq \binom{[n]}{k}$ *contain no strong* d*-simplex. Then* $|\mathcal{F}| \leq \binom{n-1}{k-1}$ *with equality only for a star.*

In Section 4, we will prove Conjecture 1.9 for all fixed $k \geq d+2 \geq 3$ and large n:

Theorem 1.10 *Let* $k \geq d+2 \geq 3$ *and* n *be sufficiently large. If* $\mathcal{F} \subseteq \binom{[n]}{k}$ *contain no strong* d*-simplex, then* $|\mathcal{F}| \leq \binom{n-1}{k-1}$ *with equality only for a star.*

The case $k = d+1$ is very interesting (partly because Proposition 1.8 shows that there are almost extremal configurations very different from a star). In Section 5 we present a short proof of the following (weaker) bound.

Theorem 1.11 *Let* $k+1 = d$ *and* $n \geq k$. *The maximum size of* k*-graph* $\mathcal{F} \subseteq \binom{[n]}{k}$ *without a strong* d*-star is at most* $\binom{n}{k-1}$.

Working harder on the case $k = d+1 = 3$, we were able to solve it completely for all n in Section 6.

Theorem 1.12 *Let* $k = 3$, $d = 2$, *and* $n \neq 4$. *The maximum size of* k*-graph* $\mathcal{F} \subseteq \binom{[n]}{k}$ *without a strong* d*-simplex is at most* $\binom{n-1}{k-1}$ *with equality only for a star. If* $n = 4$, *then the maximum size of* $\mathcal{F}$ *is* 4 *with equality only for* $\mathcal{F} = \binom{[4]}{3}$.

Independently of us, Füredi and Özkahya [9] have re-proved our Theorem 1.10 in a stronger form (but for the same set of parameters: $k \geq d+2$ and $n \geq n_0(k)$). Namely, they can additionally guarantee that (in the notation of Definition 1.5) the sets $A_1 \setminus A, \ldots, A_{d+1} \setminus A$ are pairwise disjoint, while the sets $A \setminus A_i, \ldots, A \setminus A_{d+1}$ partition A and have any specified non-zero sizes. Füredi and Özkahya's proof, which uses the Sunflower Lemma, is different from ours.

The problem of forbidding a d-simplex where we put some extra restrictions on the sizes of certain boolean combinations of edges has also been studied before, with one particularly interesting paper being that of Csákány and Kahn [3] who use a homological approach.

Frankl and Füredi's proof [8] of Chvátal's conjecture for a d-simplex for large n is very complicated. Together with the stability result in [12], we obtained a simpler proof of a stronger result. One key factor seems be that having a special edge A in a strong d-simplex $\{A, A_1, \ldots, A_{d+1}\}$ that contains an element in every d-wise intersection in the d-simplex $\{A_1 \ldots, A_{d+1}\}$ facilitates induction arguments very nicely. This observation, already made in [12], further justifies the interest in strong d-simplices.

2 Some Notation and Conventions

As is usual in the literature, a collection $\mathcal{F}$ of k-element subsets of a set V is also called a *k-uniform hypergraph* on V, where elements of V are called *vertices* (or *points*) and members of $\mathcal{F}$ are called *hyperedges* (or simply *edges*). We usually identify (hyper)graphs with their edge sets; thus, for example, $|\mathcal{F}|$ denotes the number of edges of $\mathcal{F}$.

Let $\mathcal{F} \subseteq \binom{[n]}{k}$. Recall that a strong d-simplex L in $\mathcal{F}$ consists of $d+2$ hyperedges $A, A_1, A_2, \ldots, A_{d+1}$ such that every d of $A_1, \ldots, A_{d+1}$ have non-empty intersection but $\cap_{i=1}^{d+1} A_i = \emptyset$. Furthermore, A contains an element of $\cap_{j \neq i} A_i$ for each $i \in [d+1]$. This means that we can find some $d+1$ elements $v_1, v_2, \ldots, v_{d+1}$ in A such that for each $i \in [d+1]$ $v_i \in \cap_{j \neq i} A_j$. Note that $v_1, v_2, \ldots, v_{d+1}$ are distinct because no element lies in all of $A_1, \ldots, A_{d+1}$. We call A the *special edge* for L and the set $\{v_1, v_2, \ldots, v_{d+1}\}$ a *special* $(d+1)$*-tuple* for L. (Note that there may be more than one choice of a special $(d+1)$-tuple.)

As usual, the *degree* $d_{\mathcal{F}}(x)$ (or simply $d(x)$) of a vertex x in $\mathcal{F}$ is the number of hyperedges that contain x. For a positive integer p, the *p-shadow* of $\mathcal{F}$ is defined as

$$\Delta_p(\mathcal{F}) = \{S \subseteq [n] : |S| = p, S \subseteq D \text{ for some } D \in \mathcal{F}\}.$$

Also, we let

$$\mathcal{T}_{p+1}(\mathcal{F}) = \{T : T \text{ is a special } (p+1)\text{-tuple for some strong } p\text{-simplex in } \mathcal{F}\}.$$

For each $p \in [k-2]$, let

$$\partial_p^*(\mathcal{F}) = |\Delta_p(\mathcal{F})| + |\mathcal{T}_{p+1}(\mathcal{F})|.$$

Given a vertex x in a hypergraph $\mathcal{F}$, let

$$\begin{aligned} \mathcal{F} - x &= \{D : D \in \mathcal{F}, x \notin D\}, \\ \mathcal{F}_x &= \{D \setminus \{x\} : D \in \mathcal{F}, x \in D\}, \\ \mathcal{F}(x) &= \{D \in \mathcal{F} : D \ni x\}. \end{aligned}$$

For brevity's sake, we will often write ab and abc for sets $\{a, b\}$ and $\{a, b, c\}$, etc.

3 Proof of Proposition 1.8

We have to show that if $k = d+1$, then there is no stability. Let $\epsilon > 0$ be given. Choose large m such that $\binom{m-1}{k-1} > (1 - \epsilon/2)\binom{m}{k-1}$. Let the complete star $\mathcal{H} \subseteq \binom{[m]}{k}$ consist of all k-tuples containing 1. Clearly, $\mathcal{H}$ has no $(k-1)$-simplex. Let $n \to \infty$. A result of Rödl [14] shows that we can find an m-graph $\mathcal{F} \subseteq \binom{n}{m}$ with at least $(1-\epsilon/2)\binom{n}{k-1}/\binom{m}{k-1}$ edges such that every two edges of $\mathcal{F}$ intersect in at most $k-2$ vertices. Replace every edge of $\mathcal{F}$ by a copy of the star $\mathcal{H}$.

The obtained k-graph $\mathcal{G}$ on $[n]$ has no strong $(k-1)$-simplex S. Indeed the special k-set X of S intersects every other edge of S in $k-1$ vertices; thus if X belong to some copy of the star $\mathcal{H}$, then every other edge of S belongs to the same copy, a contradiction.

The size of $\mathcal{G}$ is at least $(1-\epsilon/2)\binom{n}{k-1}/\binom{m}{k-1} \times (1-\epsilon/2)\binom{m}{k-1} > (1-\epsilon)\binom{n}{k-1}$. Also, by the packing property of $\mathcal{F}$, the number of edges of $\mathcal{G}$ containing any one vertex is at most $\binom{n-1}{k-2}/\binom{m-1}{k-2} \times \binom{m-1}{k-1} < \epsilon n^{k-1}$ when n is large. This establishes Proposition 1.8.

4 Proof of Theorem 1.10

In order to prove Theorem 1.10, we first establish a general lower bound on $\partial_p^*(\mathcal{F})$ in Theorem 4.5, which is of independent interest. Then we will use Theorem 4.5 to prove Theorem 1.10.

We need several auxiliary lemmas. The first one follows readily from Corollary 1.7.

Lemma 4.1 *For each $k \geq d+2 \geq 3$, there exists an integer $n_{k,d}$ such that for all integers $n \geq n_{k,d}$ if $\mathcal{H} \subseteq \binom{[n]}{k}$ contains no strong d-simplex then $|\mathcal{H}| \leq 2\binom{n-1}{k-1}$.*

Lemma 4.2 *For every $k \geq p+2 \geq 3$, there exists a positive constant $\beta_{k,p}$ such that the following holds.*

Let $n_{k,p}$ be defined as in Lemma 4.1. Let $\mathcal{H}$ be a k-uniform hypergraph with $n \geq n_{k,p}$ vertices and $m > 4\binom{n-1}{k-1}$ edges. Then $|\mathcal{T}_{p+1}(\mathcal{H})| \geq \beta_{k,p} m^{\frac{p}{k-1}}$.

Proof. First note that $n < \lambda_k m^{\frac{1}{k-1}}$ for some constant λ_k depending only on k. Since $m > 4\binom{n-1}{k-1}$ and $n \geq n_{k,p}$, by Lemma 4.1, $\mathcal{H}$ contains a strong p-simplex L with A being its special edge. Let us remove the edge A from $\mathcal{H}$. As long as $\mathcal{H}$ still has more than $m/2 > 2\binom{n-1}{k-1}$ edges left we can find another strong p-simplex and remove its special edge from the hypergraph. We can repeat this at least $m/2$ times. This produces at least $m/2$ different special edges. Each special edge contains a special $(p+1)$-tuple. Each special $(p+1)$-tuple is clearly contained in at most $\binom{n-p-1}{k-p-1}$ special edges. So the number of distinct $(p+1)$-tuples $\mathcal{T}_{p+1}(\mathcal{H})$ in $\mathcal{H}$ is at least $m/2\binom{n-p-1}{k-p-1}$. Using $n < \lambda_k m^{\frac{1}{k-1}}$, we get $|\mathcal{T}_{p+1}(\mathcal{H})| \geq \beta_{k,p} m^{\frac{p}{k-1}}$, for some small positive constant $\beta_{k,p}$ depending on k and p only. ■

The next lemma is a key step to our proof of Theorem 4.5.

Lemma 4.3 *Let $\mathcal{F}$ be a k-uniform hypergraph and x a vertex in $\mathcal{F}$. Suppose that $T \in \mathcal{T}_p(\mathcal{F}_x) \cap \Delta_p(\mathcal{F} - x)$. Then $T \cup \{x\} \in \mathcal{T}_{p+1}(\mathcal{F})$.*

Proof. Note that $\mathcal{F}_x$ is $(k-1)$-uniform. By our assumption, T is a special p-tuple for some strong $(p-1)$-simplex $L = \{A, A_1, \ldots, A_p\}$ in $\mathcal{F}_x$, where A is the special edge and $A \supseteq T$. By definition, $\{A_1, \ldots, A_p\}$ is $(p-1)$-intersecting but $\cap_{i=1}^p A_i = \emptyset$. Suppose that

$T = \{v_1, \ldots, v_p\}$, where for each $i \in [p]$ we have $v_i \in \cap_{j \neq i} A_j$. Since $T \in \Delta_p(\mathcal{F} - x)$, there exists $D \subseteq \mathcal{F} - x$ such that $T \subseteq D$.

For each $i \in [p]$, let $A'_i = A_i \cup \{x\}$. Let $A'_{p+1} = D$ and $A' = A \cup \{x\}$. Let $L' = \{A', A'_1, \ldots, A'_{p+1}\} \subseteq \mathcal{F}$. We claim that $\{A'_1, \ldots, A'_{p+1}\}$ is a p-simplex in $\mathcal{F}$. Indeed, $x \in \cap_{i=1}^p A'_i$. Also, for each $i \in [p]$, $v_i \in \cap_{j \in [p+1] \setminus \{i\}} A'_j$. So, $\{A'_1, \ldots, A'_{p+1}\}$ is p-intersecting. Since $\cap_{i=1}^p A_i = \emptyset$ the only element in $\cap_{i=1}^p A'_i$ is x. But $x \notin D$ since $D \in \mathcal{F} - x$. So $\cap_{i=1}^{p+1} A'_i = \emptyset$. This shows that $\{A'_1, \ldots, A'_{p+1}\}$ is a p-simplex in $\mathcal{F}$.

Now, let $T' = T \cup \{x\} = \{x, v_1, \ldots, v_p\}$. Then A' contains T'. Let $v_{p+1} = x$. For all $i \in [p+1]$ we have $v_i \in \cap_{j \in [p+1] \setminus \{i\}} A'_j$. Since A' contains $v_1, \ldots, v_{p+1}$, L' is a strong p-simplex in $\mathcal{F}$ with T' being a special $(p+1)$-tuple. That is, $T' \in \mathcal{T}_{p+1}(\mathcal{F})$. ■

Lemma 4.4 *Let $k > j \geq 2$. Let $\mathcal{F}$ be a k-graph and let x be a vertex of $\mathcal{F}$. Then*

$$\partial_j^*(\mathcal{F}) \geq \partial_j^*(\mathcal{F} - x) + \partial_{j-1}^*(\mathcal{F}_x).$$

Proof. We want to prove that

$$|\Delta_j(\mathcal{F})| + |\mathcal{T}_{j+1}(\mathcal{F})| \geq |\Delta_j(\mathcal{F} - x)| + |\mathcal{T}_{j+1}(\mathcal{F} - x)| + |\Delta_{j-1}(\mathcal{F}_x)| + |\mathcal{T}_j(\mathcal{F}_x)|. \tag{1}$$

Let $T \in \mathcal{T}_j(\mathcal{F}_x)$; that is, T is a special j-tuple in $\mathcal{F}_x$. If $T \in \Delta_j(\mathcal{F} - x)$, we say that T is of Type 1. If $T \notin \Delta_j(\mathcal{F} - x)$, we say T is of Type 2. Suppose $\mathcal{T}_j(\mathcal{F}_x)$ contains a Type 1 special j-tuples and b Type 2 special j-tuples. Then $a + b = |\mathcal{T}_j(\mathcal{F}_x)|$.

For each Type 1 special j-tuple T of $\mathcal{F}_x$, by Lemma 4.3, $T \cup \{x\} \in \mathcal{T}_{j+1}(\mathcal{F})$. Furthermore, it is not in $\mathcal{T}_{j+1}(\mathcal{F} - x)$ since $T \cup \{x\}$ contains x. Hence

$$|\mathcal{T}_{j+1}(\mathcal{F})| \geq |\mathcal{T}_{j+1}(\mathcal{F} - x)| + a. \tag{2}$$

For each Type 2 special j-tuple T of $\mathcal{F}_x$, we have $T \in \Delta_j(\mathcal{F})$ since T is contained in some special edge in $\mathcal{F}_x$ which in turn is contained in some edge of $\mathcal{F}$. Also, by our assumption about Type 2 special tuples, $T \notin \Delta_j(\mathcal{F} - x)$. Furthermore, T is not of the form $S \cup \{x\}$ since it does not contain x. Also, for each $S \in \Delta_{j-1}(\mathcal{F}_x)$, $S \cup \{x\}$ is an element in $\Delta_j(\mathcal{F})$ that is not in $\Delta_j(\mathcal{F} - x)$. Hence,

$$|\Delta_j(\mathcal{F})| \geq |\Delta_j(\mathcal{F} - x)| + |\Delta_{j-1}(\mathcal{F}_x)| + b. \tag{3}$$

When we add (2) and (3), we obtain the desired inequality (1) completing the proof of Lemma 4.4. ■

Theorem 4.5 *For all $k \geq p + 2 \geq 3$, there exists a positive constant $c_{k,p}$ such that the following holds: if $\mathcal{F}$ is a k-uniform hypergraph and $m = |\mathcal{F}|$ then $\partial_p^*(\mathcal{F}) \geq c_{k,p} m^{\frac{p}{k-1}}$.*

Proof. Let us remove all isolated vertices from $\mathcal{F}$. Let n denote the number of remaining (i.e. non-isolated) vertices of $\mathcal{F}$. Let $n_{k,p}$ be defined as in Lemma 4.1, which depends only on k and p. Suppose that $n < n_{k,p}$. Since clearly $m \leq \binom{n}{k} < \binom{n_{k,p}}{k}$, $m^{\frac{p}{k-1}}$ is upper bounded by some function of k and p. Hence, $\partial_p^*(\mathcal{F}) \geq \alpha_{k,p} m^{\frac{p}{k-1}}$, for some small enough constant $\alpha_{k,p}$. So, as long as we choose $c_{k,p}$ so that $c_{k,p} \leq \alpha_{k,p}$, the claim holds when

$n \leq n_{k,p}$. To prove the general claim, we use induction on p. For each fixed p, we use induction on n noting that when $n \leq n_{k,p}$ the claim has already been verified.

For the basis step, let $p = 1$. Let $c_{k,1} = \min\{\alpha_{k,1}, \beta_{k,1}, 1/4\}$, where $\beta_{k,1}$ is defined in Lemma 4.2. First, suppose that $m \leq 4\binom{n-1}{k-1} < 4n^{k-1}$. Then $n > (m/4)^{\frac{1}{k-1}} > m^{\frac{1}{k-1}}/4$. We have $\partial_1^*(\mathcal{F}) \geq |\Delta_1(\mathcal{F})| = n \geq c_{k,1}m^{\frac{1}{k-1}}$.

Next, suppose that $m > 4\binom{n-1}{k-1}$. By Lemma 4.2, $\partial_1^*(\mathcal{F}) \geq |\mathcal{T}_2(\mathcal{F})| \geq \beta_{k,1}m^{\frac{1}{k-1}} \geq c_{k,1}m^{\frac{1}{k-1}}$. This completes the proof of the basis step.

For the induction step, let $j \geq 2$. Suppose the claim holds for $p < j$. We prove the claim for $p = j$. We use induction on n. Let

$$c_{k,j} = \min\left\{\alpha_{k,j},\, \beta_{k,j},\, \frac{1}{4k}\left(\frac{1}{2}\right)^{\frac{k-j-1}{k-1}} c_{k-1,j-1}\right\}.$$

Suppose the claim has been verified for k-uniform hypergraphs on fewer than n vertices. Let $\mathcal{F}$ be a k-uniform on n vertices. Suppose $\mathcal{F}$ has m edges. Suppose first that $m > 4\binom{n-1}{k-1}$. By Lemma 4.2, $\partial_j^*(\mathcal{F}) \geq |\mathcal{T}_{j+1}(\mathcal{F})| \geq \beta_{k,j}m^{\frac{j}{k-1}} \geq c_{k,j}m^{\frac{j}{k-1}}$.

Next, suppose that $m \leq 4\binom{n-1}{k-1} < 4n^{k-1}$. Then $n > m^{\frac{1}{k-1}}/4$. Hence, the average degree of $\mathcal{F}$ is $km/n < 4km^{\frac{k-2}{k-1}}$. Let x be a vertex in $\mathcal{F}$ of minimum degree d. Then $d < 4km^{\frac{k-2}{k-1}}$.

Note that $\mathcal{F}_x$ is $(k-1)$-uniform with d edges. By induction hypothesis, we have $\partial_{j-1}^*(\mathcal{F}_x) \geq c_{k-1,j-1}d^{\frac{j-1}{k-2}}$. Also, $\mathcal{F} - x$ is a k-uniform hypergraph on fewer than n vertices (and has $m - d$ edges). By our assumption, $\partial_j^*(\mathcal{F} - x) \geq c_{k,j}(m-d)^{\frac{j}{k-1}}$. Hence, by Lemma 4.4 we have

$$\partial_j^*(\mathcal{F}) \geq c_{k,j}(m-d)^{\frac{j}{k-1}} + c_{k-1,j-1}d^{\frac{j-1}{k-2}}. \tag{4}$$

Recall that $d \leq 4km^{\frac{k-2}{k-1}}$. Also, $d \leq km/n$. Since we assume n is large (as a function of k), we may further assume that $d \leq m/2$.

Claim 4.6 *We have* $c_{k,j}(m-d)^{\frac{j}{k-1}} + c_{k-1,j-1}d^{\frac{j-1}{k-2}} \geq c_{k,j}m^{\frac{j}{k-1}}$.

*Proof of Claim.*By the Mean Value Theorem, there exists $y \in (m-d, m) \subseteq (m/2, m)$ such that $c_{k,j}m^{\frac{j}{k-1}} - c_{k,j}(m-d)^{\frac{j}{k-1}} = c_{k,j} \cdot d \cdot \frac{j}{k-1} \cdot y^{\frac{j}{k-1}-1}$. It suffices to prove that $c_{k-1,j-1}d^{\frac{j-1}{k-2}} \geq c_{k,j} \cdot d \cdot \frac{j}{k-1}y^{\frac{j}{k-1}-1}$. This is equivalent to $c_{k-1,j-1}y^{\frac{k-j-1}{k-1}} \geq c_{k,j}d^{\frac{k-j-1}{k-2}}$. Since $y \geq m/2$ and $d \leq 4km^{\frac{k-2}{k-1}}$, one can check that the last inequality holds as long as $c_{k,j} \leq \frac{1}{4k}(\frac{1}{2})^{\frac{k-j-1}{k-1}}c_{k-1,j-1}$, which holds by our choice of $c_{k,j}$. ■

By Equation (4) and Claim 2, we have $\partial_j^*(\mathcal{F}) \geq c_{k,j}m^{\frac{j}{k-1}}$. This completes the proof. ■

Lemma 4.7 *Let* $k \geq d + 2 \geq 3$. *Let* $\mathcal{F} \subseteq \binom{[n]}{k}$ *contain no strong* d*-simplex. Let* $x \in [n]$. *Let* $U = \{u_1, \ldots, u_d\} \in \Delta_d(\mathcal{F} - x) \cap \Delta_d(\mathcal{F}_x)$. *Let* $A, B \in \mathcal{F}$ *with* $x \in A, x \notin B$ *and* $U \subseteq A \cap B$. *Let* $W \subseteq [n] \setminus (A \cup B)$ *such that* $|W| = k - d$. *For each* $i \in [d]$, *let* $E_W^i = (\{x\} \cup U \cup W) \setminus \{u_i\}$. *Then for at least one* $i \in [d]$, *we have* $E_W^i \notin \mathcal{F}$.

Proof. Suppose otherwise that for all $i \in [d]$, $E_W^i \in \mathcal{F}$. Consider the collection $\{A, B, E_W^1, \ldots, E_W^d\}$. We have $\cap_{i=1}^d E_W^i = \{x\} \cup W$. For each $i \in [d]$, $\cap_{j \neq i} E_W^j = \{x, u_i\} \cup W$ and so $(\cap_{j \neq i} E_W^j) \cap B = \{u_i\}$. This also implies that $(\cap_{i=1}^d E_W^i) \cap B = \emptyset$. Hence $\{B, E_W^1, \ldots, E_W^d\}$ is d-intersecting but not $(d+1)$-intersecting. That is, it is a d-simplex. As A contains an element of each d-wise intersection among $\{B, E_W^1, \ldots, E_W^d\}$, these $d+2$ sets for a strong d-simplex, contradicting $\mathcal{F}$ having no strong d-simplex. ■

Now, we are ready to prove Theorem 1.10.

Proof of Theorem 1.10: Given d and k, let n be large. Suppose on the contrary that $\mathcal{F} \subseteq \binom{[n]]}{k}$ contains no strong d-simplex, $|\mathcal{F}| \geq \binom{n-1}{k-1}$, and $\mathcal{F}$ is not a star. We derive a contradiction. By Theorem 1.6, there exists an element $x \in [n]$ such that $|\mathcal{F} - x| = o(n^{k-1})$ (that is, almost all edges of $\mathcal{F}$ contain x). Let

$$\begin{aligned} \mathcal{B} &= \mathcal{F} - x \\ \mathcal{M} &= \left\{ D \in \binom{[n]}{k} : x \in D, D \notin \mathcal{F} \right\}. \end{aligned}$$

We call members of $\mathcal{B}$ *bad edges* and members of $\mathcal{M}$ *missing edges*. So, bad edges are those edges in $\mathcal{F}$ not containing x and missing edges are those k-tuples containing x which are not in $\mathcal{F}$. Since $\binom{n-1}{k-1} \leq |\mathcal{F}| = \binom{n-1}{k-1} - |\mathcal{M}| + |\mathcal{B}|$, we have $|\mathcal{B}| \geq |\mathcal{M}|$. Let $b = |\mathcal{B}|$. By the definition of x,

$$b = o(n^{k-1}). \tag{5}$$

By Theorem 4.5, $|\Delta_d(\mathcal{B})| + |\mathcal{T}_{d+1}(\mathcal{B})| = \partial_d^*(\mathcal{B}) \geq c_{k,d} b^{\frac{d}{k-1}}$. Since $\mathcal{B} \subseteq \mathcal{F}$, $\mathcal{B}$ contains no strong d-simplex. So, $|\mathcal{T}_{d+1}(\mathcal{B})| = 0$. It follows that

$$|\Delta_d(B)| \geq c_{k,d} b^{\frac{d}{k-1}}.$$

Let

$$S_1 = \Delta_d(\mathcal{B}) \setminus \Delta_d(\mathcal{F}_x) \text{ and } S_2 = \Delta_d(\mathcal{B}) \cap \Delta_d(\mathcal{F}_x).$$

We consider two cases.

Case 1. $|S_1| \geq |\Delta_d(\mathcal{B})|/2$.

For any $C \in S_1$ and a set $W \subseteq [n] \setminus (C \cup \{x\})$ of size $k-d-1$, the k-tuple $D = \{x\} \cup C \cup W$ does not belong to $\mathcal{F}$ because $C \notin \Delta_d(\mathcal{F}_x)$. So $D \in \mathcal{M}$. Doing this for each $C \in S_1$ yields a list of $\binom{n-d-1}{k-d-1}|S_1|$ edges (with multiplicity) in $\mathcal{M}$. An edge $D = \{x, y_1, \ldots, y_{k-1}\}$ may appear at most $\binom{k-1}{d}$ times in this list, as it is counted once for each d-subset of $\{y_1, \ldots, y_{k-1}\}$ that appears in S_1. Therefore,

$$b \geq |\mathcal{M}| \geq \frac{\binom{n-d}{k-d-1}|S_1|}{\binom{k-1}{d}} \geq c \cdot b^{\frac{d}{k-1}} n^{k-d-1}, \tag{6}$$

for some properly chosen small positive constant c (depending on k only). Solving Equation (6) for b, we get $b \geq c' \cdot n^{k-1}$ for some small positive constant c'. This contradicts Equation (5) for sufficiently large n.

Case 2. $|S_2| \geq |\Delta_d(\mathcal{B})|/2$.

By Lemma 4.7, for every d-tuple $C \in S_2$ we may find two edges $A, B \in \mathcal{F}$ such that every $(k-d)$-set $W \subseteq [n] \setminus (A \cup B)$ contributes an edge to $\mathcal{M}$. So, we obtain a collection of at least $\binom{n-2k}{k-d}|S_2|$ members of $\mathcal{M}$. Pick an edge $D = \{x, y_1, \ldots, y_{k-1}\}$ and consider its multiplicity in this collection. D may appear each time $(d-1)$-subset of $\{y_1, \ldots, y_{k-1}\}$ belongs to some d-tuple in S_2. There are $\binom{k-1}{d-1}$ such subsets and each may be completed to form a d-tuple in $n-d+1$ ways by picking the remaining vertex. So,

$$b \geq |\mathcal{M}| \geq \frac{\binom{n-2k}{k-d}|S_2|}{(n-d+1)\binom{k-1}{d-1}} \geq c'' \cdot b^{\frac{d}{k-1}} \cdot n^{k-d-1},$$

for some small positive constant c''. From this, we get $b \geq c''' \cdot n^{k-1}$, for some positive constant c''', which again contradicts Equation (5) for sufficiently large n. This completes the proof of Theorem 1.10. ■

5 Proof of Theorem 1.11

Here is the main lemma.

Lemma 5.1 *Let $k \geq 2$ and let $\mathcal{H} \subseteq \binom{[n]}{k}$ be arbitrary. Then $\partial^*_{k-1}(\mathcal{H}) \geq |\mathcal{H}|$.*

Proof. We use induction on k. Lemma 4.4 gives us the induction step. So all that remains to do is to establish the base case $k = 2$.

Let $k = 2$. It is enough to prove the inequality for each connectivity component of the given graph $\mathcal{H}$. So we can assume that $\mathcal{H}$ is connected. Here, $\partial^*_{k-1}(\mathcal{H})$ counts the number of vertices of $\mathcal{H}$ plus the number of special pairs. If $n \leq 3$, then an easy case analysis gives the result (with equality being attained by the triangle only).

So, suppose that $n \geq 4$. Pick a spanning tree T in $\mathcal{H}$ such that $\sum_{x \in V(T)} (d_T(x))^2$ is smallest possible. Take any edge xy in $\mathcal{H}$ but not in T. We claim that we can always choose distinct $x', y' \notin \{x, y\}$ such that xx' and yy' are edges of $\mathcal{H}$. If not, then both x and y have degree 2 in $\mathcal{H}$ and, moreover, they share the same common neighbor z. Also, both edges xz and yz belong to the tree T. Since the connected graph $\mathcal{H}$ has more than 3 vertices, the degree of z in T is at least 3. But if we replace the edge xz of T by xy, then we strictly decrease $\sum_{x \in V(T)} (d_T(x))^2$, a contradiction.

Thus every pair in $\mathcal{H} \setminus T$ is special and

$$\partial^*_1(\mathcal{H}) \geq n + (|\mathcal{H}| - |T|) = |\mathcal{H}| + 1,$$

that is, we even get a strictly larger bound. This proves the base case $k = 2$ and the lemma. ■

Now, Theorem 1.11 trivially follows from Lemma 5.1: if $\mathcal{F} \subseteq \binom{[n]}{k}$ has no strong $(k-1)$-simplex, then it has no special k-tuples and, by Lemma 5.1

$$|\mathcal{F}| \leq \Delta_{k-1}(\mathcal{F}) \leq \binom{n}{k-1}.$$

6 Proof of Theorem 1.12

Informally speaking, we want to improve Theorem 1.11 by an extra additive term $n-1$. If we had been able to strengthen the conclusion of Lemma 5.1 for $k=2$ to $\partial_1^*(\mathcal{H}) \geq |\mathcal{H}|+1$, this would have sufficed for the upper bound on Theorem 1.12. However, the union of disjoint triangles shows that that this strengthening is impossible. Yet, one can show that (modulo isolated vertices) this is the only 'counterexample'. So, in order to prove Theorem 1.12 we unfold the induction step of Lemma 5.1 (from $k=2$ to $k=3$) trying to argue that whenever the current link graph $\mathcal{F}_x$ is the union of triangles, we gain extra $+1$ somewhere else. Another difficulty we face is that some vertices of $\mathcal{F}$ may become isolated and we do not have the gain of $+1$ for these vertices. In order to take care of all these issues, we have to account for a few other special configurations $\mathcal{F}_x$, introducing more exceptional cases and making the proof much longer (and requiring that $\mathcal{F}$ has no strong 2-simplex). But the main idea is the same as in the proof of Section 5.

6.1 Auxiliary Results

Let us state some auxiliary results that we will need.

Let us define a function $f : \mathbb{N} \to \mathbb{N}$. To define $f(s)$ for an integer $s \geq 2$, pick the maximum integer l such that $\binom{l}{2} \leq s$ and let $f(s) = l$. The cases $s = 0$ and $s = 1$ do not fit into the general pattern: we let $f(0) = 0$ and $f(1) = 1$.

Lemma 6.1 *Let G be a graph with $v \geq 2$ non-isolated vertices and e edges. Let $s = e - v + 1$. Let p be the number of* special pairs*, that is edges $xy \in G$ such that for some distinct $x', y' \notin \{x, y\}$ we have $xx', yy' \in G$. Also, let us agree that each isolated triangle contributes an extra summand 2 to p.*

If $s \geq 0$, then

$$p + v - e - f(s) - 1 \geq 0. \tag{7}$$

Proof. We prove the lemma by induction on $v \geq 2$. The proof is rather routine. But since the cases $s \leq 1$ and isolated triangles have to be treated in a special manner, we include the complete case analysis.

If $v \leq 3$, then (modulo isolated vertices) G is one of the following graphs: K_2, K_3, or the 2-path P_2. The left-hand side of (7) becomes correspondingly $0+2-1-0-1$, $2+3-3-1-1$, and $0+3-2-0-1$ and (7) is in fact equality here.

Let G be an arbitrary graph with $v \geq 4$ non-isolated vertices and suppose that the lemma holds for all smaller values of v. Clearly, we can remove all isolated vertices from G. Let $f = f(s)$.

If G has an isolated edge xy, let $G' = G - x - y$. Then $v' = v - 2 \geq 2$, $e' = e - 1$, $p' = p$, $s' = s + 1 \geq 0$, and $f' \geq f$. Thus

$$p + v - e - f - 1 \geq p' + (v' + 2) - (e' + 1) - f' - 1 \geq 1 \geq 0,$$

as required, where we used the induction assumption applied to G'. So, assume that G has no isolated edges.

Suppose next that G has a vertex x of degree 1. Let $G' = G - x$. Since we have ruled out isolated edges, we have $v' = v - 1 \geq 3$. Also, $e' = e - 1$, $p' \leq p$, $s' = s$, and $f' = f$.

By induction,,we obtain the required:

$$p+v-e-f-1 \geq p'+(v'+1)-(e'+1)-f'-1 \geq 0.$$

Thus we can additionally assume that the minimum degree of G is at least 2.

If G has an isolated triangle, say on abc, then for $G' = G-a-b-c$ we have $p' = p-2$, $v' = v-3$, $e' = e-3$, $s' = s$, and $f' = f$. Since $v \geq 4$ and G has no isolated vertices, we have $v' \geq 2$. Thus the induction hypothesis applies and we get

$$p+v-e-f-1 = (p'+2)+(v'+3)-(e'+3)-f+1 \geq 2 \geq 0, \tag{8}$$

as required. So, suppose that G does not have an isolated triangle.

Suppose there is a triangle on abc such that $d(a) = d(b) = 2$. By above, $d(c) \geq 3$. Let $G' = G-a-b-c$. Then $p' \leq p-2$ (we lose the special pairs ac, bc and perhaps some others). Since $\delta(G) \geq 2$, we have $v' = v-3 \geq 2$. Also, $e' \leq e-3$, $s' \geq s$, and $f' \geq f$. The induction hypothesis applies to G' and are done (like in (8)).

Since we have ruled all above cases, it follows that every edge xy of G is a special edge, that is, $p = e$. If G is not connected, we can define a new graph G' by identifying two vertices from two different components of G, with the effect that $v' = v-1$, $e' = e$, and $s' = s-1$. Again, we are done by induction applied to G'.

So assume that G is connected. Given v, we have $s \leq \binom{v}{2} - (v-1) = \binom{v-1}{2}$. Thus $f(s) \leq v-1$. Since $p = e$, we conclude that $p+v-e-f-1 \geq 0$, finishing the proof of the lemma. ■

Lemma 6.2 *For any integers* $s_1, s_2 \geq 0$*, we have*

$$f(s_1+s_2) \leq f(s_1)+f(s_2). \tag{9}$$

Proof. Without loss of generality, assume that $s_1 \leq s_2$. If $s_1 = 0$, then inequality (9) becomes $f(s_2) \leq f(s_2)$. If $s_1 = 1$, then inequality (9) becomes $f(1+s_2) \leq 1+f(s_2)$, which is easy to verify. So suppose that $s_1 \geq 2$. For $i = 1, 2$, denote $l_i = f(s_i)$. By the definition of f we have $s_i \leq \binom{l_i+1}{2} - 1$ for $i = 1, 2$. If inequality (9) is false, then we get $f(s_1+s_2) \geq l_1+l_2+1$, which implies that $s_1+s_2 \geq \binom{l_1+l_2+1}{2}$. Thus

$$\binom{l_1+1}{2} - 1 + \binom{l_2+1}{2} - 1 \geq \binom{l_1+l_2+1}{2}.$$

When we simplify, we get $-3/2 \geq l_1 l_2$, a contradiction. ■

6.2 Proof of Theorem 1.12: Putting All Together

Let $\mathcal{F}$ be a 3-graph on $[n]$ with $\binom{n-1}{2}$ edges that does not contain a strong 2-simplex nor isolated vertices. We will argue that $\mathcal{F}$ is the *complete star* (i.e. it consists of all k-tuples that contain some vertex x). Since, for $n \neq 4$, the addition of any edge to this hypergraph creates a strong 2-simplex, the theorem will follow.

We will iteratively apply the following procedure.

Initially, let $\mathcal{F}_1 = \mathcal{F}$, $i = 1$, and $I = \emptyset$.

Let $i \geq 1$ and $\mathcal{F}_i$ be the current 3-graph without a strong 2-simplex with $n-i+1$ vertices none of which is isolated. Add i to I and let x_i denote a vertex of $\mathcal{F}_i$ of the minimal degree; call this degree d_i. We are going to specify a set $R_i \subseteq \mathcal{F}_i$ of edges and a set $P_i \subseteq \binom{V(\mathcal{F}_i)}{2}$ of pairs (see below). Then we let $\mathcal{F}_i' = \mathcal{F}_i \setminus R_i$ and let V_i consist of all isolated vertices of $\mathcal{F}_i'$. Since R_i will contain all edges of $\mathcal{F}_i$ incident to x_i, the vertex x_i belongs to V_i. We label the elements of $V_i \setminus \{x_i\}$ as $x_{i+1}, \ldots, x_j$ (in any order); thus $j = i + |V_i| - 1$. Then we let $\mathcal{F}_{j+1}$ be obtained from $\mathcal{F}_i$ by removing all edges in R_i and all vertices in V_i. If $\mathcal{F}_{j+1}$ has no edges (and thus no vertices by the definition of V_i), then we stop; otherwise we let $i = j+1$ and iterate.

Let us describe how P_i and R_i are constructed. Let

$$P_i' = \mathcal{T}_2((\mathcal{F}_i)_{x_i}) \cup \{yx_i \mid y \in \Delta_1((\mathcal{F}_i)_{x_i})\},$$

where, as before, $(\mathcal{F}_i)_{x_i}$ is the link 2-graph of x_i with respect to $\mathcal{F}_i$, $\mathcal{T}_2$ is the set of special pairs in strong 1-simplices, and Δ_j is the j-shadow.

We have the following four mutually exclusive cases.

Case A $(\mathcal{F}_i)_{x_i}$ has at least one isolated triangle. Initially, we let $R_i = \mathcal{F}_i(x_i)$ (the set of all edges of $\mathcal{F}_i$ containing x_i) and $P_i = P_i'$. Then, for every isolated triangle in $(\mathcal{F}_i)_{x_i}$, say on vertices abc, we add pairs ab, ac, bc to P_i and, if $abc \in \mathcal{F}_i$, we also add the triple abc to R_i.

Case B $(\mathcal{F}_i)_{x_i}$ is an isolated edge ab plus some isolated vertices (and no other edges). We let $R_i = R_i'$ consist of the single edge abx_i. If the pair ab belongs to the 2-shadow of $\mathcal{F}_i - x_i$, we let $P_i = P_i' = \{x_ia, x_ib\}$; otherwise, we let $P_i = P_i' \cup \{ab\} = \{x_ia, x_ib, ab\}$.

Case C $(\mathcal{F}_i)_{x_i}$ is a star with center y with at least 2 edges and every triple of $\mathcal{F}_i$ that contains y also contains x_i. We let $R_i = (\mathcal{F}_i)_{x_i}$ and let P_i consist of the pair x_iy plus all pairs x_iz and yz over all z such that $x_iyz \in \mathcal{F}_i$.

Case D None of Cases 1–3 holds. We let $R_i = \mathcal{F}_i(x_i)$ and $P_i = P_i'$.

Note that $P_i' \subseteq P_i$ and $\mathcal{F}_i(x_i) \subseteq R_i$ in each case. Let us state the properties that we will need as separate claims.

Claim 6.3 *For every two consecutive $i, j \in I$, $i < j$, we have $P_i \subseteq \Delta_2(\mathcal{F}_i) \setminus \Delta_2(\mathcal{F}_j)$. In particular, the sets $P_i \subseteq \binom{[n]}{2}$ over $i \in I$ are pairwise disjoint.*

*Proof of Claim.*The inclusion $P_i \subseteq \Delta(\mathcal{F}_i)$ is obvious from the definition. (In fact, we have $P_i \subseteq \Delta_2(R_i)$.) On the other hand, let $ab \in P_i$ with $x \notin \{a, b\}$. If $ab \in P_i'$, then it cannot belong to the 2-shadow of $\mathcal{F}_i - x_i \supseteq \mathcal{F}_j$ because this would create a strong 2-simplex in $\mathcal{F}_i$. The same applies when ab is an edge of an isolated triangle (when we are in Case A). Otherwise, we are in one of the special Cases B–C and the claim follows from the definition of P_i and R_i. ■

Let us call $i \in I$ *good* if $|P_i| \geq |R_i| + |V_i|$ and *bad* otherwise.

Claim 6.4 *If $i \in I$ is bad, then one of the following situations takes place:*

Case I: *$(\mathcal{F}_i)_{x_i}$ is an isolated triangle on abc plus isolated vertices, $abc \in \mathcal{F}_i$ and $|V_i| = 3$.*

Case II: *Same as Case I except $|V_i| = 4$ (thus $V_i = \{x_i, a, b, c\}$).*

Case III: *$(\mathcal{F}_i)_{x_i}$ is an isolated edge ab, plus isolated vertices, and $V_i = \{x_i, a, b\}$.*

*Proof of Claim.*Let $i \in I$.

First, let us show that if x_i has degree at least 4 in $\mathcal{F}_i$, then i is good.

Suppose first that $V_i = \{x_i\}$. Let t be the number of isolated triangles in $(\mathcal{F}_i)_{x_i}$. Any isolated K_3 contributes at least 6 pairs to P_i and at most 4 triples to R_i (so it increases the difference $|P_i| - |R_i|$ by 2).

Let H be obtained from $(\mathcal{F}_i)_{x_i}$ by removing all isolated triangles and isolated vertices. Let H have v vertices and c connectivity components. Let F be a maximal spanning forest in H (thus F has $v - c$ edges) such that $\sum_{x\in V(F)}(d_F(x))^2$ is minimized. Let $xy \in H \setminus F$. By the maximality of F, each of x, y has degree at least 1 in F. If there are edges $xx', yy' \in F$ with x, y, x', y' pairwise distinct, then xy is a special pair in a 1-simplex of H, so $xy \in P_i$. If not, then x, y have the same unique F-neighbor z. But if $d_F(z) = 2$, then xyz forms an isolated triangle (which contradicts the definition of H) and if $d_F(z) \geq 3$, then by replacing xz by xy in F, we strictly decrease $\sum_{x\in V(F)}(d_F(x))^2$, a contradiction again. Thus $H \setminus F \subseteq P_i$. We conclude that

$$|P_i|-|R_i|-|V_i| \geq (|H\setminus F|+v+6t)-(|H|+4t)-1 \geq |H|-(v-c)+v-|H|-1 \geq 0. \tag{10}$$

Thus the index i is indeed good.

Next, suppose that $|V_i| \geq 2$. Let $y \in V_i \setminus \{x\}$. It is impossible that there is some edge $D \in \mathcal{F}_i$ that contains y but not x. Indeed, then $D \in R_i$ and the only possibility is that y is a vertex of an isolated triangle of $(\mathcal{F}_i)_{x_i}$. But then the degree of y in $\mathcal{F}_i$ is at most 3, contradicting the fact that x_i was chosen to be a vertex of $\mathcal{F}_i$ of the smallest degree. Thus $(\mathcal{F}_i)_{x_i}$ is a star centered at y. Let e be the number of edges in $(\mathcal{F}_i)_{x_i}$. Since $e \geq 4 > 1$, we have $V_i = \{x, y\}$. This is Case C, in which every edge yz of $(\mathcal{F}_i)_{x_i}$ contributes two pairs, namely xz, yz to P_i. Here

$$|P_i| - |R_i| - |V_i| = (2e + 1) - e - 2 = e - 1 \geq 0, \tag{11}$$

and i is good.

Thus it remains to consider the cases when d_i, the degree of x_i in $\mathcal{F}_i$ is at most 3. The case analysis is fairly routine here but we include it for the sake of completeness.

Suppose first that $d_i = 3$ and $(\mathcal{F}_i)_{x_i}$ is an isolated triangle on abc plus isolated vertices (Thus we are in Case A.) If $abc \notin \mathcal{F}_i$, then none of a, b, c belongs to V_i (otherwise this vertex has degree $2 < d_i$, a contradiction); here $|P_i| - |R_i| - |V_i| = 6 - 3 - 1 \geq 0$ so i is good. If $abc \in \mathcal{F}_i$, then $|P_i| - |R_i| - |V_i| = 6 - 4 - |V_i|$, so i is good unless $|V_i| \geq 3$ (when we are either in Case I or II).

Suppose that $d_i \leq 3$ and $(\mathcal{F}_i)_{x_i}$ is a forest.

Suppose also that $|V_i| \geq 2$. Here $(\mathcal{F}_i)_{x_i}$ is a star, say with e edges. If $e \geq 2$, then we have Case C, $|V_i| = 2$, and the estimate (11) applies, showing that i is good. If $e = 1$ (i.e. we are in Case B), say yz is the only edge of $(\mathcal{F}_i)_{x_i}$, then the pair yz cannot belong to the 2-shadow of $\mathcal{F}_i - x_i$ (otherwise $y, z \notin V_i$ and $|V_i| = 1$). By definition, P_i includes pairs x_iy, x_iz, and yz. Also, R_i consists of the single triple x_iyz. Thus $|P_i| - |R_i| - |V_i| = 3 - 1 - |V_i|$ and i is bad only if $|V_i| = 3$, which is exactly Case III.

It remains to consider the case $V_i = \{x_i\}$. Here, (10) applies (with $t = 0$) and i is good. ■

Let $x_+ = x$ if $x \geq 0$ and $x_+ = 0$ if $x < 0$.

Lemma 6.5 *Let $i \in I$ be good. Then*

$$|P_i| \geq |R_i| + |V_i| + f((|R_i| - n + i + 1)_+). \tag{12}$$

Proof. We have some cases. Let $G = (\mathcal{F}_i)_{x_i}$. We will use notation of Lemma 6.1: namely, the variables v, e, s, p are defined with respect to the graph G. Let t be the number of isolated triangles in G. For example, $p + v = |P_i'| + 3t$.

Case 1 All edges in R_i contain x_i.

Here $e = |R_i|$.

Case 1.1 $V_i = \{x_i\}$.

We have $s = e - v + 1 \geq |R_i| - n + i + 1$. By the monotonicity of f, we have $f(s) \geq f((|R_i| - n + i + 1)_+)$. The claim follows from Lemma 6.1. (Note that each isolated triangle contributes at least 2 to $p - e + v$.)

Case 1.2 V_i contains another vertex $y \neq x_i$.

Since x_i has the minimal degree in $\mathcal{F}_i$, the 2-graph G is a star centered at y. In the notation of Lemma 6.1 applied to G, we have $v = e + 1$, $s = 0$, and $f(s) = 0$. The required inequality (12) follows, for example, from our assumption that i is good.

Case 2 There is $abc \in R_i$ such that $x_i \notin abc$.

This implies that $t \geq 1$ (i.e. G has isolated triangles) and we are in Case A.

Case 2.1 G is an isolated triangle one abc, plus isolated vertices.

Let $j = |V_i|$. We have $j \leq 2$ for otherwise we are in Cases I or II and i is not good. Here V_i contains vertices $x_i, \ldots, x_{i+j-1}$. The removals associated with the index $i \in I$ cannot result in the empty 3-graph (otherwise $|V_i| = 4$ and i is bad). So, if $j = 1$, then $i \leq n - 4$. Also, if $j = 2$, then $i \leq n - 5$ (because x_i has the smallest degree in $\mathcal{F}_i$). We can re-write these observations as $i \leq n - j - 3$. Also, $|R_i| = 4$, $|P_i| = 6$, and $f((|R_i| - n + i + 1)_+) = f((5 - n + i)_+)$. This is at most $f((2 - j)_+) = f(2 - j)$ by the monotonicity of f. It is enough to show that

$$6 \geq 4 + j + f(2 - j).$$

But j can assume only values 1 and 2, when the required inequality becomes equality.

Case 2.2 $t \geq 1$ and G has at least 4 edges.

Since x_i has the minimal degree in $\mathcal{F}_i$, we have $|V_i| = 1$. All edges from R_i that do not contain x_i come from isolated triangles in G. The contribution of each such triangle on abc to $|P_i| - |R_i| - |V_i|$ is 3 if $abc \notin \mathcal{F}_i$ and 2 if $abc \in \mathcal{F}_i$. In any case, it is at least 2. But 2 is exactly how much each isolated triangle contributes to p in the setting of Lemma 6.1. So we can remove all isolated triangles from G. If the obtained graph is empty, we have $s = 1$, $f(s) = 1$, $t \geq 2$, and $|P_i| - |R_i| - |V_i| - f(s) = 6t - 4t - 1 - 1 \geq 2 \geq 0$. Otherwise, we apply Lemma 6.1 the same way as in Case 1.1. ■

Let $B = B_1 \cup B_2 \cup B_3 \subseteq I$ be the set of all bad indices, partitioned into 3 sets depending on which one of Cases I–III takes place. Let $b_i = |B_i|$. Let

$$s = \sum_{i \in I \setminus B} (|R_i| - n + i + 1)_+. \tag{13}$$

We have (with explanations following afterwards):

$$\begin{aligned}
|\mathcal{F}| &= \sum_{i\in I}|R_i| \\
&= 4b_1+4b_2+b_3+\sum_{i\in I\setminus B}|R_i| \\
&\le 4b_1+4b_2+b_3+\sum_{i\in I\setminus B}(|P_i|-|V_i|-f((|R_i|-n+i+1)_+)) \\
&\le b_1+2b_2+b_3-f(s)+\sum_{i\in I}|P_i|-\sum_{i\in I}|V_i| \\
&\le b_1+2b_2+b_3-f(s)+\binom{n}{2}-n.
\end{aligned}$$

The first two equalities follow from the fact that R_i, $i\in I$, partition $\mathcal{F}$. The first inequality follows from Lemma 6.5. The second inequality is derived using Lemma 6.2; moreover we have also used the following identities: for $j=1,2,3$, the sums $\sum_{i\in B_j}|P_i|$ and $\sum_{j\in B_j}|V_j|$ are respectively $6b_1,6b_2,3b_4$ and $3b_1,4b_2,3b_3$; thus

$$\sum_{i\in I\setminus B}(|P_i|-|V_i|)=-3b_1-2b_2+\sum_{i\in I}|P_i|-\sum_{i\in I}|V_i|.$$

The last inequality uses the fact that P_i's are pairwise disjoint subsets of $\binom{[n]}{2}$ while V_i's partition $[n]$.

Thus

$$\binom{n-1}{2}=|\mathcal{F}|\le b_1+2b_2+b_3-f(s)+\binom{n-1}{2}-1 \tag{14}$$

At the very end (the last step when all edges disappear), we must have either Case II or Case III (when either $n-4\in B_2$ or $n-3\in B_3$).

Roughly speaking, the idea now is that each bad index i has a very small $|R_i|$ (at most 4). On the other hand, we have

$$\sum_{i\in I}|R_i|=|\mathcal{F}|=\binom{n-1}{2}=(n-2)+(n-1)+\ldots+1. \tag{15}$$

Thus if many indices are bad then the 'positive surplus' s defined by (13) is large and the term $-f(s)$ compensates for this. Let us formally carry out this argument.

Let us first analyze the case when $n-2\in B_3$ (that is, at the last step we remove the single remaining edge). Letting $R_i=\emptyset$ for $i\notin I$, we rewrite (15) as

$$0=\sum_{i=1}^{n-2}(|R_i|-n+i+1)=\sum_{i=1}^{n-2}(|R_i|-n+i+1)_+-\sum_{i=1}^{n-2}(n-|R_i|-i-1)_+=:s_+-s_-. \tag{16}$$

The contribution of the bad indices to the term $s_-=\sum_{i=1}^{n-2}(n-|R_i|-i-1)_+$ is $\sum_{i\in B}(\sum_{x_j\in V_i}(n-j-1)-|R_i|)_+$. This is smallest possible when all $3b_1+4b_2+3b_3$ vertices associated with the bad indices come at the very end. Thus this is at least

$$\left(\sum_{i=3}^{3b_1+4b_2+3b_3}(i-2)\right)-3b_1-4b_2-3b_3=\binom{3b_1+4b_2+3b_3-1}{2}-3b_1-4b_2-3b_3.$$

By (16), we conclude that $s \geq s'$, where

$$s' := \binom{3b_1+4b_2+3b_3-1}{2} - 3b_1 - 4b_2 - 3b_3. \tag{17}$$

This and (14) imply that

$$b_1 + 2b_2 + b_3 - 1 \geq f(s) \geq f(s'). \tag{18}$$

By the definition of f, we conclude that

$$s' = \binom{3b_1+4b_2+3b_3-1}{2} - 3b_1 - 4b_2 - 3b_3 \leq \binom{b_1+2b_2+b_3}{2} - 1. \tag{19}$$

Claim 6.6 *The only feasible solution to (19) in integers $b_1 \geq 0$, $b_2 \geq 0$, $b_3 \geq 1$ is $(b_1, b_2, b_3) = (0, 0, 1)$.*

*Proof of Claim.*Let us rewrite (19) as $Q(b_1, b_2, b_3) \leq 0$ where Q is some quadratic form. We have $Q(0, 0, 1) = -1$. Also, the partial derivatives of Q are

$$\begin{aligned}\frac{\partial Q}{\partial b_1} &= 8b_1 + 10b_2 + 8b_3 - 7\\ \frac{\partial Q}{\partial b_2} &= 10b_1 + 12b_2 + 10b_3 - 9\\ \frac{\partial Q}{\partial b_3} &= 8b_1 + 10b_2 + 8b_3 - 7.\end{aligned}$$

Clearly, each derivative is positive when $b_1 + b_2 + b_3 \geq 1$. Hence the minimum of Q over integers $b_1 \geq 0$, $b_2 \geq 0$, and $b_3 \geq 1$ with $(b_1, b_2, b_3) \neq (0, 0, 1)$ is achieved at one of the points $(1, 0, 1)$, $(0, 1, 1)$, or $(0, 0, 2)$. But the values of Q at these points are 4, 6, and 4 respectively, each contradicting inequality $Q \leq 0$. The claim is proved. ■

Thus $b_1 = b_2 = 0$ and $b_3 = 1$. Note that $f(s) = 0$ (otherwise (18) is false) so $s = 0$. Thus $s_- = 0$. If $|V_i| \geq 2$ for at least one good index $i \in I$, then we are in one of Cases A–C which easily implies $s_- > 0$, a contradiction. Thus $s_+ = 0$ and each R_i has exactly $n - i - 1$ triples. Let us prove by induction on $i = n-2, n-3, \ldots, 1$ that $\mathcal{F}_i$ is the complete star. This is true when $i = n - 2$ when we have 3 vertices and 1 edge. When $i = n - 3$, we have 4 vertices and exactly $1 + 2 = 3$ triples, so $\mathcal{F}_{n-3}$ is the complete star. Suppose that $\mathcal{F}_{i+1}$ is the star with $v = n - i \geq 4$ vertices (and $\binom{v-1}{2}$ edges) centered at some vertex x. When we add x_i and $|R_i| = v - 1$ edges back, then every edge x_iyz has to contain x: otherwise the triples x_iyz, xyw, xzw form a 2-simplex with xyz being the special triple, where w is an arbitrary vertex of $\mathcal{F}_i$ different from x, y, or z. Thus $\mathcal{F}_i$ is a star centered at x and it is the complete star because its size is $\binom{v-1}{2} + (v - 1) = \binom{v}{2}$.

Thus $\mathcal{F} = \mathcal{F}_1$ is the complete star.

Finally, it remains to consider the case when $n - 3 \in B_2$. If $n \leq 4$, we are done so suppose that $n \geq 5$. Thus $\mathcal{F}_{n-3}$ is the complete 3-graph on 4 vertices. Let $U = V(\mathcal{F}_{n-3}) = \{x_{n-3}, x_{n-2}, x_{n-1}, x_n\}$. No other edge of $\mathcal{F}$ can intersect U in 2 vertices as this creates a strong 2-simplex. Thus $n - 4 \notin I$, $n \geq 6$, and, as we have agreed, $R_{n-4} = \emptyset$. While the regime $|R_i| = n - i - 1$ predicts 6 edges on the last 5 vertices, we encountered only 4. Thus $s_+ \geq 2$. The above inequalities remain valid except we can strengthen (17) to

$$s' := \max\left(\binom{3b_1 + 4b_2 + 3b_3 - 1}{2} - 3b_1 - 4b_2 - 3b_3, 2\right). \tag{20}$$

Now, the old inequality (19) and the restrictions $b_1 \geq 0$, $b_2 \geq 1$, and $b_3 \geq 0$ imply that $(b_1, b_2, b_3) = (0, 1, 0)$. But then $f(s') \geq f(2) = 2$ by (20), which contradicts (18). This finishes the proof of Theorem 1.12.

References

[1] J.-C. Bermond and P. Frankl, *On a conjecture of Chvátal on m-intersecting hypergraphs*, Bull. London Math. Soc. **9** (1977), 309–312.

[2] V. Chvátal, *An extremal set-intersection theorem*, J. London Math. Soc. **9** (1974/75), 355-359.

[3] R. Csákány and J. Kahn, *A homological approach to two problems on finite sets*, J. Algebraic Comb. **9** (1999), 141–149.

[4] P. Erdős, *Topics in combinatorial analysis*, Proceedings of the Second Louisiana Conference on Combinatorics, Graph Theory, and Computing (R. C. Mullin et al, ed.), Pages 2–20, Louisiana State University, 1971.

[5] P. Erdős, H. Ko and R. Rado, *Intersection theorem for systems of finite sets*, Quart. J. Math. Oxford Set. **12** (1961) 313-320.

[6] P. Frankl, *On Sperner families satisfying an additional condition*, J. Combin. Theory (A) **20** (1976), 1–11.

[7] P. Frankl, *On a problem of Cvátal and Erdős on hypergraphs containing no special simplex*, J. Combin. Theory (A) **30** (1981), 169–182.

[8] P. Frankl and Z. Füredi, *Exact solution of some Turán-type problems*, J. Combinatorial Theory Series A **45 (2)** (1987), 226-262.

[9] Z. Füredi and L. Özkahya, *Unavoidable subhypergraphs: a-clusters*, Extended Abstract (EuroComb 2009), 2009.

[10] P. Keevash and D. Mubayi, *Set systems without a simplex or a cluster*, Combinatorica, to appear.

[11] D. Mubayi, *Structure and stability of triangle-free systems*, Trans. Amer. Math. Soc. **359** (2007) 275-291.

[12] D. Mubayi and R. Ramadurai, *Simplex Stability*, Combinatorics, Probability and Computing **18** (2009), 441-454.

[13] D. Mubayi and J. Verstraëte, *Proof of a conjecture of Erdős on triangles in set systems*, Combinatorica **25** (2005), 599-614.

[14] V. Rödl, *On a packing and covering problem*, Europ. J. Combin. **5** (1985), 69–78.

[15] P. Turán, *On an extremal problem in graph theory (in Hungarian)*, Mat. Fiz. Lapok **48** (1941), 436–452.