Stability theorems for the feasible region of cancellative hypergraphs

Xizhi Liu

December 30, 2019

Abstract

A hypergraph is cancellative if it does not contain three sets $A, B, C$ such that the symmetric difference of $A$ and $B$ is contained in $C$. We show that for every $r \geq 3$ a cancellative $r$-graph $\mathcal{H}$ has a stability property whenever the sizes of $\mathcal{H}$ and the shadow of $\mathcal{H}$ satisfy certain inequalities. In particular, our result for $r = 3$ generalizes a stability theorem of Keevash and Mubayi and it shows that for every $k \equiv 1$ or $3 \pmod 6$ a 3-graph $H$ is structurally close to a balanced blow up of a Steiner triple system on $k$ vertices whenever the shadow density of $H$ is close to $(k-1)/k$ and the edge density of $H$ is close to $(k-1)/k^2$. Our result for $r \geq 3$ extends a stability theorem of Keevash about the Kruskal-Katona theorem to cancellative hypergraphs, and also addresses an old conjecture of Bollobás about the maximum size of a cancellative $r$-graph.

1 Introduction

Let $r \geq 2$ and $\mathcal{F}$ be a family of $r$-graphs. An $r$-graph is $\mathcal{F}$-free if it does not contain any member of $\mathcal{F}$ as a subgraph. The Turán number $\text{ex}(n, \mathcal{F})$ of $\mathcal{F}$ is the maximum size of an $\mathcal{F}$-free $r$-graph on $n$ vertices, and the Turán density of $\mathcal{F}$ is $\pi(\mathcal{F}) := \lim_{n \to \infty} \text{ex}(n, \mathcal{F})/\binom{n}{r}$.

It is one of the central problems in extremal combinatorics to determine $\text{ex}(n, \mathcal{F})$ for various families $\mathcal{F}$.

Much is known about $\text{ex}(n, \mathcal{F})$ when $r = 2$ and one the most famous results in this regard is Turán’s theorem [19], which states that for $\ell \geq 2$ the Turán number $\text{ex}(n, K_{\ell+1})$ is uniquely achieved by $T(n, \ell)$ which is the $\ell$-partite graph on $n$ vertices with the maximum number of edges. However, for $r \geq 3$ determining $\text{ex}(n, \mathcal{F})$, even $\pi(\mathcal{F})$, is notoriously hard. Compared to the case $r = 2$, very little is known about $\text{ex}(n, \mathcal{F})$ for $r \geq 3$, and we refer the reader to [9] for results before 2011.

In 1960’s, Katona tried to generalize Turán’s theorem to 3-graphs and conjectured that the maximum size of a 3-graph on $n$ vertices that does not contain three sets $A, B, C$ with $A \Delta B \subset C$ is achieved by the balanced 3-partite 3-graph, i.e. every two part sizes differ by at most one. Katona’s conjecture was later proved by Bollobás [1]. In order to state Bollobás’ result formally let us introduce some notations.

Let $r \geq 2$ and $\mathcal{T}_r$ be the family of $r$-graphs with at most $2r-1$ vertices and three edges $A, B, C$ such that $A \Delta B \subset C$. An $r$-graph is cancellative iff it is $\mathcal{T}_r$-free. Let $\ell \geq r$ and $V_1 \cup \cdots \cup V_t$ be a partition of $[n] := \{1, \ldots, n\}$ with each $V_i$ of size either $\lfloor n/\ell \rfloor$ or $\lceil n/\ell \rceil$. The generalized Turán graph $T_r(n, \ell)$ is the collection of all $r$-subsets of $[n]$ that have at most one vertex in each $V_i$. Let $t_r(n, \ell) = |T_r(n, \ell)| \sim \binom{\ell}{r} \binom{n}{\ell}^r$.

Theorem 1.1 (Bollobás, [1]). A cancellative 3-graph on $n$ vertices has size at most $t_3(n, 3)$, with equality only for $T_3(n, 3)$. 

*Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL, 60607 USA. Email: xliu246@uic.edu
Moreover, Bollobás conjectured that a similar result holds for all \( r \geq 4 \).

**Conjecture 1.2** (Bollobás, [11]). *For every \( r \geq 4 \) a cancellative \( r \)-graph on \( n \) vertices has size at most \( t_r(n, r) \), with equality only for \( T_r(n, r) \).*

Sidorenko [17] proved Conjecture 1.2 for \( r = 4 \), but Shearer [16] gave a construction showing that Conjecture 1.2 is false for all \( r \geq 4 \). However, \( ex(n, T_r) \) is still unknown for all \( r \geq 5 \), even asymptotically.

In another direction, Keevash and Mubayi [11] proved a stability theorem for cancellative 3-graphs. Given an \( r \)-graph \( \mathcal{H} \) we use \( V(\mathcal{H}) \) to denote the vertex set of \( \mathcal{H} \) and \( v(\mathcal{H}) = |V(\mathcal{H})| \).

**Theorem 1.3** (Keevash and Mubayi, [11]). *For every \( \delta > 0 \) there exists \( \epsilon > 0 \) and \( n_0 \) such that the following holds for all \( n \geq n_0 \). Every cancellative 3-graph \( \mathcal{H} \) with \( n \) vertices and at least \((1-\epsilon) t_3(n, 3) \) edges has a partition \( V(\mathcal{H}) = V_1 \cup V_2 \cup V_3 \) such that all but at most \( \delta n^3 \) edges in \( \mathcal{H} \) have exactly one vertex in each \( V_i \).

Actually the original statement of Theorem 1.3 is stronger and we refer the reader to Theorem 1.5 in [11] for details. A similar stability theorem for cancellative 4-graphs follows from Pikhurko’s results in [15].

Let \( \mathcal{H} \) be an \( r \)-graph on \( n \) vertices. The shadow of \( \mathcal{H} \) is

\[
\partial \mathcal{H} := \left\{ A \in \binom{V(\mathcal{H})}{r-1} : \exists B \in \mathcal{H} \text{ such that } A \subseteq B \right\}.
\]

The edge density of \( \mathcal{H} \) is \( d(\mathcal{H}) := |\mathcal{H}|/\binom{n}{r} \) and the shadow density of \( \mathcal{H} \) is \( d(\partial \mathcal{H}) := |\partial \mathcal{H}|/\binom{n}{r-1} \). The classical Kruskal-Katona theorem gives a tight upper bound for \( |\mathcal{H}| \) as a function of \( |\partial \mathcal{H}| \), and we state the following technically simpler version of the Kruskal-Katona theorem which is due to Lovász.

**Theorem 1.4** (see Lovász [14]). *Let \( \mathcal{H} \) be an \( r \)-graph, and suppose that \(|\partial \mathcal{H}| = \binom{\frac{z}{r-1}}{r-1} \) for some real number \( z \geq r \). Then \(|\mathcal{H}| \leq \binom{z}{r-1} \).*

The feasible region \( \Omega(\mathcal{F}) \) of \( \mathcal{F} \) is the set of points \((x, y) \in [0, 1]^2\) such that there exists a sequence of \( \mathcal{F} \)-free \( r \)-graphs \( (\mathcal{H}_k)_{k=1}^\infty \) with \( \lim_{k \to \infty} v(\mathcal{H}_k) = \infty \), \( \lim_{k \to \infty} d(\partial \mathcal{H}_k) = x \) and \( \lim_{k \to \infty} d(\mathcal{H}_k) = y \). Mubayi and the author introduced this notation recently in [12] as a way of studying the extremal properties of \( \mathcal{F} \)-free hypergraphs that goes well beyond just the determination of \( \pi(\mathcal{F}) \). In particular, we proved that \( \Omega(\mathcal{F}) \) is completely determined by a left-continuous almost everywhere differentiable function \( g(\mathcal{F}) : \text{proj}\Omega(\mathcal{F}) \to [0, 1] \), where

\[
\text{proj}\Omega(\mathcal{F}) = \{ x : \exists y \in [0, 1] \text{ such that } (x, y) \in \Omega(\mathcal{F}) \},
\]

and

\[
g(\mathcal{F}, x) = \max \{ y : (x, y) \in \Omega(\mathcal{F}) \}, \text{ for all } x \in \text{proj}\Omega(\mathcal{F}).
\]

Note that for fixed \( r \geq 3 \), the Kruskal-Katona theorem (and some other observations) implies that \( g(\emptyset, x) = x^{r/(r-1)} \) for all \( x \in [0, 1] \). For cancellative hypergraphs, the following results were proved in [12].

**Theorem 1.5** ([12]). *Let \( r \geq 2 \) and let \( \mathcal{H} \) be a cancellative \( r \)-graph. Then

\[
|\mathcal{H}| \leq \left( \frac{|\partial \mathcal{H}|}{r} \right)^{\frac{r}{r-1}}.
\]
In particular, for every \( x \in \text{proj}\Omega(T_r) \),

\[
g(T_r, x) \leq \left( \frac{x^r}{r!} \right)^{\frac{1}{r-1}}.
\]

Moreover, equality holds for all \( x \in [0, (r-1)!/r^{r-2}] \).

Let \( 6\mathbb{N} + \{1, 3\} \) denote the set of all positive integers \( k \) with \( k \equiv 1 \text{ or } 3 \pmod{6} \).

**Theorem 1.6** ([12]). Let \( H \) be a cancellative 3-graph on \( n \) vertices. Then

\[
|H| \leq \frac{(n^2 - 2|\partial H|)|\partial H|}{3n} + 3n^2.
\]

In particular, \( g(T_3, x) \leq x(1-x) \) for all \( x \in [0,1] \). Moreover, for every \( k \in 6\mathbb{N} + \{1, 3\} \), \( g(T_3, (k-1)/k) = (k-1)/k^2 \).

Theorems 1.5 and 1.6 imply that \( \pi(T_3) \leq \max_{x \in [0,1]} \{ g(T_3, x) \} = \max_{x \in [0,1]} \{ \min\{ (x^3/6)^{1/2}, x(1-x) \} \} = \frac{2}{9} \),

which is a weak version of Theorem 1.1. On the other hand, for every \( k \in 6\mathbb{N} + \{1, 3\} \) the lower bound for \( g(T_3, (k-1)/k) \) is given by balanced blow ups of Steiner triple systems (which will be explained in detail later), and in [12] the following problem was posed.

**Problem 1.7.** For every \( k \in 6\mathbb{N} + \{1, 3\} \) with \( k \geq 7 \), is the point \( ((k-1)/k, (k-1)/k^2) \) a local maximum of \( g(T_3) \)?

---

**Figure 1:** \( \Omega(T_3) \) is bounded by \( \min\{ (x^3/6)^{1/2}, x(1-x) \} \) according to Theorems 1.5 and 1.6.

---

**Definition 1.8.** For fixed \( r \geq 3 \) and a family \( F \) of \( r \)-graphs, the boundary of \( \Omega(F) \) is

\[
\partial \Omega(F) := \{ (x, g(F, x)) : x \in \text{proj}\Omega(F) \}.
\]

In this paper we study the stability property of cancellative \( r \)-graphs \( H \) for \( r \geq 3 \) when \( (d(\partial H), d(H)) \) is close to \( \partial \Omega(T_r) \). Our result for \( r \geq 3 \) is an extension of a stability theorem by Keevash [8] about the Kruskal-Katona theorem to cancellative hypergraphs (see Theorems 1.13 and 1.14), and it gives more information about Conjecture 1.2. Our result for \( r = 3 \) (see Theorem 1.15) contains Theorem 1.3 as a special case, and it might be helpful in solving Problem 1.7.

On the other hand, studying the stability property of points in \( \partial \Omega(F) \) is also helpful in understanding the local property of \( g(F) \). For example, in [12] the stability property of the the family \( D^r \) (we refer the reader to [12] for the definition of \( D^r \)) was successfully applied.
to show that the function \( g(D') \) has a discontinuity. In [13], the stability property of an \( \mathcal{M} \)-free (we refer the reader to [13] for the definition of \( \mathcal{M} \)) 3-graph \( \mathcal{H} \) when \( (d(\partial H), d(H)) \) is close to \((5/6, 4/9)\) or \((8/9, 4/9)\) was used to show that \( \{(5/6, 4/9), (8/9, 4/9)\} \) are (the only) global maximums of the function \( g(M) \).

Before stating our results formally let us introduce some definitions.

**Definition 1.9.** Let \( r \geq 3 \) and \( \mathcal{F} \) be a family of \( r \)-graphs, \( (x_0, y_0) \in \Omega(\mathcal{F}) \) and \( \epsilon > 0 \). Let

\[
B^\epsilon_\mathcal{F}(x_0, y_0) = \{(x, y) \in \Omega(\mathcal{F}) : \text{dist}((x, y), (x_0, y_0)) < \epsilon \},
\]

where \( \text{dist}((x, y), (x_0, y_0)) = \sqrt{(x-x_0)^2 + (y-y_0)^2} \) is distance of \((x, y)\) and \((x_0, y_0)\) in the Euclidean space \( \mathbb{R}^2 \).

**Definition 1.10** (Graph edit distance). Let \( r \geq 2 \) and \( \mathcal{H}_1, \mathcal{H}_2 \) be two \( r \)-graphs with \( v(\mathcal{H}_1) = v(\mathcal{H}_2) \). The edit-distance between \( \mathcal{H}_1 \) and \( \mathcal{H}_2 \), denoted by \( ed(\mathcal{H}_1, \mathcal{H}_2) \), is the minimum integer \( d \) such that \( \mathcal{H}_1 \) can be transformed into a copy of \( \mathcal{H}_2 \) by removing and adding \( d \) edges.

**Definition 1.11** (t-stable points). Let \( r \geq 3 \) and \( \mathcal{F} \) be a family of \( r \)-graphs. A point \((x_0, y_0) \in \Omega(\mathcal{F})\) is said to be \( t \)-stable for \( t \geq 1 \) if there exists \( m_0 \) and \( G^1_m, \ldots, G^t_m \) for all integer \( m > m_0 \) such that the following holds. For every \( \delta > 0 \) there exists \( \epsilon > 0 \) and \( n_0 \) such that every \( \mathcal{F} \)-free \( r \)-graph \( \mathcal{H} \) on \( n \geq n_0 \) vertices with \( (d(\partial H), d(H)) \in B_\epsilon(x_0, y_0) \) satisfies \( ed(\mathcal{H}, G^i_m)/{\binom{n}{r}} < \delta \) for some \( i \in [t] \). In particular, \( 1 \)-stable points are called stable points. If \((x_0, y_0) \in \Omega(\mathcal{F})\) is not \( t \)-stable for any \( t > 0 \), then it is called \( \infty \)-stable.

**Definition 1.12** (Stability number of a point). Let \( r \geq 3 \) and \( \mathcal{F} \) be a family of \( r \)-graphs, \( (x_0, y_0) \in \Omega(\mathcal{F}) \). The stability number of \((x_0, y_0)\), denoted by \( \xi_\mathcal{F}(x_0, y_0) \), is the minimum integer \( t \) such that \((x_0, y_0)\) is \( t \)-stable. If there is no such \( t \), then we set \( \xi_\mathcal{F}(x_0, y_0) = \infty \).

For the case \( r \geq 3 \) and \( F = \emptyset \), Keevash [13] proved a corresponding stability theorem of Theorem 1.14.

**Theorem 1.13** (Keevash. [13]). For every \( r \geq 2 \) and \( \delta > 0 \) there exists \( \epsilon > 0 \) such that every \( r \)-graph \( \mathcal{H} \) with \( |\partial H| = \binom{r}{r-1}\) and \( |\mathcal{H}| > (1-\epsilon)\binom{n}{r} \) contains a set \( S \) of size \( \lfloor z \rfloor \) such that all but at most \( \delta n^r \) edges of \( \mathcal{H} \) are contained in \( S \).

Let \( x \in [0,1] \) and \( \alpha = x^{1/(r-1)} \). Let \( K^r(n, x) \) be the disjoint union of a complete 3-graph on \( \lfloor \alpha n \rfloor \) vertices and a set of \( n - \lfloor \alpha n \rfloor \) isolated vertices. Theorem 1.13 says that for fixed \( r \geq 3 \), for every \( x \in [0,1] \), \( (x, g(0, x)) \in \Omega(\emptyset) \) is stable with respect to \( K^r(n, x) \).

Fix \( x \in [0, (r-1)!/r^{r-2}] \) and let \( \alpha' = (xr^{r-2}/(r-1)!)^{1/(r-1)} \). The \( r \)-graph \( H^r(n, x) \) is the disjoint union of \( T_r([\alpha' n], r) \) and a set of \( n - \lfloor \alpha' n \rfloor \) isolated vertices. Our first result extends Theorem 1.13 to cancellative hypergraphs and it shows that for every \( r \geq 3 \), \( (x, g(T_r, x)) \in \Omega(T_r) \) is stable with respect to \( H^r(n, x) \).

**Theorem 1.14.** Let \( r \geq 3 \), \( x \in [0, (r-1)!/r^{r-2}] \) and \( y = g(T_r, x) \). For every \( \delta > 0 \) there exists \( \epsilon > 0 \) and \( n_0 \) such that the following holds for all \( n \geq n_0 \). Suppose that \( H \) is cancellative \( r \)-graph on \( n \) vertices with \( (d(\partial H), d(H)) \in B^\epsilon_{T_r}(x, y) \). Then \( ed(\mathcal{H}, H^r(n, x)) < \delta n^r \). In particular, \( \xi_{T_r}(x, y) = 1 \).

A Steiner triple system (STS) on \( k \) vertices is a 3-graph on \( k \) vertices such that every pair of vertices is contained in exactly one edge. It is known that a \( k \)-vertex STS exists iff \( k \in 6\mathbb{N} + \{1, 3\} \). (e.g. see [20]). Let STS(\( k \)) denote the family of all Steiner triple systems on \( k \) vertices. For example, STS(3) comprises only one 3-graph \( K^3_3 \), STS(6) comprises of only one 3-graph, which is the Fano plane, and STS(9) comprises of only one 3-graph, which is the affine plane of order 3. For \( k \in 6\mathbb{N} + \{1, 3\} \) let \( s_k \) denote the maximum
In particular, \( \xi \)

**Theorem 1.15.** Let \( k \in 6\mathbb{N} + \{1, 3\} \). For every \( \delta > 0 \) there exists \( \epsilon > 0 \) and \( n_0 \) such that the following holds for all \( n \geq n_0 \). Suppose that \( H \) is cancellative 3-graph on \( n \) vertices with 
\[ (d(\partial H), d(H)) \in B_{T_3}((k-1)/k, (k-1)/k^2). \]
Then \( \text{ed}(\mathcal{H}, \mathcal{G}) < \delta n^r \) for some \( \mathcal{G} \in \mathcal{S}(n, k) \). In particular, \( \xi_{T_3} \left( \frac{k-1}{k}, \frac{k-1}{k^2} \right) = s_k \).

Moreover, we are able to determine exactly the maximum size of a cancellative 3-graph \( H \) with \( n \) vertices and \( |\partial H| = t_2(n, k) \) when \( n \) is large.

**Theorem 1.16.** Let \( k \in 6\mathbb{N} + \{1, 3\} \) and \( n \) be sufficiently large. Suppose that \( H \) is a cancellative 3-graph on \( n \) vertices with \( |\partial H| = t_2(n, k) \). Then \( |H| \leq s(n, k) \), and equality holds only if \( H \in \mathcal{S}(n, k) \).

The remainder of this paper is organized as follows. In Section 2 we prove Theorem 1.14. In Section 3 we prove Theorems 1.15 and 1.16.

## 2 Stable points in \( \partial \Omega(T_r) \) for all \( r \geq 3 \)

In this section we will prove the following statement, which implies Theorem 1.14:

**Theorem 2.1.** Let \( r \geq 3 \) and \( c > 0 \) be a constant. For every \( \delta > 0 \) there exists \( \epsilon > 0 \) and \( n_0 \) such that the following holds for all \( n \geq n_0 \). Suppose that \( H \) is a cancellative \( r \)-graph on \( n \) vertices with \( |\partial H| \geq cn^{r-1} \) and \( |H| > (1 - \epsilon) \left( |\partial H|/r \right)^{r/(r-1)} \). Then, \( \text{ed}(\mathcal{H}, \mathcal{H}'(n, x)) < \delta n^r \), where \( x = |\partial H|/\binom{n}{r-1} \).

The proof of Theorem 2.1 contains two parts: Lemma 2.5 and Lemma 2.14. Lemma 2.5 reduces the stability of \((x, g(T_r, x))\) to the stability of \( \binom{(r-1)!}{r-1} r! \), and Lemma 2.14 shows that \( \binom{(r-1)!}{r-1} r! \) is 1-stable respects \( T_r(n, r) \).
For $1 \leq i \leq r - 1$, the $i$-th shadow of an $r$-graph $\mathcal{H}$ is

$$\partial_i \mathcal{H} := \left\{ A \in \binom{V(\mathcal{H})}{r-i} : \exists B \in \mathcal{H} \text{ such that } A \subset B \right\}.$$}

Let $v \in V(\mathcal{H})$. The link of $v$ in $\mathcal{H}$ is

$$L_{\mathcal{H}}(v) := \left\{ A \in \binom{V(\mathcal{H})}{r-1} : \{v\} \cup A \subset \mathcal{H} \right\},$$

and $d_{\mathcal{H}}(v) = |L_{\mathcal{H}}(v)|$. We will omit the subscript if it is clear from context.

**Lemma 2.2** (e.g. see [12]). Let $r \geq 3$, $\mathcal{H}$ be a cancellative $r$-graph, and $S \subset V(\mathcal{H})$. Suppose that $(\partial_{r-2} \mathcal{H})[S]$ is complete. Then $L(v) \cap L(u) = \emptyset$ for all $\{u, v\} \subset S$ and, in particular, $\sum_{v \in S} d_{\mathcal{H}}(v) \leq |\partial \mathcal{H}|$.

The next lemma follows from the proof of Theorem 4.4 in [12]. Given $S \subset V(\mathcal{H})$, let $\sigma(S) = \sum_{v \in S} d_{\mathcal{H}}(v)$. Let $\hat{\sigma} = \max \{ \sigma(H) : H \in \mathcal{H} \}$ and fix $E \in \mathcal{H}$ such that $\sigma(E) = \hat{\sigma}$.

**Lemma 2.3.** Let $r \geq 3$ and $\mathcal{H}$ be a cancellative $r$-graph. Then

$$|\mathcal{H}| \leq \frac{|\partial \mathcal{H}|^{-\frac{2}{r}}}{r(r-1)^{1/(r-1)}} \left( \left( |\partial \mathcal{H}| - \frac{\hat{\sigma}}{r} \right)^{\frac{1}{r-1}} \right)^{\frac{1}{r-1}},$$

$$|\mathcal{H}| \leq \frac{|\partial \mathcal{H}|^{-\frac{2}{r}}}{r(r-1)^{1/(r-1)}} \left( \sum_{v \in E} \sum_{S \in L(v)} d_{\mathcal{H}}(v) (\hat{\sigma} - d_{\mathcal{H}}(v)) + (|\partial \mathcal{H}| - \hat{\sigma}) \hat{\sigma} \right)^{\frac{1}{r-1}},$$

$$|\mathcal{H}| \leq \frac{|\partial \mathcal{H}|^{-\frac{2}{r}}}{r(r-1)^{1/(r-1)}} \left( \sum_{v \in E} \sum_{S \in L(v)} \sigma(S) + \sum_{S \in \partial \mathcal{H} \setminus \cup_{v \in E} L(v)} \sigma(S) \right)^{\frac{1}{r-1}},$$

and

$$\frac{1}{r(r-1)} \sum_{v \in V(\mathcal{H})} \frac{d_{\mathcal{H}}(v)}{r} \leq \frac{|\partial \mathcal{H}|}{r}.$$  \hfill (4)

The following lemma will be used intensively in our proofs (including Section 3).

**Lemma 2.4.** Let $f : \mathbb{R} \to \mathbb{R}$ be a function, $\delta_1$ and $\delta_2$ be two nonnegative real numbers, and $S \subset \mathbb{R}$ be a finite set. Let $E = \left( \sum_{s \in S} f(s) \right)/|S|$ and $S' = \{ s \in S : f(s) < E - \delta_1 \}$. Suppose that $\max_{s \in S} \{ f(s) \} < E + \delta_2$. Then

$$|S'| < \frac{\delta_2}{\delta_1 + \delta_2} |S|.$$  

**Proof.** By assumption,

$$|S| E = \sum_{s \in S} f(s) = \sum_{s' \in S'} f(s') + \sum_{s \in S \setminus S'} f(s) < |S'| (E - \delta_1) + (|S| - |S'|) (E + \delta_2) = |S| E + \delta_2 |S| - (\delta_1 + \delta_2) |S'|,$$

which implies that $|S'| < \delta_2 |S|/(\delta_1 + \delta_2)$. \hfill \blacksquare

The next lemma reduces the proof of Theorem 2.1 to the case $|\partial \mathcal{H}| \sim r^{-1}/r^{r-2}$.  

6
Lemma 2.5. Let $r \geq 3$, $c > 0$ be a constant, $\epsilon > 0$ be sufficiently small, and $n$ be sufficiently large. Suppose that $\mathcal{H}$ is a cancellative $r$-graph on $n$ vertices with $|\partial \mathcal{H}| \geq cn^{r-1}$ and $|\mathcal{H}| > (1 - \epsilon) (|\partial \mathcal{H}|/r)^{r/(r-1)}$. Then, there exists $U \subset V(\mathcal{H})$ with

$$
1 - \frac{16r^4 \epsilon^{1/2}}{c^1/(r-1)} \frac{r-2}{r} |\partial \mathcal{H}|^{r-1} < |U| < \left( 1 + 2r^5/2 \epsilon^{1/2} \right) \frac{r-2}{r} |\partial \mathcal{H}|^{r-1}
$$

such that

$$
|\partial (\mathcal{H}[U])| > \left( 1 - 8r^2 \epsilon^{1/2} \right) |\partial \mathcal{H}| \quad \text{and} \quad |\mathcal{H}[U]| > \left( 1 - 16r^4 \epsilon^{1/2} \right) \frac{1}{c^{1/(r-1)}} |\mathcal{H}|.
$$

Proof. We prove this lemma through a series of claims.

Claim 2.6. $(1 - 2r \epsilon) |\partial \mathcal{H}| < \bar{\sigma} \leq |\partial \mathcal{H}|$.

Proof of Claim 2.6. The inequality $\bar{\sigma} \leq |\partial \mathcal{H}|$ follows from Lemma 2.2, so we may focus on the lower bound for $\bar{\sigma}$.

It follows from our assumption and (1) that

$$
(1 - \epsilon) \left( \frac{|\partial \mathcal{H}|}{r} \right)^{r-1} < |\mathcal{H}| \leq \frac{|\partial \mathcal{H}|^{r-2}}{r(r-1)^{1/(r-1)}} \left( \left( |\partial \mathcal{H}| - \bar{\sigma} \frac{r}{r} \right) \bar{\sigma} \right)^{1/(r-1)}.
$$

Consequently,

$$
\left( |\partial \mathcal{H}| - \bar{\sigma} \frac{r}{r} \right) \bar{\sigma} > (1 - \epsilon) r^{-1} \frac{r-1}{r} \epsilon r^2 > (1 - (r - 1) \epsilon) r^{-1} |\partial \mathcal{H}|^2,
$$

which implies that $\bar{\sigma} > (1 - 2r \epsilon) |\partial \mathcal{H}|$.

Claim 2.7. $|d(v) - \bar{\sigma}/r| < (2r \epsilon)^{1/2} \bar{\sigma}$ for all $v \in E$.

Proof of Claim 2.7. First, we prove that

$$
\sum_{v \in E} d(v) (\bar{\sigma} - d(v)) > \left( \frac{r - 1}{r} - 2r \epsilon \right) (\bar{\sigma})^2. \quad (5)
$$

Suppose that (5) is not true. Then

$$
|\mathcal{H}| \leq \left( \frac{|\partial \mathcal{H}|^{r-2}}{r(r-1)^{1/(r-1)}} \left( \sum_{v \in E} d(v) (\bar{\sigma} - d(v)) + (|\partial \mathcal{H}| - \bar{\sigma}) \bar{\sigma} \right) \right)^{1/(r-1)}
$$

$$
\leq \frac{|\partial \mathcal{H}|^{r-2}}{r(r-1)^{1/(r-1)}} \left( \left( \frac{r - 1}{r} - 2r \epsilon \right) (\bar{\sigma})^2 + (|\partial \mathcal{H}| - \bar{\sigma}) \bar{\sigma} \right)^{1/(r-1)}
$$

$$
\leq \frac{|\partial \mathcal{H}|^{r-2}}{r(r-1)^{1/(r-1)}} \left( \left( |\partial \mathcal{H}| - \left( \frac{1}{r} + 2r \epsilon \right) \bar{\sigma} \right) \bar{\sigma} \right)^{1/(r-1)}
$$

Claim 2.8

$$
\leq \frac{2r^2}{r(r-1)^{1/(r-1)}} \left( \frac{r - 1}{r} - 2r \epsilon \right) |\partial \mathcal{H}|^2 \bar{\sigma} \frac{1}{r} \left( \frac{|\partial \mathcal{H}|}{r} \right)^{r-1}
$$

$$
= \left( 1 - 2r \epsilon \right) \left( \frac{|\partial \mathcal{H}|}{r} \right)^{r-1} < (1 - \epsilon) \left( \frac{|\partial \mathcal{H}|}{r} \right)^{r-1},
$$

a contradiction. Therefore, (5) is true.
Now suppose that Claim 2.7 is not true. Assume that \( E = \{v_1, \ldots, v_r\} \) and without loss of generality we may assume that \(|d(v_1) - \hat{\sigma}/r| \geq (2\varepsilon)^{1/2} \hat{\sigma} \). Then

\[
\sum_{i \in [r]} d(v_i) (\hat{\sigma} - d(v_i)) = d(v_1) (\hat{\sigma} - d(v_1)) + \sum_{i=2}^r d(v_i) (\hat{\sigma} - d(v_i)) \\
\leq d(v_1) (\hat{\sigma} - d(v_1)) + \left( \sum_{i=2}^r d(v_i) \right) \left( \hat{\sigma} - \frac{\sum_{i=2}^r d(v_i)}{r-1} \right) \\
= d(v_1) (\hat{\sigma} - d(v_1)) + (\hat{\sigma} - d(v_1)) \left( \hat{\sigma} - \frac{r-2}{r-1} d(v_1) \right) \\
= \frac{r-2}{r-1} (\hat{\sigma} - d(v_1)) \left( \hat{\sigma} + \frac{r}{r-2} d(v_1) \right) \\
\leq \frac{r-1}{r-1} (\hat{\sigma})^2 - \frac{2r^2}{r-1} \epsilon (\hat{\sigma})^2 < \frac{r-1}{r} (\hat{\sigma})^2 - 2r \epsilon (\hat{\sigma})^2 ,
\]

which contradicts (5). \( \blacksquare \)

For every \( v \in E \) let \( \mathcal{L}_v = \{ S \in L(v) : \sigma(S) \geq (1 - \epsilon^{1/2}) (\hat{\sigma} - d(v)) \} \).

**Claim 2.8.** For every \( v \in E \), \( |\mathcal{L}_v| > (1 - 4r^2 \epsilon^{1/2}) d(v) \).

**Proof of Claim 2.8** First we show that for every \( v \in E \)

\[
\sum_{S \in L(v)} \sigma(S) > (1 - 4r^2 \epsilon) d(v) (\hat{\sigma} - d(v)) . \tag{6}
\]

Suppose that (6) is not true and fix \( u \in E \) with \( \sum_{S \in L(u)} \sigma(S) \leq (1 - 4r^2 \epsilon) d(u) (\hat{\sigma} - d(u)) \). Then

\[
\sum_{v \in E} \sum_{S \in L(v)} \sigma(S) = \sum_{S \in L(u)} \sigma(S) + \sum_{v \in E \setminus \{u\}} \sum_{S \in L(v)} \sigma(S) \\
\leq (1 - 4r^2 \epsilon) d(u) (\hat{\sigma} - d(u)) + \sum_{v \in E \setminus \{u\}} d(v) (\hat{\sigma} - d(v)) \\
\leq (1 - 2r \epsilon) \sum_{v \in E} d(v) (\hat{\sigma} - d(v)) - 2r \epsilon \left( 2r d(u) (\hat{\sigma} - d(u)) - \sum_{v \in E} d(v) (\hat{\sigma} - d(v)) \right) \\
\text{Claim } \text{< from } (5) \text{ > } (1 - 2r \epsilon) \sum_{v \in E} d(v) (\hat{\sigma} - d(v)) \\
\leq (1 - 2r \epsilon) \left( \sum_{v \in E} d(v) \right) \left( \hat{\sigma} - \frac{\sum_{v \in E} d(v)}{r} \right) = (1 - 2r \epsilon) \frac{r-1}{r} (\hat{\sigma})^2 ,
\]

and it follows from (6) that

\[
|\mathcal{H}| \leq \frac{1}{r(r-1)^{1/2}} |\partial \mathcal{H}|^{\frac{2}{r-1}} \left( \sum_{v \in E} \sum_{S \in L(v)} \sigma(S) + \sum_{S \in \partial \mathcal{H} \cup \cup_{v \in E} L(v)} \sigma(S) \right)^{\frac{1}{r-1}} \\
< \frac{1}{r(r-1)^{1/2}} |\partial \mathcal{H}|^{\frac{2}{r-1}} \left( (1 - 2r \epsilon) \frac{r-1}{r} (\hat{\sigma})^2 + (|\mathcal{H}| - \hat{\sigma}) \hat{\sigma} \right)^{\frac{1}{r-1}} \\
< \frac{1}{r(r-1)^{1/2}} |\partial \mathcal{H}|^{\frac{2}{r-1}} \left( \left( |\mathcal{H}| - \frac{1}{r} + 2(r-1) \epsilon \right) \hat{\sigma} \right)^{\frac{1}{r-1}} \\
\leq (1 - 2r \epsilon) \frac{1}{r} \left( \frac{|\partial \mathcal{H}|}{r} \right)^{\frac{1}{r-1}} < (1 - \epsilon) \left( \frac{|\partial \mathcal{H}|}{r} \right)^{\frac{1}{r-1}} ,
\]

8
a contradiction. Therefore, (6) holds for all \( v \in E \), and it follows from Lemma 2.2 that

\[
|L_v| > d(v) - \frac{\hat{\sigma} - d(v) - \sum_{S \in L(v)} \sigma(S)}{d(v)} \frac{1}{\varepsilon} (\hat{\sigma} - d(v)) d(v)
\]

\[
> d(v) - \frac{\hat{\sigma} - d(v) - (1-4r^2\epsilon^2)d(v)(\hat{\sigma} - d(v))}{\varepsilon/2 (\hat{\sigma} - d(v))} d(v) > d(v) - 4r^2\epsilon^2d(v).
\]

\[\]

Let \( \mathcal{G} = \{ S \in \partial \mathcal{H} : \sigma(S) > \left(\frac{r-1}{r} - 2r^1/2\epsilon^1/2\right)|\partial \mathcal{H}| \}. \]

**Claim 2.9.** \( |\mathcal{G}| > (1 - 8r^2\epsilon^1/2)|\partial \mathcal{H}| \).

**Proof of Claim 2.9.** By definition, for every \( v \in E \) and \( S \in L_v \),

\[
\sigma(S) > \left(1 - \epsilon^{1/2}\right)(\hat{\sigma} - d(v)) \quad \text{Claim 2.11} \quad \left(1 - \epsilon^{1/2}\right) \left(\frac{r-1}{r} - (2\epsilon)^{1/2}\right) \hat{\sigma}
\]

\[
\geq \left(1 - \epsilon^{1/2}\right) \left(\frac{r-1}{r} - (2\epsilon)^{1/2}\right) (1 - 2\epsilon)|\partial \mathcal{H}| \quad \text{Claim 2.12} \quad \geq \left(\frac{r-1}{r} - 2r^{1/2}\epsilon^{1/2}\right)|\partial \mathcal{H}|.
\]

Therefore, by Lemma 2.2,

\[
|\mathcal{G}| \geq \sum_{v \in E} |L_v| \quad \text{Claim 2.8} \quad \sum_{v \in E} (1 - 4r^2\epsilon^1/2)d(v) = (1 - 4r^2\epsilon^1/2)\hat{\sigma} \quad \text{Claim 2.13} \quad > (1 - 8r^2\epsilon^1/2)|\partial \mathcal{H}|.
\]

Let \( \Delta(\mathcal{H}) = \max \{ d(v) : v \in V(\mathcal{H}) \} \) be the maximum degree of \( \mathcal{H} \).

**Claim 2.10.** \( \Delta(\mathcal{H}) < \left(\frac{1}{r} + 2r^{1/2}\epsilon^{1/2}\right)|\partial \mathcal{H}| \).

**Proof of Claim 2.10.** Suppose this is not true and let \( u \in V(\mathcal{H}) \) such that

\[
d(u) = \Delta(\mathcal{H}) \geq \left(\frac{1}{r} + 2r^{1/2}\epsilon^{1/2}\right)|\partial \mathcal{H}|.
\]

Then, for every \( S \in L(u) \),

\[
\sigma(S) \leq \hat{\sigma} - d(u) \leq |\partial \mathcal{H}| - \left(\frac{1}{r} + 2r^{1/2}\epsilon^{1/2}\right)|\partial \mathcal{H}| = \left(\frac{r-1}{r} - 2r^{1/2}\epsilon^{1/2}\right)|\partial \mathcal{H}|.
\]

Therefore, \( L(u) \cap \mathcal{G} = \emptyset \), and hence

\[
|\mathcal{G}| \leq |\partial \mathcal{H}| - |d(u)| < \frac{r-1}{r}|\partial \mathcal{H}| < (1 - 8r^2\epsilon^1/2)|\partial \mathcal{H}|,
\]

which contradicts Claim 2.9.

Let \( U = \partial_{r-2} \mathcal{G} \) and note that \( U \subset V(\mathcal{H}) \).

**Claim 2.11.** \( |U| < (1 + 2r^{5/2}\epsilon^{1/2}) r^{-1/4} |\partial \mathcal{H}|^{1/4} \).

9
Proof of Claim 2.11. First we show that for every \( v \in U \),

\[
d(v) \geq \left( \frac{1}{r} - 2r^{3/2} \epsilon^{1/2} \right) |\partial H|. \tag{7}
\]

Suppose that there exists \( u \in U \) such that (7) is not true for \( u \). Then choose a set \( S \in \mathcal{G} \) such that \( u \in S \). By the definition of \( \mathcal{G} \),

\[
\sigma(S) > \left( \frac{r-1}{r} - 2r^{1/2} \epsilon^{1/2} \right) |\partial H|,
\]

so by Pigeonhole principle, there exists \( u' \in S \setminus \{u\} \) such that

\[
d(u') \geq \sigma(S) - d(u) > \left( \frac{1}{r} + 2r^{1/2} \epsilon^{1/2} \right) |\partial H|,
\]

which contradicts Claim 2.10. Therefore, (7) holds for all \( v \in U \), and it follows from

\[
\sum_{v \in U} d(v) \leq r|H| \text{ and Theorem 1.5 that}
\]

\[
|U| \leq \frac{r^2}{r-2} \leq \frac{r^2}{r-2} \left( \frac{\epsilon}{r-1} \right)^{r-1} \frac{1}{r-2} < \left( 1 + 2r^{5/2} \epsilon^{1/2} \right) r^{-1} |\partial H|^{r^{-1}}.
\]

Let \( \hat{\mathcal{G}} = \{ G \in H : \partial G \subset \mathcal{G} \} \). Recall from Lemma 2.5 that \( |\partial H| > cn^{r-1} \) and \( c > 0 \) is a constant.

Claim 2.12. \( |\hat{\mathcal{G}}| \geq \left( 1 - \frac{16r^4 \epsilon^{1/2}}{c^{1/(r-1)}} \right) |H|. \)

Proof of Claim 2.12. By assumption,

\[
|H| > (1 - \epsilon) \left( \frac{|\partial H|}{r} \right)^{r^{-1}} > (1 - \epsilon) \left( \frac{cn^{r-1}}{r} \right)^{r^{-1}},
\]

which implies that

\[
|\partial H| > r \left( \frac{|H|}{1-\epsilon} \right)^{r^{-1}}, \tag{8}
\]

and

\[
n > \frac{r^{-1}}{c^{1/(r-1)}} \left( \frac{|H|}{1-\epsilon} \right)^{1/r}. \tag{9}
\]

Therefore,

\[
|\hat{\mathcal{G}}| = |\mathcal{G}| - |\mathcal{G} \setminus \hat{\mathcal{G}}| > |H| - |\partial H| \geq |H| - 8r^2 \epsilon^{1/2} n |\partial H| \geq |H| - 8r^2 \epsilon^{1/2} n \left( \frac{r^{-1}}{c^{1/(r-1)}} \right) |H| \left( \frac{1}{1-\epsilon} \right)
\]

\[
> |H| - \frac{16r^4 \epsilon^{1/2}}{c^{1/(r-1)}} |H|.
\]

\[\blacksquare\]
Claim 2.13. $|U| > \left(1 - \frac{32r^4\epsilon^{1/2}}{c^1/(r-1)}\right) r^{-\frac{r}{r-2}} |\partial\mathcal{H}|^\frac{1}{r-1}$.

Proof of Claim 2.13. It follows from Claims 2.10 and 2.12 and \(\sum_{u \in U} d_H(u) \geq r|\hat{G}|\) that

\[
|U| \geq \frac{r|\hat{G}|}{\Delta(H)} > \frac{r \left(1 - \frac{16r^4\epsilon^{1/2}}{c^1/(r-1)}\right) |H|}{\left(\frac{1}{2} + 2r^{-1/2}\epsilon^{1/2}\right) |\partial H|} > \frac{r \left(1 - \frac{16r^4\epsilon^{1/2}}{c^1/(r-1)}\right) (1 - \epsilon) \left(\frac{|\partial H|}{r}\right)^{\frac{r}{r-2}}}{\left(\frac{1}{2} + 2r^{-1/2}\epsilon^{1/2}\right) |\partial H|} > \left(1 - \frac{32r^4\epsilon^{1/2}}{c^1/(r-1)}\right) r^{-\frac{r}{r-2}} |\partial H|^{\frac{1}{r-1}}.
\]

Now we are ready to finish the proof of Lemma 2.5. First, Claims 2.11 and 2.13 imply that

\[
\left(1 - \frac{32r^4\epsilon^{1/2}}{c^1/(r-1)}\right) r^{-\frac{r}{r-2}} |\partial H|^{\frac{1}{r-1}} < |U| < \left(1 + 2r^{5/2}\epsilon^{1/2}\right) r^{-\frac{r}{r-2}} |\partial H|^{\frac{1}{r-1}},
\]

On the other hand, since \(G \subset \partial(H[U])\) and \(\hat{G} \subset H[U]\), it follows from Claim 2.9 that

\[
|\partial(H[U])| \geq |G| > \left(1 - 8r^2\epsilon^{1/2}\right) |\partial H|,
\]

and it follows from Claim 2.12 that

\[
|H[U]| \geq |\hat{G}| > \left(1 - \frac{16r^4\epsilon^{1/2}}{c^1/(r-1)}\right) |H|.
\]

Let \(n' = |U|\) and \(H' = H[U]\). Then Lemma 2.5 implies that

\[
\left(1 - 4r^{7/2}\epsilon^{1/2}\right) \frac{(n')^{r-1}}{r^{r-2}} < |\partial H'| < \left(1 + \frac{32r^5\epsilon^{1/2}}{c^1/(r-1)}\right) \frac{(n')^{r-1}}{r^{r-2}},\tag{10}
\]

and

\[
|H'| \geq \left(1 - \frac{16r^4\epsilon^{1/2}}{c^1/(r-1)}\right) |H| > \left(1 - \frac{32r^5\epsilon^{1/2}}{c^1/(r-1)}\right) (1 - \epsilon) \left(\frac{|\partial H|}{r}\right)^{\frac{r}{r-2}} \geq \left(1 - \frac{32r^5\epsilon^{1/2}}{c^1/(r-1)}\right) (1 - \epsilon) \left(\frac{|\partial H'|}{r}\right)^{\frac{r}{r-2}} > \left(1 - \frac{40r^5\epsilon^{1/2}}{c^1/(r-1)}\right) \frac{(n')^{r}}{r^{r}}.\tag{11}
\]

On the other hand, by Theorem 1.5,

\[
|H'| \leq \left(\frac{|\partial H'|}{r}\right)^{\frac{r}{r-2}} < \left(1 + \frac{64r^5\epsilon^{1/2}}{c^1/(r-1)}\right) \frac{(n')^{r}}{r^{r}}.
\tag{12}
\]

Therefore, \((d(\partial H'), d(H')) \in B_{T_r}^{*}(\frac{(r-1)^{\frac{r}{r-2}}}{r^{r-2}}, \frac{c}{r})\), where \(c' = \frac{128r^5\epsilon^{1/2}}{c^1/(r-1)}\).

Our next lemma shows that if an \(r\)-graph \(H\) satisfies \((d(\partial H), d(H)) \in B_{T_r}^{*}(\frac{(r-1)^{\frac{r}{r-2}}}{r^{r-2}}, \frac{c}{r})\) for some sufficiently small \(\epsilon > 0\), then \(H\) is structurally close to \(T_r(n, r)\). This will be used to show that \(ed(H', T_r(n', r))\) is small.
Lemma 2.14. Let $r \geq 2$. For every $\delta > 0$ there exists $\epsilon > 0$ and $n_0$ such that the following holds for all $n \geq n_0$. Suppose that $\mathcal{H}$ is a cancellative $r$-graph on $n$ vertices with

$$
\frac{n^{r-1}}{r^{r-2}} - cn^{r-1} < |\partial \mathcal{H}| < \frac{n^{r-1}}{r^{r-2}} + cn^{r-1} \text{ and } |\mathcal{H}| > (1 - \epsilon) \left( \frac{|\partial \mathcal{H}|}{r} \right)^\frac{r-2}{r-1}.
$$

Then $ed(\mathcal{H}, T_r(n, r)) < \delta n^r$.

Proof. The proof of this lemma is by induction on $r$. When $r = 2$, this is Simonovits’ stability theorem [18]. So we may assume that $r \geq 3$.

Let

$$
V_L = \left\{ v \in V(\mathcal{H}) : d(v) > (1 - \epsilon^{1/2}) \left( \frac{|\partial L(v)|}{r - 1} \right)^\frac{r-2}{r-1} \right\},
$$

\[\hat{V}_L = \left\{ v \in V(\mathcal{H}) : d(v) > \left( \frac{1}{r} - 3r^{3/2} \epsilon^{1/2} \right) |\partial \mathcal{H}| \right\},\]

$V_S = V(\mathcal{H}) \setminus V_L$, and $\hat{V}_S = V(\mathcal{H}) \setminus \hat{V}_L$. Note that for every $v \in V_S$,

$$
|\partial L(v)| \geq \frac{(r - 1)(d(v))^{\frac{r-2}{r-1}}}{(1 - \epsilon^{1/2})^{\frac{r-2}{r-1}}}.
$$

Claim 2.15. $|\hat{V}_L| > (1 - 64r^5 \epsilon^{1/2}) n$, and hence $|\hat{V}_S| < 64r^5 \epsilon^{1/2} n$.

Proof of Claim 2.15. Since $|\mathcal{H}| > (1 - \epsilon)(|\partial \mathcal{H}|/r)^{(r-1)/(r-2)}$ and $|\partial \mathcal{H}| > (1/r^{r-2} - \epsilon)n^{r-1}$, it follows from Lemma 2.13 and the proof of Claim 2.11 that there exists $U \subset V(\mathcal{H})$ with

$$
|U| > \left( 1 - \frac{32r^4 \epsilon^{1/2}}{(1/r^{r-2} - \epsilon)^{1/(r-1)}} \right) \left( \frac{r-2}{r-1} \right)^{\frac{r-2}{r-1}} |\partial \mathcal{H}|^{\frac{r-2}{r-1}} > \left( 1 - \frac{32r^4 \epsilon^{1/2}}{(1/r^{r-2} - \epsilon)^{1/(r-1)}} \right) \left( \frac{1}{r^{r-2} - \epsilon} \right)^{\frac{r-2}{r-1}} n > \left( 1 - 64r^5 \epsilon^{1/2} \right) n
$$

such that $d(v) \geq (1/r^{r-2} - 2 \epsilon^{3/2} \epsilon^{1/2})|\partial \mathcal{H}|$ for all $v \in U$. Therefore, $|\hat{V}_L| \geq |U| > (1 - 64r^5 \epsilon^{1/2}) n$, and hence $|V_S| = n - |\hat{V}_L| < 64r^5 \epsilon^{1/2} n$.

Claim 2.16. $|V_L| > (1 - 66r^5 \epsilon^{1/2}) n$.

Proof of Claim 2.16. It is easy to see that for every $v \in V(\mathcal{H})$, $L(v)$ is also cancellative, so by Theorem 1.5, $d(v) \leq (|\partial L(v)|/(r - 1))^{(r-1)/(r-2)}$. Therefore,

$$
|\mathcal{H}| = \frac{1}{r} \sum_{v \in V(\mathcal{H})} d(v)
= \frac{1}{r} \left( \sum_{v \in V_L} (d(v))^{\frac{1}{r-1}} (d(v))^{\frac{r-2}{r-1}} + \sum_{v \in V_S} (d(v))^{\frac{1}{r-1}} (d(v))^{\frac{r-2}{r-1}} \right)
\leq \frac{1}{r(r - 1)} \left( \sum_{v \in V_L} (d(v))^{\frac{1}{r-1}} |\partial L(v)| + (1 - \epsilon^{1/2})^{\frac{r-2}{r-1}} \sum_{v \in V_S} (d(v))^{\frac{1}{r-1}} |\partial L(v)| \right)
= \frac{1}{r(r - 1)} \sum_{v \in V(\mathcal{H})} (d(v))^{\frac{1}{r-1}} |\partial L(v)| - \frac{1 - (1 - \epsilon^{1/2})^{\frac{r-2}{r-1}}}{r(r - 1)} \sum_{v \in V_S} (d(v))^{\frac{1}{r-1}} |\partial L(v)|,
$$
which together with (13) gives
\[
|\mathcal{H}| \leq \frac{1}{r} \frac{1}{|v(v)\mathcal{H}|} \sum_{v \in V(H)} (d(v))^{\frac{r}{r-1}} |\partial L(v)| - \frac{1 - (1 - \epsilon^{1/2})^{\frac{r}{r-1}}}{1 - \epsilon^{1/2}} \frac{1}{r} \sum_{v \in V(G)} d(v)
\]
\[
\leq \frac{1}{r} \frac{1}{|v(v)\mathcal{H}|} \sum_{v \in V(H)} (d(v))^{\frac{r}{r-1}} |\partial L(v)| - \epsilon^{1/2} \frac{1}{2r} \sum_{v \in V(G)} d(v)
\]
\[
\leq \left( \frac{|\partial |}{r} \right)^{\frac{r}{r-1}} - \frac{1}{2r} \sum_{v \in V(G)} d(v).
\]
Since $|\mathcal{H}| > (1 - \epsilon) (|\partial|)^{r/(r-1)}$, the inequality above implies
\[
\left( \frac{|\partial |}{r} \right)^{\frac{r}{r-1}} - \epsilon |\partial|^{\frac{r}{r-1}} \leq \left( \frac{|\partial |}{r} \right)^{\frac{r}{r-1}} - \frac{1}{2r} \sum_{v \in V(G)} d(v)
\]
\[
< \left( \frac{|\partial |}{r} \right)^{\frac{r}{r-1}} - \epsilon \frac{1}{2r} \left( |V| - |\hat{V}| \right) \left( \frac{1}{r} - 2r^{3/2} \epsilon^{1/2} \right) |\partial|,
\]

Claim 2.13 implies that $|V| < 66r^5 \epsilon^{1/2} n$. Therefore, $|V_L| = n - |V| > (1 - 66r^5 \epsilon^{1/2}) n$.

Claims 2.13 and 2.14 imply that $|V_L \cap \hat{V}_L| > (1 - 130r^5 \epsilon^{1/2}) n$, and since
\[
|\mathcal{H}| > (1 - \epsilon) \left( \frac{|\partial |}{r} \right)^{\frac{r}{r-1}} > (1 - \epsilon) \left( \frac{1}{r^{r-2} - \epsilon} n^{r-1} \right)^{\frac{r}{r-1}} > \frac{n^r}{2r^r},
\]
there exists $\hat{E} \in \mathcal{H}[V_L \cap \hat{V}_L]$. By the definition of $V_L$ and $\hat{V}_L$, for every $v \in \hat{E}$,
\[
d(v) > (1 - \epsilon^{1/2}) \left( \frac{|\partial L(v)|}{r} \right)^{\frac{r}{r-1}} - \frac{1}{r} \frac{1}{|v(v)\mathcal{H}|} \sum_{v \in V(G)} (d(v))^{\frac{r}{r-1}} |\partial L(v)|
\]
and
\[
d(v) > \left( \frac{1}{r} - 2r^{3/2} \epsilon^{1/2} \right) |\partial| > \left( \frac{1}{r} - 2r^{3/2} \epsilon^{1/2} \right) \left( \frac{1}{r^{r-2} - \epsilon} n^{r-1} \right)
\]
\[
> \left( \frac{1}{r^{r-1} - 4r^3 \epsilon^{1/2}} \right) n^{r-1}.
\]

On the other hand, since $\sum_{v \in \hat{E}} d(v) \leq |\partial|$, (13) implies that for all $v \in \hat{E}$,
\[
d(v) < |\partial| - (r - 1) \left( \frac{1}{r} - 3r^{3/2} \epsilon^{1/2} \right) |\partial| < \left( \frac{1}{r} + 2r^5 \epsilon^{1/2} \right) \left( \frac{1}{r^{r-2} + \epsilon} n^{r-1} \right)
\]
\[
< \left( \frac{1}{r^{r-1} + 4r^5 \epsilon^{1/2}} \right) n^{r-1}.
\]

Fix $v \in \hat{E}$ and let $\mathcal{H}_{v} = L(v)$, and note that $\mathcal{H}_{v}$ is a cancellative $(r - 1)$-graph. So (14) and Theorem 1.5 imply that
\[
(1 - \epsilon^{1/2}) \left( \frac{|\partial L_{v}|}{r - 1} \right)^{\frac{r}{r-1}} < |\mathcal{H}_{v}| \leq \left( \frac{|\partial L_{v}|}{r - 1} \right)^{\frac{r}{r-1}},
\]
and (16) and (10) give
\[ \left( \frac{1}{r-1} - 4r^{3/2} \epsilon^{1/2} \right) n^{r-1} < |\mathcal{H}_v| < \left( \frac{1}{r-1} + 4r^{5/2} \epsilon^{1/2} \right) n^{r-1}. \]  
(18)

(17) and (18) imply that
\[ \left( 1 - 8r^{r+1/2} \epsilon^{1/2} \right) \frac{r-1}{r} n^{r-2} < |\partial \mathcal{H}_v| < \left( 1 + 8r^{r+3/2} \epsilon^{1/2} \right) \frac{r-1}{r} n^{r-2}. \]  
(19)

Since \( \mathcal{H}_v \) is also cancellative, by (17) and Lemma 2.3 there exists \( U_v \subset V(\mathcal{H}_v) \subset V(\mathcal{H}) \) with
\[ \left( 1 - 16r^{r+3/2} \epsilon^{1/2} \right) \frac{r-1}{r} n < |U_v| < \left( 1 + 16r^{r+3/2} \epsilon^{1/2} \right) \frac{r-1}{r} n, \]  
(20)
such that
\[ \left( 1 - 8r^2 \epsilon^{1/2} \right) |\partial \mathcal{H}_v| < |\partial (\mathcal{H}_v[U_v])| \leq |\partial \mathcal{H}_v|, \]  
(21)
and
\[ \left( 1 - 16r^5 \epsilon^{1/2} \right) |\mathcal{H}_v| < |\mathcal{H}_v[U_v]| < |\mathcal{H}_v|. \]  
(22)

Let \( n_v = |U_v| \) and \( \mathcal{H}_v' = \mathcal{H}_v[U_v] \). Then, (19), (20) and (21) imply that
\[ \left( 1 - 32r^{r+3/2} \epsilon^{1/2} \right) \frac{n_v^{r-2}}{(r-1)^{r-3}} < |\partial \mathcal{H}_v'| < \left( 1 + 32r^{r+3/2} \epsilon^{1/2} \right) \frac{n_v^{r-2}}{(r-1)^{r-3}}, \]
and (17), (21) and (22) imply that
\[ |\mathcal{H}_v'| > \left( 1 - 32r^5 \epsilon^{1/2} \right) \left( \frac{|\partial \mathcal{H}_v'|}{r-1} \right)^{\frac{r-1}{r}}. \]

Therefore, by the induction hypothesis, there is a sufficiently small \( \delta' > 0 \) (and we may assume that \( \delta' \geq \epsilon^{1/2} \)) such that
\[ \text{ed} (\mathcal{H}_v', T_{r-1} (n_v, r-1)) < \delta' n_v^{r-1}. \]

In other words, there exists a partition \( U_v = V_1 \cup \cdots \cup V_{r-1} \) such that all but at most \( \delta' n_v^{r-1} \) sets in \( \mathcal{H}_v' \) have exactly one vertex in each \( V_i \). Let
\[ \mathcal{K}_v = \{ A \in \mathcal{H}_v : |A \cap V_i| = 1 \text{ for all } i \in [r-1] \}, \]
and it is easy to see that
\[ |\mathcal{K}_v| > |\mathcal{H}_v'|-\delta' n_v^{r-1} \geq |\mathcal{H}_v| - (8r^2 \epsilon^{1/2} + \delta') n^{r-1}. \]  
(23)

For every \( S \subset V(\mathcal{H}) \) let \( N(S) = \{ u \in V(\mathcal{H}) \setminus S : \exists A \in \mathcal{H} \text{ such that } \{ u \} \cup S \subset A \}. \)

**Claim 2.17.** For every \( S \in \mathcal{K}_v, N(S) \cap U_v = \emptyset \), i.e. \( N(S) \subset V(\mathcal{H}) \setminus U_v. \)

**Proof.** Suppose this is not true. Let \( S \in \mathcal{K}_v \) such that \( N(S) \cap U_v \neq \emptyset \) and \( u \in N(S) \cap U_v. \) Since \( \{ v \} \cup \mathcal{H} \cap \mathcal{H} \neq 0 \) and \( \{ u \} \cup S \in \mathcal{H} \) and \( \{ v \} \cup S \cap N(\mathcal{H}) \neq \emptyset \), \( \{ u, v \} \) is not contained in any edge of \( \mathcal{H} \). However, this is a contradiction, since \( u \in U_v \subset N(v) \). \( \blacksquare \)

Let \( S_L = \{ S \in \partial \mathcal{H} : |N(S)| > \left( \frac{1}{2} - (\delta')^{1/2} \right) n \}. \)
Claim 2.18. \( |S_L| > (1 - 32r^{r+3/2}e^{1/4}) \frac{n^{r-1}}{r-2} \).

Proof. By Claim 2.17 for every \( u \in \hat{E} \), all but at most \( (8r^2e^{1/2} + \delta')n^{r-1} \) sets \( S \in L(u) \) satisfies \( |N(S)| \leq |V(H) \setminus u| < (1/r + 16r^{r+3/2}e^{1/2})n \). Therefore, all but at most

\[
r(8r^2e^{1/2} + \delta')n^{r-1} + |\partial H \setminus \bigcup_{u \in E} L(u)| < (40r^5e^{1/2} + \delta')n^{r-1} \tag{24}
\]

edges \( S \in \partial H \) satisfy \( N(S) \leq (1/r + 16r^{r+3/2}e^{1/2})n \). It follows that

\[
r|H| = \sum_{S \in \partial H} N(S)
= \sum_{S \in S_L} N(S) + \sum_{S \in \partial H \setminus S_L} N(S)
\leq |S_L| \left( \frac{1}{r} + 16r^{r+3/2}e^{1/2} \right) n + |\partial H \setminus S_L| \left( \frac{1}{r} - (\delta')^{1/2} \right) n + (40r^5e^{1/2} + \delta')n^r
< |\partial H| \left( \frac{1}{r} - (\delta')^{1/2} \right) n + |S_L| \left( 16r^{r+3/2}e^{1/2} + (\delta')^{1/2} \right) n + (40r^5e^{1/2} + \delta')n^r
< \left( \frac{1}{r^{r-1}} + \frac{\epsilon}{r} - \frac{(\delta')^{1/2}}{r^{r-2}} \right) n^r + |S_L| \left( 16r^{r+3/2}e^{1/2} + (\delta')^{1/2} \right) n + (40r^5e^{1/2} + \delta')n^r.
\]

On the other hand, \( |\partial H| > (1 - \epsilon) \left( |\partial H|/r \right)^{r/(r-1)} > n^r/r^r - \epsilon n^r \). Therefore,

\[
|S_L| > \frac{(\delta')^{1/2}/r^{r-2} - \epsilon/r - \epsilon}{(\delta')^{1/2} + 16r^{r+3/2}e^{1/2}} n^r > \frac{1 - 32r^{r+3/2}e^{1/4}}{r^{r-2}} n^{r-1}.
\]

Here we used \( \delta' \geq e^{1/2} \).

Now we are ready to finish the proof of Lemma 2.14. Let \( \delta = 256r^{5/2}(\delta')^{1/2} \). Recall that we already have a partition \( U_v = V_1 \cup \cdots \cup V_{r-1} \). Let \( V_r = V(H) \setminus U_v \), and we claim that all but at most \( \delta n^r \) edges in \( H \) have exactly one vertex in each \( V_i \). Indeed, by Claim 2.18 all but at most

\[
|\partial H| - |S_L| < \left( \frac{1}{r^{r-2}} + \epsilon \right) n^{r-1} - \left( 1 - 32r^{r+3/2}e^{1/4} \right) \frac{n^{r-1}}{r^{r-2}} < 64r^5/2e^{1/4}n^{r-1}
\]

sets \( S \in K \) satisfying \( |N(S)| > (1/r - (\delta')^{1/2})n \). It follows from Claim 2.17 that at least

\[
\left( |\mathcal{K}_v| - 64r^{5/2}e^{1/4}n^{r-1} \right) \left( \frac{1}{r} - (\delta')^{1/2} \right) n
\geq \left( |\mathcal{K}_v| - (8r^2e^{1/2} + \delta')n^{r-1} - 64r^{5/2}e^{1/4}n^{r-1} \right) \left( \frac{1}{r} - (\delta')^{1/2} \right) n.
\]

\[
\geq \left( \frac{1}{r^{r-1}} - 4r^{3/2}e^{1/2} - (8r^2e^{1/2} + \delta') - 64r^{5/2}e^{1/4} \right) \left( \frac{1}{r} - (\delta')^{1/2} \right) n^r.
\]

\[
> \left( \frac{1}{r^r} - 128r^{5/2}(\delta')^{1/2} \right) n^r > |H| - \delta n^r
\]

edges in \( H \) have exactly one vertex in each \( V_i \). Here we used \( |H| \leq (|\partial H|/r)^{r/(r-1)} < n^r/r^r + \epsilon n^r \) in the last inequality.

Now we are ready to prove Theorem 2.1.


Proof of Theorem 2.1. Let $\mathcal{H}$ be a cancellative $r$-graph on $n$ vertices that satisfies assumptions in Theorem 2.1. Let $\delta$ be obtained from Lemma 2.1.4 by replacing $\epsilon$ with $\frac{128r^5\epsilon^{1/2}}{c_1/r(r-1)}$. Let $\delta = \frac{64r^5\epsilon^{1/2}}{c_1/r(r-1)} + \delta$. Let $x = |\partial\mathcal{H}|(\frac{n}{x-1})$ and $\alpha = (x^r - 2)/(r-1)!$. First, by Lemma 2.3, there exists $U \subset V(\mathcal{H})$ such that $\mathcal{H}' := \mathcal{H}$ and $n' := |U|$ satisfy (10), (11), and (12), and

\[ |an - n'| < \left( \frac{r^2 - 2|\partial\mathcal{H}|}{(\frac{n}{x-1})^2 (r-1)!} \right)^{1/16} n - r^{-2} |\partial\mathcal{H}|^\frac{1}{r-1} + \frac{16r^4\epsilon^{1/2}}{c_1/(r-1)} r^{-1} |\partial\mathcal{H}|^{-1} < \frac{32r^5\epsilon^{1/2}}{c_1/(r-1)} n. \]

Then, applying Lemma 2.1.4 to $\mathcal{H}'$, we obtain $\text{ed}(\mathcal{H}', T_r(n', r)) < \delta n^r$. Therefore,

\[
\text{ed}(\mathcal{H}, \mathcal{H}'(n, x)) \leq |\mathcal{H}| - |\mathcal{H}'| + \text{ed}(\mathcal{H}', T_r(|U|, r)) + |an - n'|n^{-1} < \frac{16r^4\epsilon^{1/2}}{c_1/(r-1)} |\mathcal{H}| + \delta n^r + \frac{32r^5\epsilon^{1/2}}{c_1/(r-1)} n^{-1} < \delta n^r.
\]

\[ \square \]

3 $s_k$-stable points in $\partial\Omega(T_3)$

In this section we prove Theorems 1.15 and 1.16.

Let $G$ be a graph. The clique number $\omega(G)$ of $G$ is the largest integer $\omega$ such that there is a copy of $K_\omega$ in $G$. For an $r$-graph $\mathcal{H}$ and $S \subset V(\mathcal{H})$, we use $\mathcal{H}[S]$ to denote the induced subgraph of $\mathcal{H}$ on $S$. The following results will be used in our proofs.

Theorem 3.1 (Graph removal lemma, see [5]). For every graph $F$ and every $\epsilon > 0$ there exists $\delta > 0$ such that every graph on $n$ vertices which contains at most $\delta n^{\omega(F)}$ copies of $F$ can be made $F$-free by removing at most $\epsilon n^2$ edges.

Theorem 3.2 (Stability of $K_{\ell+1}$, see [8]). Let $\ell \geq 2$ and $G$ be a $K_{\ell+1}$-free graph with $n$ vertices and $t_2(n, \ell) - m$ edges for some $m \geq 0$. Then $G$ contains an $\ell$-partite subgraph with at least $t_2(n, \ell) - 2m$ edges.

Theorem 3.3 ([12]). Let $\mathcal{H}$ be a cancellative 3-graph on $n$ vertices. Let $U \subset V(\mathcal{H})$ be a set of size $m$. Suppose that $|(\partial\mathcal{H})[U]| = xm^2/2$ for some real number $x$ with $0 \leq x \leq (m - 1)/m$. Then

\[ |\mathcal{H}[U]| \leq \frac{(1 - x)x}{6}m^3 + 3m^2. \]

The following algorithm will be used in the proofs.

Algorithm 1 (Withdraw cliques with threshold $\kappa$)

- **Input:** A 3-graph $\mathcal{H}$ and an integer $\kappa$.
- **Initial step:** Let $\mathcal{H}_0 = \mathcal{H}$, $G_0 = G = \partial\mathcal{H}$, and $\omega_1 = \omega(G_0)$. If $\omega_1 < \kappa$, then terminate this algorithm. Otherwise, we repeat the following operation.
- **Iteration:** For $i \geq 1$, if $\omega_i < \kappa$ or $G_{i-1} = \emptyset$, then terminate this process. Otherwise, choose $S_i \subset V(G_{i-1})$ with $|S_i| = \omega_i$ such that $G_{i-1}[S_i] \cong K_{\omega_i}$. Let $T_i = V(G_{i-1}) \setminus S_i$, $G_i = G_{i-1}[T_i]$, and $\mathcal{H}_i = \mathcal{H}[T_i]$.
- **Output:** A descending chain of induced subgraphs of $\mathcal{H}$ for some $t \geq 0$,

\[ \mathcal{H} = \mathcal{H}_0 \supset \mathcal{H}_1 \supset \cdots \supset \mathcal{H}_t. \]
Lemma 3.4. Let $\mathcal{H}$ be a cancellative 3-graph on $n$ vertices. Applying Algorithm 1 to $\mathcal{H}$ with threshold $\kappa$ and suppose that it stops after $t$ steps and we obtain a sequence of induced subgraphs of $\mathcal{H}$, namely, $\mathcal{H} = \mathcal{H}_0 \supset \mathcal{H}_1 \supset \cdots \supset \mathcal{H}_t$. For every $i \in [t]$ let $e_i$ denote the number of edges in $G_i$ that have at least one vertex in $S_i$, $W_i = \sum_{j=1}^i \omega_j$, and $E_i = \sum_{j=1}^i e_j$. Then, the following inequalities hold.

$$|\mathcal{H}| \leq |\mathcal{H}_t| + t|\partial \mathcal{H}|.$$  \hfill (25)

$$E_i \leq (W_i - i) n.$$  \hfill (26)

Proof. It follows from Lemma 3.17 that $|\mathcal{H}_{i-1}| \leq |\mathcal{H}_i| + |G_{i-1}|$ for all $i \in [t]$. So

$$|\mathcal{H}| \leq |\mathcal{H}_t| + \sum_{i=0}^{t-1} |G_i| \leq |\mathcal{H}_t| + t|\partial \mathcal{H}|.$$  

On the other hand, by Algorithm 1, for every $i \in [t]$, every vertex in $T_i$ is adjacent to at most $\omega_i - 1$ vertices in $S_i$. Therefore, $e_i \leq |T_i| (\omega_i - 1) + \binom{\omega_i}{2}$, and hence

$$E_i = \sum_{j=1}^i e_j \leq \sum_{j=1}^i \left( |T_i| (\omega_i - 1) + \binom{\omega_i}{2} \right) \leq \sum_{j=1}^i (\omega_i - 1) n = (W_i - i) n.$$  

3.1 Stability result

In this section we prove the following statement, which implies Theorem 1.15.

Theorem 3.5. Let $k \geq 6N + \{1, 3\}$ and $k \geq 3$. For every $\delta > 0$ there exists an $\epsilon > 0$ and $n_0$ such that the following holds for all $n \geq n_0$. Suppose that $\mathcal{H}$ is a cancellative 3-graph on $n$ vertices with $|\partial \mathcal{H}| \geq (1 - \epsilon)(k - 1)n^2/(2k)$ and $|\mathcal{H}| \geq (1 - \epsilon)(k - 1)n^3/(6k^2)$. Then, $\mathcal{H}$ can be transformed into a subgraph of a 3-graph in $S(n, k)$ by removing at most $\delta n^3$ edges.

The proof of Theorem 3.5 is consist of the following steps. First we show that the number of copies of $K_{k+1}$ in $\partial \mathcal{H}$ is very small, and by Theorem 3.1 we can get a $K_{k+1}$-free graph $\partial \mathcal{H}'$ from $\partial \mathcal{H}$ by removing very few edges. Since $|\partial \mathcal{H}'|$ is still very close to $t_2(n, k)$, by Theorem 3.2 the structure of $\partial \mathcal{H}'$ is very close to the Turán graph $T_2(n, k)$, and so is $\partial \mathcal{H}$. The final step is to show that the structure of $\mathcal{H}$ is close to a 3-graph in $S(n, k)$ using the structure of $\partial \mathcal{H}$.

Proof of Theorem 3.5. We will prove Theorem 3.5 through a series of claims, and we will omit the floor and ceiling signs when they are not crucial in the proof. Let $G = \partial \mathcal{H}$ and $e = |G|$. First, we give an upper bound for $e$ in the following claim.
Claim 3.6. \( e < \left( \frac{k-1}{k} + 4\epsilon \right) n^2 \).

Proof. It follows from our assumption and Theorem 3.3 that
\[
(1 - e)\frac{(k-1)n^3}{6k^2} < |S| \leq \frac{e(n^2 - 2e)}{3n} + 3n^2.
\]
Since \( k \geq 3 \) and \( n \) is sufficiently large, the inequality above implies that
\[
e < \frac{1}{4} \left( 1 + \left( \frac{k - 2}{k} \right)^2 + 4\epsilon \left( \frac{k - 1}{k^2} + \frac{72}{n} \right)^{1/2} \right) n^2 < \left( \frac{k - 1}{k} + 4\epsilon \right) n^2.
\]

Our next claim gives an upper bound for the clique number of \( G \).

Claim 3.7. \( \omega(G) < 10k\epsilon n \).

Proof of Claim 3.7. Let \( a = 10k\epsilon \) and suppose that \( \omega(G) \geq 10k\epsilon n \). Then choose \( S \subset V(H) \) with \(|S| = an\) such that \( G[S] \cong K_{an} \). Let \( T = V(H) \setminus S \) and let \( e_s \) denote the number of edges in \( G \) that have at least one vertex in \( S \) and notice that \( e_s < an^2 \). Let \( x' = 2(e - e_s)/((a - 1)\pi^2) \). Then by Lemma 2.2 and Theorem 3.3,
\[
|S| = |H| + \sum_{v \in S} d(v) \\
\leq \frac{(1 - x')x'}{6}((1 - a)^3n^3 + 3(1 - a)^2n^2 + e) \\
= \left( \frac{(1 - a)^2n^2 - 2(e - e_s)(e - e_s)}{3(1 - a)n} \right) + 3(1 - a)^2n^2 + e \\
= \frac{-2e_s^2 + (4 - (1 - a)^2n^2)e_s + (1 - a)^2n^2e - 2e_s^2 + 3(1 - a)^2n^2 + e}{3(1 - a)n}.
\]

Since \(-2e_s^2 + (4 - (1 - a)^2n^2)e_s\) is increasing in \( e_s \) when \( e_s \leq e - (1 - a)^2n^2/4 \) and
\[
e - \frac{(1 - a)^2n^2}{4} > (1 - e)\left( \frac{(k - 1)n}{2k} - \frac{(1 - a)^2n^2}{4} \right) \\
\geq (1 - e)\left( \frac{n^2}{3} - \frac{(1 - a)^2n^2}{4} \right) > an^2,
\]
we may substitute \( e_s = an^2 \) into (27) and obtain
\[
|S| \leq \frac{-2e^2 + ((1 + a)^2n^2 + 3(1 - a)n)e - (1 + a^2)an^4}{3(1 - a)n} + 3(1 - a)^2n^2.
\]

Since \(-2e^2 + ((1 + a^2)n^2 + 3(1 - a)n)e\) is decreasing in \( e \) when \( e \geq (1 + a^2)n^2/4 + 3(1 - a)n/4 \) and
\[
\frac{(1 + a^2)n^2}{4} + \frac{3(1 - a)n}{4} < \frac{n^2}{4} + \frac{n^2}{100} < (1 - e)\frac{n^2}{3} < (1 - e)\frac{(k - 1)n^2}{2k},
\]
we may substitute \( e = (1 - e)(k - 1)n^2/(2k) \) into (28) and obtain
\[
|S| < \frac{(1 - e)(k - 1) - 2ka + ka^2 + (k - 1)e}{6(1 - a)k^2}n^3 + \left( 1 - e \right)\frac{(k - 1)2k + 3(1 - a)^2}{2k}n^2 \\
\leq \frac{(k - 1)(1 - a) - ka}{6(1 - a)k^2}n^3 + 4n^2 \\
\leq \frac{k - 1}{6k^2}n^3 + \frac{e}{6}n^3 - \frac{a}{6k}n^3 + 4n^2 < \frac{k - 1}{6k^2}n^3 - \epsilon n^3,
\]
a contradiction. \[ \blacksquare \]
The next step is to show that the number of copies of $K_{k+1}$ in $G$ is small. If $G$ is $K_{k+1}$-free, then there is nothing to prove. So we may assume that $\omega(G) \geq k + 1$. Applying Algorithm 1 to $H$ with threshold $k + 1$. Suppose that the algorithm stops after $t$ steps and we obtain a sequence of induced subgraphs of $H$, namely, 

$$H = H_0 \supset H_1 \supset \cdots \supset H_t.$$ 

For convenience, we will keep using the notations in Algorithm 1 and Lemma 3.4

**Claim 3.8.** $W_t < 30k^2e_n$.

**Proof of Claim 3.8** Let $\beta = 20k^2e$ and assume that $W_t \geq \beta n$. By Claim 3.7 $\omega_i < 10ken$ for all $1 \leq i \leq t$. So there exists $t' < t$ such that $\beta n - 10ken < W_{t'} < \beta n + 10ken$, and without loss of generality we may assume that $W_t = \beta n$ (since we may replace $W_t$ by $W_{t'}$ and the exact value of $\beta$ is not crucial in the proof).

Let $x' = 2(e - E_t)/(n - W_t)^2$ and it follows from Theorem 3.3 and (26) that 

$$|H| \leq \frac{x'(1 - x')}{6}(n - W_t)^3 + 3(n - W_t)^2 + te$$

$$= -2E_t^2 + \left(4e - (n - W_t)^2\right)E_t + (n - W_t)^2e - 2e^2 + 3(n - W_t)^2 + te. \quad (29)$$

Similar to the proof of Claim 3.7 and by (26), we may substitute $E_t = (n - W_t)n$ into (29) and obtain 

$$|H| \leq -2n^2t^2 + (n(n + W_t)^2 - (n + 3W_t)e)t - (e - W_t)n(2e - n^2 - W_t^2) + 3(n - W_t)^2. \quad (30)$$

Since $t \leq W_t/(k + 1)$ and $-2n^2t^2 + (n(n + W_t)^2 - (n + 3W_t)e)t$ is increasing in $t$ when 

$$t \leq (n(n + W_t)^2 - (n + 3W_t)e) / (4n^2),$$

we may substitute $t = W_t/(k + 1)$ into (30) and obtain 

$$|H| \leq \frac{(k + 1)(-2(k + 1)e^2 + ((k + 1)n^2 + (2k + 1)W_t n + (k - 2)W_t^2)e)}{3(k + 1)^2(n - W_t)}$$

$$- \frac{((k + 1)(n^2 + W_t^2) - 2W_t n)kW_t n}{3(k + 1)^2(n - W_t)} + 3(n - W_t)^2. \quad (31)$$

Since $-2(k + 1)^2e^2 + (k + 1)\left((k + 1)n^2 + (2k + 1)W_t n + (k - 2)W_t^2\right)e$ is decreasing in $e$ when 

$$e \geq \frac{(k + 1)n^2 + (2k + 1)W_t n + (k - 2)W_t^2}{4(k + 1)},$$

we may substitute $e = (1 - \epsilon)(k - 1)n^2/(2k)$ into (31) and obtain 

$$|H| \leq (1 - \epsilon)\frac{k - 1}{6k^2}n^3 - \frac{(k + 1)^2W_t n^3 - k(k^3 + 2k^2 - k + 2)W_t^2 n^2 + 2k^3(k + 1)W_t^3 n}{6k^2(k + 1)^2(n - W_t)}$$

$$+ \epsilon n^3 + 3(n - W_t)^2$$

$$< (1 - \epsilon)\frac{k - 1}{6k^2}n^3 - \frac{(k + 1)^2W_t n^3}{12k^2(k + 1)^2n} + 2\epsilon n^3$$

$$< (1 - \epsilon)\frac{k - 1}{6k^2}n^3 - \frac{2\beta}{12k^2(k + 1)^2n} + 2\epsilon n^3 < (1 - \epsilon)\frac{k - 1}{6k^2}n^3,$$  

a contradiction. Here we used $\beta = 30k^2e$. \[\blacksquare\]
Our next claim gives an upper bound for the number of copies of $K_{k+1}$ in $G$.

**Claim 3.9.** The number of copies of $K_{k+1}$ in $G$ is less than $30k^2 \epsilon n^{k+1}$.

**Proof of Claim 3.9.** Since we are applying Algorithm 1 to $H$ with threshold $k+1$, $G_t$ is $K_{k+1}$-free. So every copy of $K_{k+1}$ in $G$ has at least one vertex in $V(H) \setminus T_t$. By Claim 3.8, $|V(H) \setminus T_t| = W_t < 30k^2 \epsilon n$. Therefore, the number of copies of $K_{k+1}$ in $G$ is less than $30k^2 \epsilon n^{(3k-1)/k} < 30k^2 \epsilon n^{k+1}$.

By Theorem 3.1 and Claim 3.9, we can obtain a $K_{k+1}$-free graph $G'$ from $G$ by removing at most $bn^2$ edges, where $b = b(k, \epsilon) > 0$ is a constant only related to $k, \epsilon$ and it is sufficiently small. In other words, $G$ contains a $K_{k+1}$-free subgraph $G'$ with at least $|G| - bn^2 > (k - 1)n^2/(2k) - (\epsilon + b)n^2$ edges. Therefore, by Theorem 3.2, there exists a partition $V(H) = V_1 \cup \cdots \cup V_k$ such that at least $(k - 1)n^2/(2k) - 2(\epsilon + b)n^2$ edges in $G$ have at most one vertex in each $V_i$. Let $\epsilon_1 = \epsilon + b$.

$$K = \{A \in G : |A \cap V_i| \leq 1 \text{ for all } 1 \leq i \leq k\},$$

and

$$\mathcal{K} = \{A \in H : |A \cap V_i| \leq 1 \text{ for all } 1 \leq i \leq k\}.$$

Note that

$$|K| > \frac{k - 1}{2k} n^2 - 2\epsilon_1 n^2,$$

and by Claim 3.6 and (32),

$$\mathcal{K} \geq |H| - |G \setminus K| n > (1 - \epsilon) \frac{k - 1}{6k^2} n^3 - \left(\frac{k - 1}{2k} + 4\epsilon - \left(\frac{k - 1}{2k} - 2\epsilon_1\right)\right) n^3
\geq \frac{k - 1}{6k^2} n^3 - 7\epsilon_1 n^3,$$

(33)

Our next step is to show that the structure of $\mathcal{K}$ is close to some member of $\mathcal{S}(n, k)$. Recall that for every $\{u, v\} \subset V(K)$ (which equals $V(H)$), the neighborhoods of $uv$ in $\mathcal{K}$ is

$$N_\mathcal{K}(uv) = \{w \in V(K) : \{u, v, w\} \in \mathcal{K}\},$$

(34)

and we will omit the subscript if it is clear from context.

**Claim 3.10.** $|V_i| - n/k < 2\epsilon_1^{1/2} n$ for all $i \in [k]$.

**Proof of Claim 3.10.** Fix $i \in [k]$ and let $\beta = |V_i|$. Then

$$\frac{k - 1}{2k} n^2 - 2\epsilon_1 n^2 < |K| \leq (1 - \beta)\beta n^2 + \frac{k - 2}{2(k - 1)}(1 - \beta)^2 n^2,$$

which implies that $n/k - 2\epsilon_1^{1/2} n < \beta < n/k + 2\epsilon_1^{1/2} n$.

The set of missing edges of $K$ is

$$M_K := \{\{u, v\} \subset V(K) : \exists i, j \in [k], i \neq j, u \in V_i, v \in V_j, \text{ and } uv \not\in K\},$$

and it follows from Turán’s theorem that

$$|M_K| \leq \frac{k - 1}{2k} n^2 - |K| < 2\epsilon_1 n^2.$$

(35)

Let $\alpha(K)$ denote the independent number of $K$, i.e. $\alpha(K)$ is the largest integer $\alpha$ such that there exists an independent set $S \subset V(K)$ in $K$ with $|S| = \alpha$. Our next claim gives an upper bound for $\alpha(K)$.
Claim 3.11. \( \alpha(K) < n/k + 4\epsilon_1^{1/2}kn \).

Proof of Claim 3.11. Suppose that \( \alpha(K) \geq n/k + 3\epsilon_1^{1/2}n \) and let \( S \subset V(H) \) be an independent set in \( K \) with size \( \alpha(K) \). Since there is no edge between \( S \cap V_i \) and \( S \cap V_j \) for all \( \{i, j\} \subset [k] \), it follows from Claim 3.10 that

\[
|M_K| \geq \left( \alpha(K) - \left( \frac{n}{k} + 2\epsilon_1^{1/2}n \right) \right) \left( \frac{n}{k} - 2\epsilon_1^{1/2}n \right) > 2\epsilon_1 n^2,
\]

which contradicts \( \ref{claim:3.10} \).

Let \( uv \in G \). Since \( H \) is cancellative, it is easy to see that \( N_H(uv) \) is an independent set in \( H \), i.e. every edge in \( H \) has at most one vertex in \( N(uv) \). Therefore, \( N_H(uv) \) is an independent set in \( K \), and it follows from Claim 3.11 that \( |N_H(uv)| < n/k + 4\epsilon_1^{1/2}kn \).

Since \( N_K(uv) \subset N_H(uv) \),

\[
|N_K(uv)| < n/k + 4\epsilon_1^{1/2}kn.
\]

Let \( B_E = \{ uv \in K : |N_K(uv)| \leq n/(2k) \} \). Our next claim shows that \( |B_E| \) is small.

Claim 3.12. \( |B_E| < 8\epsilon_1^{1/2}k^2n^2 \).

Proof of Claim 3.12. It follows from \( \sum_{uv \in K} |N_K(uv)| \geq 3|K| \) and Lemma 2.4 that

\[
|B_E| < \frac{n}{k} + 4\epsilon_1^{1/2}kn - \frac{3|K|}{2k},
\]

\[
< 2k \left( \frac{n}{k} + 4\epsilon_1^{1/2}kn - \frac{3k}{2k}n^3 - 7\epsilon_1n^3 \right) \left( \frac{k}{2k}n - 2\epsilon_1n^2 \right) k - 1 n^2
\]

\[
< 8\epsilon_1^{1/2}k^2n^2.
\]

Our next claim shows that for each \( uv \in K \) most of \( N_K(uv) \) is contained in some \( V_i \).

Claim 3.13. Let \( uv \in K \). Then for every \( i \in [k] \) either \( |N_K(uv) \cap V_i| < 4\epsilon_1n^2/|N_K(uv)| \) or \( |N_K(uv) \cap V_i| > |N_K(uv)| - 4\epsilon_1n^2/|N_K(uv)| \). In particular, if \( |N_K(uv)| > n/\hat{c} \) for some constant \( \hat{c} < (2\epsilon_1k)^{-1/2} \), then there exists a unique \( i \in [k] \) such that \( |N_K(uv) \cap V_i| > |N_K(uv)| - 4\epsilon_1\hat{c}n \).

Proof of Claim 3.13. Fix \( uv \in K \) and \( i \in [k] \). Let \( \alpha = |N_K(uv)| \) and \( \beta = |N_K(uv) \cap V_i| \). Since \( N_K(uv) \) is an independent set, there is no edge between \( N_K(uv) \cap V_i \) and \( N_K(uv) \setminus V_i \). So \( |M_K| \geq \beta(\alpha - \beta) \), and it follows from \( \ref{claim:3.11} \) that \( \beta(\alpha - \beta) < 2\epsilon_1n^2 \). Therefore, \( \beta < 4\epsilon_1n^2/\alpha \) or \( \beta > \alpha - 4\epsilon_1n^2/\alpha \).

Now, suppose that \( |N_K(uv)| > n/\hat{c} \) and \( (2\epsilon_1k)^{-1/2} \). Since

\[
k \frac{4\epsilon_1n^2}{|N_K(uv)|} < \frac{4\epsilon_1kn^2}{n/\hat{c}} = 4\epsilon_1\hat{c}kn < \frac{n}{\hat{c}} < |N_K(uv)|,
\]

there exists \( i \in [k] \) such that

\[
|N_K(uv) \cap V_i| > |N_K(uv)| - 4\epsilon_1n^2/|N_K(uv)| > |N_K(uv)| - 4\epsilon_1n^2/n/\hat{c} > |N_K(uv)| - 4\epsilon_1\hat{c}n.
\]

Since \( |N_K(uv)| - 4\epsilon_1\hat{c}n > |N_K(uv)|/2 \), such \( i \) is unique.

Recall that the link of \( v \in V(K) \) in \( K \) is \( L_K(v) = \{ S \subset V(K) : \{v\} \cup S \in H \} \) and \( d_K(v) = |L_K(v)| \). Let \( \Delta(K) = \max_{v \in V(K)} \{ d_K(v) \} \) be the maximum degree of \( K \). Our next claim gives an upper bound for \( \Delta(K) \).
Claim 3.14. $\Delta(K) < \frac{k-1}{2k^2} n^2 + 3\epsilon_1/2 kn^2$.

Proof of Claim 3.14. Fix $v \in V(K)$ and it suffices to show that $d_K(v) < \frac{k-1}{2k^2} n^2 + 3\epsilon_1/2 n^2$. Without loss of generality, we may assume that $v \in V_1$. Let $w \in N_K(v) \subset \bigcup_{i=2}^{k} V_i$, and let $d_v(w)$ denote the degree of $w$ in $L(v)$. By (36), $d_v(w) = |N_K(vw)| < n/k + 4\epsilon_1/2 kn$. Therefore,

$$d_K(v) = \frac{1}{2} \sum_{w \in \bigcup_{i=2}^{k} V_i} d(w) < \frac{k-1}{2} \left( \frac{n}{k} + 2\epsilon_1/2 n \right) \left( \frac{n}{k} + 4\epsilon_1/2 kn \right)$$

$$< \frac{k-1}{2k^2} n^2 + 3\epsilon_1/2 kn^2.$$ 

\[ \Box \]

Let $B_V = \left\{ v \in V(H) : d_K(v) < \frac{k-1}{2k^2} n^2 - 20\epsilon_1/2 k^2 n^2 \right\}$. Our next claim gives an upper bound for $|B_V|$.

Claim 3.15. $|B_V| < \frac{n}{6k}$.

Proof of Claim 3.15. It follows from $\sum_{v \in V(K)} d_K(v) = 3|K|$ and Lemma 2.4 that

$$|B_V| \leq \frac{k-1}{2k^2} n^2 + 3\epsilon_1/2 kn^2 - \frac{3|K|}{n} \frac{n}{k} < \frac{n}{6k}.$$ 

\[ \Box \]

It is easy to see that the link of every vertex in a $k$-vertex Steiner triple system is a matching with $(k-1)/2$ edges, i.e. a graph consisting of $(k-1)/2$ pairwise disjoint edges. Also, notice that a blow up of an edge is a bipartite graph. Our next claim shows that for every $v \in V(H) \setminus B_V$, $L_K(v)$ is almost (i.e. after removing a small number of edges) consisting of $(k-1)/2$ pairwise vertex disjoint bipartite graphs.

Claim 3.16. For every $v \in V(H) \setminus B_V$ there exists $L'_K(v) \subset L_K(v)$ with $|L'_K(v)| > d_K(v) - 255\epsilon_1/2 k^4 n^2$ such that $L'_K(v)$ is consisting of $(k-1)/2$ pairwise vertex disjoint bipartite graphs.

Proof of Claim 3.16. Fix $v \in V(H) \setminus B_V$ and without loss of generality, we may assume that $v \in V_1$. Note that $L_K(v)$ is a graph on $N_K(v) \subset \bigcup_{i=2}^{k} V_i$. For every $u \in N_K(v)$ let $d_v(u)$ denote the degree of $u$ in $L_K(v)$. Let

$$B_v = \left\{ u \in \bigcup_{i=2}^{k} V_i : d_v(u) \leq \frac{n}{k} - 250\epsilon_1/2 k^3 n \right\}.$$ 

Then, it follows from Lemma 2.4 that

$$|B_v| \leq \frac{n}{k} - 4\epsilon_1/2 kn - \frac{2d_K(v)}{|\bigcup_{i=2}^{k} V_i|} \frac{n}{k} - 4\epsilon_1/2 kn - \left( \frac{n}{k} - 250\epsilon_1/2 k^3 n \right)$$

$$\leq \frac{n}{k} - 4\epsilon_1/2 kn - \left( \frac{n}{k} - 250\epsilon_1/2 k^3 n \right) < \frac{n}{6k}.$$ 

22
Therefore, for every $i \in [k] \setminus \{1\}$,

$$|V_i \setminus B_v| \geq \left( \frac{n}{k} - 2\epsilon_1^{1/2} n \right) - \frac{n}{6k} > \frac{2n}{3k}.$$  \hfill (37)

Fix $i \in [k] \setminus \{1\}$ and let $u \in V_i \setminus B_v$. By Claim 3.10 there exists $i' \in [k] \setminus \{1, i\}$ such that

$$|N_K(uv) \cap V_{i'}| > |N_K(uv)| - \frac{4\epsilon_1 n^2}{|N_K(uv)|} = d_v(u) - \frac{4\epsilon_1 n^2}{d_v(u)} > \left( \frac{n}{k} - 250\epsilon_1^{1/2} k^3 n \right) - \frac{4\epsilon_1 n^2}{n/k - 250\epsilon_1^{1/2} k^3 n} > \frac{n}{k} - 251\epsilon_1^{1/2} k^3 n$$

Therefore, by (37) and the Pigeonhole principle, there exists $U_i \subset V_i \setminus B_v$ with $|U_i| > 2n/(3k(2k - 2))$ such that all vertices in $U_i$ have at least $n/k - 251\epsilon_1^{1/2} k^3 n$ neighbors in $V_{i'}$ for some $i' \in [k] \setminus \{1, i\}$. Define the bipartite graph $G_{i,i'}$ as

$$G_{i,i'} = \{ uv \in L_K(v) : u \in U_i, w \in V_{i'} \},$$

and notice that

$$|G_{i,i'}| > |U_i| \left( \frac{n}{k} - 251\epsilon_1^{1/2} k^3 n \right). \hfill (38)$$

Let $V_{i'} = \left\{ u \in V_{i'} : d_{G_{i,i'}}(u) > |U_i|/2 > n/(2k(k - 2)) \right\}$. Then

$$|V_{i'}|/|U_i| + \left( |V_{i'}| - |V_{i'}'| \right) \frac{|U_i|}{2} > |G_{i,i'}| > |U_i| \left( \frac{n}{k} - 251\epsilon_1^{1/2} k^3 n \right),$$

which implies that

$$|V_{i'}| > 2 \left( \frac{n}{k} - 251\epsilon_1^{1/2} k^3 n - \frac{|V_{i'}|}{2} \right) \quad \text{Claim 3.10} > 2 \left( \frac{n}{k} - 251\epsilon_1^{1/2} k^3 n - \frac{1}{2} \left( \frac{n}{k} + 2\epsilon_1^{1/2} n \right) \right) > \frac{n}{k} - 504\epsilon_1^{1/2} k^3 n$$

For every $u \in V_{i'}$, since $|N_K(uv)| \geq d_{G_{i,i'}}(u) > n/(2k(k - 2))$, it follows that

$$|N_K(uv) \cap V_i| \geq d_{G_{i,i'}}(u) > \frac{|U_i|}{2} > \frac{n}{2k(k - 2)} > \frac{4\epsilon_1 n^2}{n/(2k(k - 2))} > \frac{4\epsilon_1 n^2}{|N_K(uv)|}.$$

Therefore, by Claim 3.13

$$|N_K(uv) \cap V_i| > |N_K(uv)| - \frac{4\epsilon_1 n^2}{|N_K(uv)|} > |N_K(uv)| - \frac{4\epsilon_1 n^2}{n/(2k(k - 2))} > |N_K(uv)| - 8\epsilon_1 k^2 n.$$

So, by the Pigeonhole principle, there exists $U_{i'} \subset V_{i'} \setminus B_v$ with $|U_{i'}| > 2n/(2k(k - 2))$ such that every $w \in U_{i'}$ satisfies

$$|N_K(uv) \cap V_i| > \left( \frac{n}{k} - 250\epsilon_1^{1/2} k^3 n \right) - 8\epsilon_1 k^2 n > \frac{n}{k} - 251\epsilon_1^{1/2} k^3 n.$$

Let $G_{i,i} = \{ wu \in L_K(v) : w \in U_{i'}, u \in V_i \}$, and a similar argument as above shows that there exists $V_i' \subset V_i$ with $|V_i'| > n/k - 504\epsilon_1^{1/2} k^3 n$ such that for all $u \in V_i'$

$$|N_K(uv) \cap V_i'| > |N_K(uv)| - 8\epsilon_1 k^2 n.$$
Therefore, we can remove at most
\[8c_1k^2n \left( |V'_i| + |V''_i| \right) + |V_i \setminus V'_i|n + |V''_i \setminus V'_i|n\]
Claim 3.10
\[< 16c_1k^2n \left( \frac{n}{k} - 251\epsilon_1^{1/2}k^3n \right) + 2n \left( \frac{n}{k} + 2\epsilon_1^{1/2}n \right) - \left( \frac{n}{k} - 251\epsilon_1^{1/2}k^3n \right)\]
\[< 505\epsilon_1^{1/2}k^3n^2\]
edges from \(L_K(v)\) such that for every \(uw \in L_K(v)\) either \(\{u, w\} \cap (V_i \cup V_{i'}) = \emptyset\) or \(u \in V_i\) and \(w \in V_{i'}\).

Repeating the same argument as above, we know that members in \(\{2, \ldots, k\}\) form \((k - 1)/2\) disjoint pairs \(\{i, i'\}, \ldots, \{j, j'\}\) such that one can remove at most
\[\frac{k - 1}{2} \times 505\epsilon_1^{1/2}k^3n^2 < 255\epsilon_1^{1/2}k^4n^2\]
edges from \(L_K(v)\) such that the resulting graph \(L'_K(v)\) is consist of \((k - 1)/2\) pairwise vertex disjoint bipartite graphs with sets of parts \(\{V_i, V_{i'}\}, \ldots, \{V_j, V_{j'}\}\).

Now we are ready to finish the proof of Theorem 3.5. For every \(v \in V(H)\) define a function \(\varphi_v : [k] \to [k]\) as follows:

(a). If \(v \in V_{i_0}\) for some \(i_0 \in [k]\), then \(\varphi_v(i_0) = i_0\).

(b). For every \(i \in [k] \setminus \{i_0\}\) let \(i'\) be given by the proof of Claim 3.16 and let \(\varphi_v(i) = i'\).

Note that if \(\varphi_v(i) = i'\), then \(\varphi_v(i') = i\).

Fix \(i \in [k]\) and without loss of generality, we may assume that \(i = 1\). Since the number of different functions \(\varphi : [k] \to [k]\) that satisfy (a) and (b) is at most \(k!\), by Pigeonhole principle and Claims 3.10 and 3.15, there exists \(\tilde{V}_1 \subset V_1 \setminus B\) with \(|\tilde{V}_1| > n/(2kk!\) such that \(\varphi_u \equiv \varphi_v\) for all \(u, v \in \tilde{V}_1\) and \(d_K(w) > \frac{k - 1}{2k^2}n^2 - 20\epsilon_1^{1/2}k^2n^2\) for all \(w \in \tilde{V}_1\).

Fix \(v \in \tilde{V}_1, i \in [k] \setminus \{1\}\), and suppose that \(\varphi_v(i) = i'\) for some \(i' \in [k] \setminus \{1, i\}\). Define an auxiliary bipartite graph \(M\) with two parts \(C = \tilde{V}_1\) and \(D = V_i \setminus V_{\varphi_v(i)}\) as follows: for every \(a \in C\) and \(b \in D\), \(ab\) is an edge in \(M\) iff \(b \in L_K(a)\). By Claim 3.16, there exists \(L'_K(v) \subset L_K(v)\) such that \(L'_K(v)\) is consisting of \((k - 1)/2\) pairwise vertex disjoint bipartite graphs and
\[|L'_K(v)| > d_K(v) - 255\epsilon_1^{1/2}k^4n^2 > \frac{k - 1}{2k^2}n^2 - 20\epsilon_1^{1/2}k^2n^2 - 255\epsilon_1^{1/2}k^4n^2\]
\[> \frac{k - 1}{2k^2}n^2 - 275\epsilon_1^{1/2}k^4n^2\]
Therefore, by Claim 3.10, the induced bipartite subgraph \(L'_K(v)[V_i, V_{\varphi_v(i)}]\) of \(L_K(v)\) satisfies
\[|L_K(v)V_i, V_{\varphi_v}()| > |L'_K(v)| - \left( \frac{k - 1}{2} - 1 \right) \left( \frac{n}{k} + 2\epsilon_1^{1/2}n \right)^2\]
\[> \frac{k - 1}{2k^2}n^2 - 275\epsilon_1^{1/2}k^4n^2 - \left( \frac{n}{k} + 2\epsilon_1^{1/2}n \right)^2\]
\[> \frac{n^2}{k^2} - 276\epsilon_1^{1/2}k^4n^2\]
\[> \frac{n^2}{k^2} - 276\epsilon_1^{1/2}k^4n^2\]
So by the definition of \(M\),
\[|M| \geq |\tilde{V}_1| \left( \frac{n^2}{k^2} - 276\epsilon_1^{1/2}k^4n^2 \right),\]
and by Claim 3.10 and Lemma 2.4, at least
\[|D| - \frac{|\hat{V}_1| - \frac{|M|}{|\hat{V}_1|}}{|\hat{V}_1|} > \left(\frac{n^2}{k} - 2\epsilon_1^{1/2}n\right)^2 - \frac{2}{|\hat{V}_1|} \left(\hat{V}_1 - \frac{|\hat{V}_1| \left(n^2/k^2 - 276\epsilon_1^{1/2}k^4n^2\right)}{n/k + 2\epsilon_1^{1/2}n}\right)\]
\[> \frac{n^2}{k^2} - 560\epsilon_1^{1/2}k^4n^2\]
vertices \(u \in D\) satisfies \(d_M(u) > |\hat{V}_1|/2\). In other words, there are at least \(\frac{n^2}{k^2} - 560\epsilon_1^{1/2}k^4n^2\) pairs \((w_1, w_2) \in V_i \times V_{\varphi_v(i)}\) such that \(|N_K(w_1w_2) \cap V_i| > |\hat{V}_1|/2 > n/(4kk!)\). Therefore, by Claim 3.13,
\[|N_K(w_1w_2) \cap V_i| > |N_K(w_1w_2)| - \frac{4\epsilon_1n^2}{|N_K(w_1w_2)|} > |N_K(w_1w_2)| - \frac{4\epsilon_1n^2}{n/(4kk!)} > |N_K(w_1w_2)| - 16\epsilon_1kk!n\]
Therefore, we can remove at most
\[16\epsilon_1kk!n \left(\frac{n^2}{k^2} - 560\epsilon_1^{1/2}k^4n^2\right) + \left(\frac{n^2}{k} - 2\epsilon_1^{1/2}\right)^2 - \left(\frac{n^2}{k^2} - 560\epsilon_1^{1/2}k^4n^2\right)\] \(n < 561\epsilon_1^{1/2}k^4n^3\)
edges from \(K\) such that every remaining edge \(\{w_1, w_2, w_3\}\) satisfies that if \(w_1 \in V_i\) and \(w_2 \in V_{\varphi_v(i)}\), then \(w_3 \in V_i\). Repeating the same argument as above to all pairs \((j, \varphi_v(j))\) for \(j \in [k] \setminus \{1, i, \varphi_v(i)\}\), after removing at most
\[
\frac{k - 1}{2} \times 561\epsilon_1^{1/2}k^4n^3 < 290\epsilon_1^{1/2}k^6n^3\]edges from \(K\), each remaining edge \(\{w'_1, w'_2, w'_3\}\) satisfies that if \(w'_1 \in V_j\) and \(w'_2 \in V_{\varphi_v(j)}\) for some \(j \in [k] \setminus \{1\}\), then \(w'_3 \in V_i\).
Repeating the argument above to \(V_i\) for \(2 \leq \ell \leq k\), it is easy to see that after removing at most \(290\epsilon_1^{1/2}k^6n^3\) edges from \(K\), the remaining 3-graph \(K'\) is a subgraph of a member of \(S(n, k)\).

### 3.2 Exact result

In this section we will prove Theorem 1.16. The following lemma, which will be used in the proof of Theorem 1.16, is an extension of Lemma 3.17.

Given a graph \(G\) and \(S, T \subset V(G)\) and \(S \cap T = \emptyset\), we use \(G[S, T]\) to denote the induced bipartite subgraph of \(G\) with the set of parts \(\{S, T\}\).

**Lemma 3.17.** Let \(H\) be a cancellative 3-graph and \(S \cup T\) be a partition of \(V(H)\). Let \(e(S), e(S, T), e(T)\) denote the number of edges in \((\partial H)[S], (\partial H)[S, T], (\partial H)[T]\), respectively. Suppose that \((\partial H)[S]\) is a complete graph. Then \(|H| \leq |H[T]| + e(T) + e(S, T)/2 + e(S)/3\).

**Proof.** It suffices to prove that the \(|H \setminus H[T]| \leq e(T) + e(S, T)/2 + e(S)/3\). For \(1 \leq i \leq 3\), we say \(E \in H \setminus H_T\) is of type-\(i\) if \(|E \cap S| = i\). Since a type-1 edge contains one edge in \((\partial H)[T]\), it follows from Lemma 3.17 that the number of type-1 edges is at most \(e(T)\). Since \((\partial H)[S]\) is complete, every edge in \((\partial H)[S, T]\) is covered by at most one type-2 edge. On the other hand, every type-2 edge contains exactly 2 edges in \((\partial H)[S, T]\), so the number of type-2 edges is at most \(e(S, T)/2\). Similarly, since \((\partial H)[S]\) is complete, every edge in \((\partial H)[S]\) is covered by at most one type-3 edge. Therefore, the number of type-3 edges is at most \(e(S)/3\). Therefore, \(|H \setminus H[T]| \leq e(T) + e(S, T)/2 + e(S)/3\).
The idea of the proof of Theorem 1.16 is first to show that $\partial H$ is $K_{k+1}$-free. Then, by Turán’s theorem, $\partial H \cong T_2(n, k)$. Finally, we show that $H$ is a copy of some member in $S(n, k)$ using the structure of $\partial H$, and this part is basically the same as the corresponding part in the proof of Theorem 3.5. In order to keep the calculations simple, let us assume that $n$ is a multiple of $k$.

**Proof of Theorem 1.16.** Let $H$ be a cancellative 3-graph on $n$ vertices with $|\partial H| = t_2(n, k) = (k - 1)n^2/(2k)$ and $|H| \geq s(n, k) = (k - 1)n^3/(6k^2)$. It suffices to show that $H \in S(n, k)$.

Let $G = \partial H$. Applying Algorithm 1 to $H$ with the threshold $k + 1$. Suppose that the algorithm stops after $t$ steps and we obtain a sequence of induced subgraphs of $H$, namely, $H = H_0 \supset H_1 \supset \cdots \supset H_t$.

We will keep using the notations in Algorithm 1 and Lemma 2.4.

Our first goal is to show that $G$ is $K_{k+1}$-free. If $\omega(G) = k$, then we are done. So we may assume that $\omega(G) \geq k + 1$. Notice that $\omega_1 \geq \cdots \geq \omega_t \geq k$ and let $t'$ be the largest integer such that $\omega_{t'} \geq k + 1$. Let $e = |G|$. For every $1 \leq i \leq t$ let $e_i$ be the number of edges in $G_{i-1}$ that have at least one vertex in $S_i$, $E_i = \sum_{j=1}^i e_j$, and $W_i = \sum_{j=1}^i \omega_j$.

**Claim 3.18.** $W_{t'} < 100k(k + 1)$, and hence $t' < 100k$.

**Proof of Claim 3.18.** The proof is very similar to the proof of Claim 3.8 and in order to keep our proof short, we will omit some details in the calculations.

Let $x' = (e - E_{t'})/(n - W_{t'})^2$. Similar to the proof of Claim 3.8

$$|H| \leq \frac{x'(1 - x')(n - W_{t'})^3 + 3n^2 + t'e}{6} = \frac{-2E_{t'}^2 + (4e - (n - W_{t'})^2) E_{t'} + (n - W_{t'})^2 e - 2e^2}{3(n - W_{t'})} + 3n^2 + t'e$$

$$= \frac{-2(W_{t'}n - t'n)^2 + (4e - (n - W_{t'})^2) (W_{t'}n - t'n) + (n - W_{t'})^2 e - 2e^2}{3(n - W_{t'})} + 3n^2 + t'e$$

$$\leq \frac{-2n^2(t')^2 + (n(n + W_{t'})^2 - (n + 3W_{t'})e) t' - (e - W_{t'}n)(2e - n^2 - W_{t'}^2)}{3(n - W_{t'})} + 3n^2.$$ (39)

Since $\omega_i \geq k + 1$ for all $1 \leq i \leq t'$, $t' \leq W_{t'}/(k + 1)$. Therefore, we may substitute $t' = W_{t'}/(k + 1)$ into (39) and obtain

$$|H| \leq \frac{(k + 1)(-2(1 + k)e + ((k + 1)n^2 + (2k + 1)W_{t'}n + (k - 2)W_{t'}^2)e)}{3(k + 1)^2(n - W_{t'})}$$

$$- \frac{(k + 1)(n^2 + W_{t'}^2) - 2W_{t'}n}{3(k + 1)^2(n - W_{t'})} kW_{t'}n + 3n^2.$$ (40)

Substituting $e = (k - 1)n^2/(2k)$ into (40) we obtain

$$|H| \leq \frac{k - 1}{6k^2} n^3 - \frac{(k + 1)^2 n^2 - k(k^3 + 2k^2 - k + 2)W_{t'}n + 2k^3(k + 1)W_{t'}^2}{6k^2(k + 1)^2(n - W_{t'})} W_{t'}n + 3n^2.$$ 

If $W_{t'} \geq 100k(k + 1)$, then the inequality above implies that $|H| < (k - 1)n^3/(6k^2)$, a contradiction. Therefore, $W_{t'} < 100(k + 1)$ and $t' \leq W_{t'}/(k + 1) < 100k$.

Our next claim gives an upper bound for $|T_i|$.

**Claim 3.19.** $|T_i| < 20k^2 n^{1/2}$.
Proof of Claim 3.19. First, note that $G_t$ is $K_k$-free, so by Turán’s theorem, $|G_t| \leq (k - 2)|T_t|^2/(2(k - 1))$. For every $1 \leq i \leq t$, since every vertex in $T_i$ is adjacent to at most $\omega_i - 1$ vertices in $S_i$, $|G_{i-1}| \leq |G_i| + (\omega_i - 1)|T_i| + \binom{\omega_i}{2}$. Therefore,

$$|G| \leq \left( \sum_{i=1}^{t'} \omega_i \right) + \sum_{j=t'+1}^{t} \left( (\omega_j - 1) \left( n - \left( \sum_{i=1}^{j} \omega_i \right) \right) + \binom{\omega_j}{2} \right) + |G_t|$$

\[ \leq W_t n + \sum_{i=0}^{t-t'-1} (k - 1)(n - W_t - ik) + \frac{k - 2}{2(k - 1)}|T_i|^2 \]

\[ \leq W_t n + (t - t')(k - 1)(n - W_t) - \frac{k(k - 1)(t - t')(t - t' - 1)}{2} + \frac{k - 2}{2(k - 1)}|T_i|^2 \]

Since $t - t' = (n - W_t - |T_i|)/k$, the inequality above and $|G| = (k - 1)n^2/(2k)$ imply that

$$\frac{k - 1}{2k}n^2 \leq W_t n + (k - 1)(n - W_t) \frac{n - W_t - |T_i|}{k} - \frac{k(k - 1)}{2} \left( \frac{n - W_t - |T_i|}{k} \right)^2$$

$$\leq \frac{k - 1}{2k}n^2 - \frac{1}{2k(k - 1)}|T_i|^2 + \frac{2W_t}{k}n + k^2 n$$

Claim 3.18 $\frac{k - 1}{2k}n^2 - \frac{|T_i|^2}{2k(k - 1)} + 200k^2 n$.

which implies that $|T_i| < 20k^2 n^{1/2}$. ☐

For every $1 \leq i \leq t$ let $x_i = 2|G_i|/|T_i|^2$. Notice that the upper bound for $|\mathcal{H}|$ in Theorem 1.6 has an error term $3n^3$. Our next claim improves this error term to $O(n)$.

Claim 3.20. For every $t' \leq i \leq t$,

$$|\mathcal{H}_i| \leq \frac{x_i(1 - x_i)}{6}|T_i|^3 + k|T_i| + 1200k^4 n.$$

Proof of Claim 3.20. We proceed by backward induction on $i$. When $i = t$, by Theorem 1.5

$$|\mathcal{H}_t| \leq \frac{x_t(1 - x_t)}{6}|T_t|^3 + 3|T_t|^2 = \frac{x_t(1 - x_t)}{6}|T_t|^3 + 1200k^4 n.$$

Now assume that the claim is true for some $i + 1$ with $t' + 1 \leq i + 1 \leq t$ and we want to show that it is also true for $i$. By Lemma 3.17 and the induction hypothesis,

$$|\mathcal{H}_i| \leq |\mathcal{H}_{i+1}| + |G_{i+1}| + \frac{e_{i+1}}{2}$$

$$\leq \frac{x_{i+1}(1 - x_{i+1})}{6}|T_{i+1}|^3 + k|T_{i+1}| + 1200k^4 n + |G_{i+1}| + \frac{e_{i+1}}{2}.$$
We may substitute $e_{i+1} = (k - 1)(|T_i| - k) + \binom{k}{2}$ into the inequality above and obtain
\[
\Delta \geq \frac{8k|G_i|^2 - 4(3k|T_i| - |T_i| - k^2 - k) |T_i||G_i|}{12|T_i|(|T_i| - k)} + \frac{(2|T_i|^2 - |T_i| - k)(2|T_i| - k)(k - 1)}{12(|T_i| - k)} + k^2.
\]
Note that $8k|G_i|^2 - 4(3k|T_i| - |T_i| - k^2 - k) |T_i||G_i|$ is decreasing in $|G_i|$ when
\[
|G_i| \leq \frac{3k - 1}{4k}|T_i|^2 - \frac{k + 1}{4}|T_i|.
\]
On the other hand, since $G_i$ is $K_{k+1}$-free, by Turán’s theorem, we may substitute $|G_i| = (k - 1)|T_i|^2/(2k)$ into the inequality above and obtain $\Delta \geq (11k + 1)k/12 > 0$. Therefore,
\[
|\mathcal{H}_i| \leq \frac{x_i(1 - x_i)}{6}|T_i|^3 + k|T_i| + 1200k^4n.
\]

**Claim 3.21.** $G$ is $K_{k+1}$-free.

**Proof of Claim 3.21.** Recall that $t'$ is the largest integer such that $\omega_{t'} \geq k + 1$ and $W_{t'} = \sum_{i=1}^{t'} \omega_i$. So it suffices to show that $t' = 0$, i.e. $W_{t'} = \sum_{i=1}^{t'} \omega_i = 0$. Suppose that this is not true, i.e. $W_{t'} > 0$. By Lemma 3.4 and Claim 3.20,
\[
|\mathcal{H}| \leq |\mathcal{H}_{t'}| + t'e < \frac{x_{t'}(1 - x_{t'})}{6}|T_{t'}|^3 + 1201k^4n + t'e.
\]
Let
\[
\Delta = \frac{k - 1}{6k^2}n^3 - \left(\frac{x_{t'}(1 - x_{t'})}{6}|T_{t'}|^3 + t'e + 1201k^4n\right),
\]
and by assumption we should have $\Delta \leq 0$. Substituting $x_{t'} = 2(e - E_{t'})/(n - W_{t'})^2$ and $|T_{t'}| = n - W_{t'}$ into (41) we obtain
\[
\Delta \geq \frac{k - 1}{6k^2}n^3 - \left(-\frac{2E_{t'}^2 + (4e - (n - W_{t'}))^2}{3(n - W_{t'})} + t'e + 1201k^4n\right).
\]
Similar to the proof of Claim 3.18, we may substitute $E_{t'} = (W_{t'} - t')n$, $t' = W_{t'}/(k + 1)$, and $e = (k - 1)n^2/(2k)$ into the inequality above and obtain
\[
\Delta \geq \frac{((k + 1)^2n^2 - (k(k^2 + 2k^2 - k + 2)W_{t'}n + 2k^3(k + 1)W_{t'}^2)W_{t'}n}{6k^2(k + 1)^2(n - W_{t'})} - 1201k^4n,
\]
which is greater than 0 when $n$ is sufficiently large, a contradiction.

Since $G$ is $K_{k+1}$-free and $|G| = t_2(n, k)$, by Turán’s theorem, $G \cong T_2(n, k)$. The following claim completes the proof of Theorem 1.3.

**Claim 3.22.** $\mathcal{H} \in \mathcal{S}(n, k)$.

The proof of Claim 3.22 is basically the same as the corresponding part (starting from Claim 3.12) in the proof of Theorem 3.5. One may just replace $\epsilon_1$ by 0 in the proofs and it is easy to obtain the conclusion that $\mathcal{H} \in \mathcal{S}(n, k)$.

## 4 Acknowledgement

We are very grateful to Dhruv Mubayi for inspiring discussions on this project, and many thanks to Sayan Mukherjee for carefully reading the manuscript and for providing many helpful suggestions.
References

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

[2] F. N. Cole. The triad systems of thirteen letters. Trans. Amer. Math. Soc., 14(1):1–5, 1913.

[3] G. P. Egorychev. The solution of van der Waerden’s problem for permanents. Adv. in Math., 42(3):299–305, 1981.

[4] D. I. Falikman. Proof of the van der Waerden conjecture on the permanent of a doubly stochastic matrix. Mat. Zametki, 29(6):931–938, 957, 1981.

[5] J. Fox. A new proof of the graph removal lemma. Ann. of Math., pages 561–579, 2011.

[6] Z. Füredi. A proof of the stability of extremal graphs, Simonovits’ stability from Szemerédi’s regularity. J. of Combin. theory. Ser. B, 115:66–71, 2015.

[7] M. Hall, Jr. and J. D. Swift. Determination of Steiner triple systems of order 15. Math. Tables Aids Comput., 9:146–152, 1955.

[8] P. Keevash. Shadows and intersections: stability and new proofs. Adv. Math., 218(5):1685–1703, 2008.

[9] P. Keevash. Hypergraph Turán problems. Surveys in combinatorics, 392:83–140, 2011.

[10] P. Keevash. Counting Steiner triple systems. In European Congress of Mathematics, pages 459–481. Eur. Math. Soc., Zürich, 2018.

[11] P. Keevash and D. Mubayi. Stability theorems for cancellative hypergraphs. J. of Combin. Theory. Ser. B, 92(1):163–175, 2004.

[12] X. Liu and D. Mubayi. The feasible region of hypergraphs. arXiv preprint arXiv:1911.02090, 2019.

[13] X. Liu and D. Mubayi. A hypergraph Turán problem with no stability. arXiv preprint arXiv:1911.07969, 2019.

[14] L. Lovász. Combinatorial Problems and Exercises. North-Holland Publishing Co., Amsterdam, second edition, 1993.

[15] O. Pikhurko. An exact Turán result for the generalized triangle. Combinatorica, 28(2):187–208, 2008.

[16] J. B. Shearer. A new construction for cancellative families of sets. Electron. J. Combin., 3(1):3, 1996.

[17] A. F. Sidorenko. On the maximal number of edges in a homogeneous hypergraph that does not contain prohibited subgraphs. Mat. Zametki, 41(3):433–455, 459, 1987.

[18] M. Simonovits. A method for solving extremal problems in graph theory, stability problems. Theory of Graphs, Proc. Colloq. Tihany, Hungary 1966, 279-319 (1968), 1968.
[19] P. Turán. On an extremal problem in graph theory. *Mat. Fiz. Lapok*, 48:436–452, 1941.

[20] R. Wilson. The early history of block designs. *Rend. Semin. Mat. Messina, Ser. II*, 9:267–276, 2004.

[21] R. M. Wilson. Nonisomorphic Steiner triple systems. *Math. Z.*, 135:303–313, 1973/74.