Stability theorems for some Kruskal-Katona type results

Xizhi Liu * Sayan Mukherjee †

April 6, 2021

Abstract

The classical Kruskal-Katona theorem gives a tight upper bound for the size of an r-uniform hypergraph $\mathcal{H}$ as a function of the size of its shadow. Its stability version was obtained by Keevash who proved that if the size of $\mathcal{H}$ is close to the maximum, then $\mathcal{H}$ is structurally close to a complete r-uniform hypergraph. We prove similar stability results for two classes of hypergraphs whose extremal properties have been investigated by many researchers: the cancellative hypergraphs and hypergraphs without expansion of cliques.

Keywords. hypergraphs, shadows, the Kruskal-Katona theorem, Turán problems, stability.

1 Introduction

An r-uniform hypergraph (henceforth r-graph) $\mathcal{H}$ is a collection of r-subsets of a ground set $V(\mathcal{H})$, which is called the vertex set of $\mathcal{H}$. The r-sets that are contained in $\mathcal{H}$ are called edges of $\mathcal{H}$, and we identify $\mathcal{H}$ with its edge set. We use K_n^r to denote the complete r-graph on n vertices. The *shadow* $\partial\mathcal{H}$ of an r-graph $\mathcal{H}$ is an $(r-1)$-graph defined as

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

The classical Kruskal-Katona theorem [14, 9] gives a tight upper bound for $|\mathcal{H}|$ as a function of $|\partial\mathcal{H}|$. Here we state a technically simper version due to Lovász.

Theorem 1.1 (Lovász [20, Ex. 13.31(b)]). *Let $r \geq 2$ be an integer and $\mathcal{H}$ be an r-graph with $|\partial\mathcal{H}| = \binom{x}{r-1}$ for some real number $x \geq r$. Then $|\mathcal{H}| \leq \binom{x}{r}$. Moreover, equality holds if x is an integer and $\mathcal{H}$ is a union of K_x^r and a set of isolated vertices.*

Keevash [10] gave a nice short proof to Theorem 1.1 without using the compression technique, and moreover, his proof was extended to obtain the following stability result.

Theorem 1.2 (Keevash [10]). *Let $r \geq 2$ be an integer. For every $\delta > 0$ there exists $\epsilon > 0$ such that if $\mathcal{H}$ is an r-graph with $|\partial\mathcal{H}| = \binom{x}{r-1}$ and $|\mathcal{H}| \geq (1-\epsilon)\binom{x}{r}$ for some real number $x \geq r$, then there exists a set $V' \subset V(\mathcal{H})$ of size at most $\lceil x \rceil$ such that all but at most $\delta\binom{x}{r}$ edges of $\mathcal{H}$ are contained in V'.*

*Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL, 60607 USA. Email: xliu246@uic.edu. Research partially supported by NSF awards DMS-1763317 and DMS-1952767.

†Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL, 60607 USA. Email: smukhe2@uic.edu.

In this work, we consider extensions of Theorem 1.2 to the class of hypergraphs that omit certain forbidden substructures.

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 (not necessarily induced) subgraph. The *Turán number* $\mathrm{ex}(n, \mathcal{F})$ 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 $\pi(\mathcal{F}) = \lim_{n\to\infty} \mathrm{ex}(n, \mathcal{F})/\binom{n}{r}$. It is one of the central problems in extremal combinatorics to determine $\mathrm{ex}(n, \mathcal{F})$ for various families $\mathcal{F}$.

Much is known about $\mathrm{ex}(n, \mathcal{F})$ when $r = 2$ and one the most famous results in this regard is Turán's theorem [29], which states that for $\ell \geq 2$ the Turán number $\mathrm{ex}(n, K_{\ell+1})$ is uniquely achieved by $T(n, \ell)$ which is the ℓ-partite graph on n vertices with the maximum number of edges. However, for $r \geq 3$ determining $\mathrm{ex}(n, \mathcal{F})$, even $\pi(\mathcal{F})$, is notoriously hard in general. Compared to the case $r = 2$, very little is known about $\mathrm{ex}(n, \mathcal{F})$ for $r \geq 3$, and we refer the reader to [11] for results before 2011.

To have a better understanding of the extremal properties of $\mathcal{F}$-free hypergraphs, in [17], the following question which combines the Kruskal-Katona theorem and the hypergraph Turán problem was studied systemically.

If $\mathcal{H}$ is $\mathcal{F}$-free, what are the possible values of $|\mathcal{H}|$ for fixed $|\partial\mathcal{H}|$?

Tight upper bound for $|\mathcal{H}|$ was obtained in [17] for two specific families that have been investigated by many researchers: cancellative hypergraphs and hypergraphs without expansions of cliques.

1.1 Cancellative hypergraphs

For every integer $r \geq 2$ let $\mathcal{T}_r$ be the family of r-graphs with at most $2r - 1$ vertices and three edges A, B, C such that $A \triangle B \subset C$. Note that when $r = 2$ the family $\mathcal{T}_2$ consists of only one graph K_3. An r-graph $\mathcal{H}$ is *cancellative* if it has the property that $A \cup B = A \cup C$ implies $B = C$, where A, B, C are edges in $\mathcal{H}$. It is easy to see that an r-graph is cancellative if and only if it is $\mathcal{T}_r$-free, and a cancellative graph is simply a triangle-free graph.

Let $\ell \geq r \geq 2$ be integers and let $V_1 \cup \cdots \cup V_\ell$ be a partition of $[n]$ 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 $t_r(n, \ell) = |T_r(n, \ell)| \sim \binom{\ell}{r}(n/\ell)^r$.

In the 1960's, Katona raised the problem of determining the maximum size of a cancellative 3-graph on n vertices and conjectured that the maximum size is achieved by $T_3(n, 3)$. Bollobás [4] proved Katona's conjecture and he conjectured that a similar result holds for all $r \geq 4$. Sidorenko [27] proved it for $r = 4$, but Shearer [26] gave a construction showing that Bollobás' conjecture is false for all $r \geq 10$. The number $\pi(\mathcal{T}_r)$ is still unknown for all $r \geq 5$.

The following Kruskal-Katona type result for cancellative r-graphs was proved in [17] despite $\pi(\mathcal{T}_r)$ is only known for $r \in \{3, 4\}$.

Theorem 1.3 ([17]). *Let $r \geq 2$ be an integer and $\mathcal{H}$ be a cancellative r-graph. Suppose that $|\partial\mathcal{H}| = x^{r-1}/r^{r-2}$ for some real number $x \geq r$. Then $|\mathcal{H}| \leq (x/r)^r$. In other words, $|\mathcal{H}| \leq (|\partial\mathcal{H}|/r)^{r/(r-1)}$.*

For integers $n \geq m \geq \ell \geq r \geq 2$ let $T_r(n, m, \ell)$ be the union of $T_r(m, \ell)$ and a set of $n - m$ isolated vertices. Notice that the inequality in Theorem 1.3 is tight for $T_r(n, m, r)$ if m is a multiple of r.

We prove a corresponding stability result for Theorem 1.3.

Theorem 1.4. *Let $r \geq 2$ be an integer. For every $\delta > 0$ there exist $\epsilon > 0$ and n_0 such that the following holds for all real numbers $x \geq n_0$. Suppose that $\mathcal{H}$ is a cancellative r-graph with*

$$|\partial\mathcal{H}| = \frac{x^{r-1}}{r^{r-2}} \quad \text{and} \quad |\mathcal{H}| \geq (1-\epsilon)\left(\frac{x}{r}\right)^r. \tag{1}$$

Then there exists a set $V' \subset V(\mathcal{H})$ of size at most $\lceil x \rceil$ such that $\mathcal{H}$ is a subgraph of a complete r-partite r-graph on V' after removing at most δx^r edges.

1.2 Expansion of cliques

For $\ell \geq r \geq 2$ let $\mathcal{K}^r_{\ell+1}$ denote the collection of all r-graphs F with at most $\binom{\ell+1}{2}$ edges such that for some $(\ell+1)$-set S (called the core of F), every pair $\{u, v\} \subset S$ is covered by an edge in F. Let the r-graph $H^r_{\ell+1}$ be obtained from the complete graph $K_{\ell+1}$ by adding $r-2$ new vertices to every edge. The r-graph $H^r_{\ell+1}$ is called the *expansion* of $K_{\ell+1}$, and it is easy to see that $H^r_{\ell+1} \in \mathcal{K}^r_{\ell+1}$. Also note that $\{H^2_{\ell+1}\} = \mathcal{K}^2_{\ell+1} = \{K_{\ell+1}\}$.

The family $\mathcal{K}^r_{\ell+1}$ was introduced by Mubayi in [21] as a way to extend Turán's theorem to hypergraphs. He proved that $\mathrm{ex}(n, \mathcal{K}^r_{\ell+1}) = t_r(n, \ell)$, and moreover, $T_r(n, \ell)$ is the unique $\mathcal{K}^r_{\ell+1}$-free r-graph on n vertices with exactly $t_r(n, \ell)$ edges. Pikhurko [25] improved this result by showing that $\mathrm{ex}(n, H^r_{\ell+1}) = t_r(n, \ell)$ for sufficiently large n and $T_r(n, \ell)$ is also the unique $H^r_{\ell+1}$-free r-graph on n vertices with exactly $t_r(n, \ell)$ edges. One key tool used by Pikhurko is the following stability theorem, which extends the Erdős-Simonovits stability theorem for graphs [28].

Theorem 1.5 (Stability, see [21, 25, 15]). *Let $\ell \geq r \geq 2$ be integers. For every $\delta > 0$ there exists $\epsilon > 0$ and $n_0 = n_0(\ell, r, \delta)$ such that the following holds for all $n \geq n_0$. Suppose that $\mathcal{H}$ is an $H^r_{\ell+1}$-free r-graph on n vertices with at least $(1-\epsilon)t_r(n, \ell)$ edges. Then $\mathcal{H}$ is a subgraph of $T_r(n, \ell)$ after removing at most δn^r edges.*

In [17] the following Kruskal-Katona type result was proved for $\mathcal{K}^r_{\ell+1}$-free r-graphs.

Theorem 1.6 ([17]). *Let $\ell \geq r \geq 2$ be integers and $\mathcal{H}$ be a $\mathcal{K}^r_{\ell+1}$-free r-graph. Suppose that $|\partial\mathcal{H}| = \binom{\ell}{r-1}(x/\ell)^{r-1}$ for some real number $x \geq \ell$. Then $|\mathcal{H}| \leq \binom{\ell}{r}(x/\ell)^r$. In other words,*

$$|\mathcal{H}| \leq \binom{\ell}{r}\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{r/(r-1)}.$$

Note that the inequality in Theorem 1.6 is tight for the r-graph $T_r(n, m, \ell)$ if m is a multiple of ℓ.

We prove the following stability result for Theorem 1.6.

Theorem 1.7. *Let $\ell \geq r \geq 2$ be integers. For $\delta > 0$ there exist $\epsilon > 0$ and n_0 such that the following holds for all real numbers $x \geq n_0$. Suppose that $\mathcal{H}$ is an $\mathcal{K}^r_{\ell+1}$-free r-graph with*

$$|\partial\mathcal{H}| = \binom{\ell}{r-1}\left(\frac{x}{\ell}\right)^{r-1} \quad \text{and} \quad |\mathcal{H}| \geq (1-\epsilon)\binom{\ell}{r}\left(\frac{x}{\ell}\right)^r. \tag{2}$$

Then there exists a set $V' \subset V(\mathcal{H})$ of size at most $\lceil x \rceil$ such that $\mathcal{H}$ is a subgraph of a complete ℓ-partite r-graph on V' after removing at most δx^r edges.

This paper is organized as follows. In Section 2, we prove Theorem 1.4. Section 3, we prove Theorem 1.7. In Section 4, we include some concluding remarks.

2 Cancellative hypergraphs

We prove Theorem 1.4 in this section.

2.1 Preliminaries

For every integer $i \in [r-1]$ the *i-th shadow* $\partial_i \mathcal{H}$ of an r-graph $\mathcal{H}$ is

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

The $(r-1)$-graph $\partial_1 \mathcal{H}$ is also called the shadow of $\mathcal{H}$ and denoted by $\partial \mathcal{H}$. For a vertex $v \in V(\mathcal{H})$ the *link* of v in $\mathcal{H}$ is

$$L_{\mathcal{H}}(v) = \{A \in \partial \mathcal{H} \colon \{v\} \cup A \subset \mathcal{H}\}.$$

The degree of v is $d_{\mathcal{H}}(v) = |L_{\mathcal{H}}(v)|$. For a set $S \subset V(\mathcal{H})$ of size at most $r-1$ the *neighborhood* of S is

$$N_{\mathcal{H}}(S) = \{v \in V(\mathcal{H}) \setminus S \colon \exists A \in \mathcal{H} \text{ such that } S \cup \{v\} \subset A\}.$$

When $S = \{v\}$ for some $v \in V(\mathcal{H})$ we write $N_{\mathcal{H}}(v)$ instead of $N_{\mathcal{H}}(\{v\})$. The *degree sum* of S is defined as

$$\sigma_{\mathcal{H}}(S) = \sum_{v \in S} d_{\mathcal{H}}(v).$$

For every hypergraph $\mathcal{H}$ let

$$\hat{\sigma}_{\mathcal{H}} = \max\{\sigma_{\mathcal{H}}(E) \colon E \in \mathcal{H}\}.$$

We will omit the subscript $\mathcal{H}$ from the notations above if it is clear from context.

We say a set S of vertices is 2-*covered* in $\mathcal{H}$ if every pair of vertices in S is contained in some edge of $\mathcal{H}$, or equivalent, if S induces a complete graph in the graph $\partial_{r-2}\mathcal{H}$.

For cancellative hypergraphs we have the following lemma for 2-covered sets.

Lemma 2.1 ([17]). *Let $r \geq 2$ be an integer and $\mathcal{H}$ be a cancellative r-graph. Suppose that $S \subset V(\mathcal{H})$ is a 2-covered set. Then $L(v) \cap L(u) = \emptyset$ for every pair $\{u, v\} \subset S$. In particular, $\sigma(S) = \sum_{v \in S} d(v) \leq |\partial \mathcal{H}|$.*

The following inequalities which can be found in [17, Section 4.1] will be important for our proofs.

Lemma 2.2 ([17]). *Let $r \geq 2$ be an integer, $\mathcal{H}$ be a cancellative r-graph, and $E \in \mathcal{H}$ be an edge with $\sigma(E) = \hat{\sigma}$. Then*

$$|\mathcal{H}| \leq \frac{|\partial\mathcal{H}|^{\frac{r-2}{r-1}}}{r(r-1)^{1/(r-1)}} \left(\left(|\partial\mathcal{H}| - \frac{\hat{\sigma}}{r} \right) \hat{\sigma} \right)^{\frac{1}{r-1}}, \tag{3}$$

$$|\mathcal{H}| \leq \frac{|\partial\mathcal{H}|^{\frac{r-2}{r-1}}}{r(r-1)^{1/(r-1)}} \left(\sum_{v\in E} d(v) \left(\hat{\sigma} - d(v)\right) + \left(|\partial\mathcal{H}| - \hat{\sigma}\right) \hat{\sigma} \right)^{\frac{1}{r-1}}, \tag{4}$$

$$|\mathcal{H}| \leq \frac{|\partial\mathcal{H}|^{\frac{r-2}{r-1}}}{r(r-1)^{1/(r-1)}} \left(\sum_{v\in E} \sum_{S\in L(v)} \sigma(S) + \sum_{S\in\partial\mathcal{H}\setminus\bigcup_{v\in E} L(v)} \sigma(S) \right)^{\frac{1}{r-1}}, \tag{5}$$

and

$$\frac{1}{r(r-1)} \sum_{v\in V(\mathcal{H})} (d(v))^{\frac{1}{r-1}} |\partial L(v)| \leq \left(\frac{|\partial\mathcal{H}|}{r} \right)^{\frac{r}{r-1}}. \tag{6}$$

We will also use the following property of cancellative hypergraphs.

Lemma 2.3. *Let $r \geq 2$ be an integer and $\mathcal{H}$ be a cancellative r-graph. Then for every $v \in V(\mathcal{H})$ and every $A \in L(v)$ we have $N(A) \cap N(v) = \emptyset$. In other words, $N(A) \subset V(\mathcal{H}) \setminus N(A)$.*

Proof of Lemma 2.3. Suppose to the contrary that there exists $A \in L(v)$ and $u \in N(v)$ such that $u \in N(A)$. Since $|A| = r-1$, $|A \cup \{u\}| = r$. So, $A \cup \{u\}$ is an edge in $\mathcal{H}$. On the other hand, since $A \in L(v)$, $A \cup \{v\}$ is also an edge in $\mathcal{H}$. However, $u \in N(v)$ implies that there is an edge $B \in \mathcal{H}$ such that $\{u, v\} \subset B$, which contradicts the assumption that $\mathcal{H}$ is cancellative. ∎

The following easy lemma will help us to simplify some calculations.

Lemma 2.4. *Let V be a finite set, $f\colon V \to \mathbb{R}$ be a map, and $\delta_1, \delta_2 > 0$ be two real numbers. Let $\bar{f} = \left(\sum_{v\in V} f(v)\right)/|V|$ be the average value of f on V, and suppose that $\max_{v\in V}\{f(v)\} \leq \bar{f} + \delta_2$. Then the set $V_s = \{v \in V : f(v) \leq \bar{f} - \delta_1\}$ satisfies*

$$|V_s| \leq \frac{\delta_2}{\delta_1 + \delta_2}|V|.$$

Proof of Lemma 2.4. By assumption,

$$\begin{aligned} |V|\bar{f} = \sum_{v\in V} f(v) = \sum_{v'\in V_s} f(v') + \sum_{v\in V\setminus V_s} f(v) &\leq |V_s| \left(\bar{f} - \delta_1\right) + \left(|V| - |V_s|\right) \left(\bar{f} + \delta_2\right) \\ &= |V|\bar{f} + \delta_2|V| - (\delta_1 + \delta_2)\,|V_s|, \end{aligned}$$

which implies that $|V_s| \leq \delta_2|V|/(\delta_1 + \delta_2)$. ∎

For two nonnegative numbers $x, y \in \mathbb{R}$ and $\epsilon \in [0, 1]$ we write $x = (1 \pm \epsilon)y$ if x satisfies $(1-\epsilon)y \leq x \leq (1+\epsilon)y$.

2.2 Proof of Theorem 1.4

We prove Theorem 1.4 in this section. The most technical parts are contained in proofs of Lemma 2.5 and Lemma 2.6. In Lemma 2.5 we show that the proof of Theorem 1.4 can be reduced to the same problem with two extra assumptions that $|\partial\mathcal{H}|$ is close to $|\partial T_r(\lceil x\rceil, r)|$ and $|\mathcal{H}|$ is close to $|T_r(\lceil x\rceil, r)|$. In Lemma 2.6 we prove the stability result for hypergraphs with these two extra assumptions.

Lemma 2.5. *Let $r \geq 2$ be an integer, $\epsilon > 0$ be a sufficiently small constant, and $x > 0$ be sufficiently large real number. Suppose that $\mathcal{H}$ is a cancellative r-graph satisfying (1). Then there exists a set $U \subset V(\mathcal{H})$ of size $(1 \pm \epsilon_1)x$ such that the induced subgraph of $\mathcal{H}$ on U satisfies*

$$|\partial(\mathcal{H}[U])| \geq (1-\epsilon_1)\frac{x^{r-1}}{r^{r-2}} \quad \text{and} \quad |\mathcal{H}[U]| \geq (1-\epsilon_1)\left(\frac{x}{r}\right)^r,$$

where $\epsilon_1 = 35r^4\epsilon^{1/2}$.

Lemma 2.6. *Let $r \geq 2$ be an integer. 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*

$$|\partial\mathcal{H}| = (1 \pm \epsilon)\frac{n^{r-1}}{r^{r-2}} \quad \text{and} \quad |\mathcal{H}| \geq (1-\epsilon)\left(\frac{n}{r}\right)^r. \tag{7}$$

Then $\mathcal{H}$ is a subgraph of $T_r(n, r)$ after removing at most δn^r edges.

First let us show that Lemmas 2.5 and 2.6 imply Theorem 1.4.

Proof of Theorem 1.4 using Lemmas 2.5 and 2.6. Fix the integer $r \geq 2$ and the constant $\delta > 0$. Let $\epsilon > 0$ be a sufficiently small constant (whose value can be determined in the following proof) and $x > 0$ be a sufficiently large real number. Let $\mathcal{H}$ be a cancellative r-graph satisfying (1). By Lemma 2.5 there exists a set $U \subset V(\mathcal{H})$ of size $(1 \pm \epsilon_1)x$ such that the induced subgraph $\mathcal{H}$ satisfies

$$(1-\epsilon_1)\frac{x^{r-1}}{r^{r-2}} \leq |\partial(\mathcal{H}[U])| \leq \frac{x^{r-1}}{r^{r-2}} \quad \text{and} \quad |\mathcal{H}[U]| \geq (1-\epsilon_1)\left(\frac{x}{r}\right)^r,$$

where $\epsilon_1 = 40r^{2r}\epsilon^{1/2}$. Let $m = |U| = (1 \pm \epsilon_1)x$. Then $x = (1 \pm 2\epsilon_1)m$ and the inequalities above imply that

$$(1-4r\epsilon_1)\frac{m^{r-1}}{r^{r-2}} \leq |\partial(\mathcal{H}[U])| \leq (1+4r\epsilon_1)\frac{m^{r-1}}{r^{r-2}} \quad \text{and} \quad |\mathcal{H}[U]| \geq (1-4r\epsilon_1)\left(\frac{m}{r}\right)^r,$$

Lemma 2.6 applied to the r-graph $\mathcal{H}[U]$ shows that $\mathcal{H}[U]$ is a subgraph of $T_r(m, r)$ after removing at most $\delta_1 m^r \leq 2\delta_1 x^r$ edges, where $\delta_1 = \delta_1(4r\epsilon_1)$ is the constant guaranteed by Lemma 2.6. If $m \leq \lceil x\rceil$, then let $V' = U$ and we are done. Otherwise we replace U by any $\lceil x\rceil$-subset V' of it, and since $m \leq (1+\epsilon_1)x$, we only loss at most $\epsilon_1 x^r$ edges. Therefore, we can remove at most $\epsilon_1 (x/r)^r + 2\delta_1 x^r + \epsilon_1 x^r$ edges from $\mathcal{H}$ to obtain an r-partite r-graph on at most $\lceil x\rceil$ vertices. ∎

2.3 Proof of Lemma 2.5

We prove Lemma 2.5 in this section. Recall that for a hypergraph $\mathcal{H}$ and a set $S \subset V(\mathcal{H})$

$$\sigma_{\mathcal{H}}(S) = \sum_{v \in S} d_{\mathcal{H}}(v), \quad \text{and} \quad \hat{\sigma}_{\mathcal{H}} = \max\left\{\sigma_{\mathcal{H}}(E) \colon E \in \mathcal{H}\right\}.$$

The subscript $\mathcal{H}$ will be omitted in the following proof.

Proof of Lemma 2.5. We may assume that $r \geq 3$ since the case $r = 2$ follows from the Erdős-Simonovits stability theorem [28] (for K_3-free graphs). Let $\epsilon > 0$ be sufficiently small, $x \geq 0$ be sufficiently large, and $\mathcal{H}$ be a cancellative r-graph satisfying assumptions in Lemma 2.5. Fix an edge $E \in \mathcal{H}$ with $\sigma(E) = \hat{\sigma}$.

Claim 2.7. *We have* $(1 - 2r\epsilon)\,|\partial\mathcal{H}| < \hat{\sigma} \leq |\partial\mathcal{H}|$.

Proof of Claim 2.7. The inequality $\hat{\sigma} \leq |\partial\mathcal{H}|$ follows from Lemma 2.1, so we may focus on the lower bound for $\hat{\sigma}$. It follows from (1) and (3) that

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

So,

$$\begin{aligned}\left(|\partial\mathcal{H}| - \frac{\hat{\sigma}}{r}\right)\hat{\sigma} &\geq (1-\epsilon)^{r-1}\frac{r-1}{r}|\partial\mathcal{H}|^2 \geq (1-(r-1)\epsilon)\frac{r-1}{r}|\partial\mathcal{H}|^2 \\ &= \frac{r-1}{r}|\partial\mathcal{H}|^2 - \epsilon\frac{(r-1)^2}{r}|\partial\mathcal{H}|^2.\end{aligned}$$

Solving this quadratic inequality we obtain that $\hat{\sigma} \leq (1 - 2r\epsilon)\,|\partial\mathcal{H}|$ (the other solution is greater than $|\partial\mathcal{H}|$, which is not possible). ∎

Claim 2.8. *We have* $|d(v) - \hat{\sigma}/r| < 2r\epsilon^{1/2}\hat{\sigma}$ *for every vertex* $v \in E$.

Proof of Claim 2.8. First, we prove that

$$\sum_{v \in E} d(v)\left(\hat{\sigma} - d(v)\right) > \left(\frac{r-1}{r} - 2r\epsilon\right)\hat{\sigma}^2. \tag{8}$$

Suppose that (8) is not true. Then by (4),

$$\begin{aligned}|\mathcal{H}| &\leq \frac{|\partial\mathcal{H}|^{\frac{r-2}{r-1}}}{r(r-1)^{1/(r-1)}}\left(\sum_{v \in E} d(v)\left(\hat{\sigma} - d(v)\right) + \left(|\partial\mathcal{H}| - \hat{\sigma}\right)\hat{\sigma}\right)^{1/(r-1)} \\ &\leq \frac{|\partial\mathcal{H}|^{\frac{r-2}{r-1}}}{r(r-1)^{1/(r-1)}}\left(\left(\frac{r-1}{r} - 2r\epsilon\right)\hat{\sigma}^2 + \left(|\partial\mathcal{H}| - \hat{\sigma}\right)\hat{\sigma}\right)^{1/(r-1)} \\ &\leq \frac{|\partial\mathcal{H}|^{\frac{r-2}{r-1}}}{r(r-1)^{1/(r-1)}}\left(\left(|\partial\mathcal{H}| - \left(\frac{1}{r} + 2r\epsilon\right)\hat{\sigma}\right)\hat{\sigma}\right)^{1/(r-1)}\end{aligned}$$

It follows from $\hat{\sigma} \leq |\partial\mathcal{H}|$ (Claim 2.7) that

$$\begin{aligned}|\mathcal{H}| &\leq \frac{|\partial\mathcal{H}|^{\frac{r-2}{r-1}}}{r(r-1)^{1/(r-1)}}\left(\left(\frac{r-1}{r} - 2r\epsilon\right)|\partial\mathcal{H}|^2\right)^{\frac{1}{r-1}} \\ &= \left(1 - \frac{2r^2}{r-1}\epsilon\right)^{\frac{1}{r-1}}\left(\frac{|\partial\mathcal{H}|}{r}\right)^{\frac{r}{r-1}} < (1-\epsilon)\left(\frac{|\partial\mathcal{H}|}{r}\right)^{\frac{r}{r-1}},\end{aligned}$$

a contradiction. Therefore, (8) is true.

Now suppose that Claim 2.8 is not true. Assume that $E = \{v_1, \dots, v_r\}$ and without loss of generality we may assume that $|d(v_1) - \hat{\sigma}/r| \geq 2r\epsilon^{1/2}\hat{\sigma}$. Then by Jensen's inequality

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

which contradicts (8). ∎

For every $v \in E$ let

$$\mathcal{L}_v = \left\{S \in L(v) \colon \sigma(S) \geq \left(1 - \epsilon^{1/2}\right)\left(\hat{\sigma} - d(v)\right)\right\}.$$

Claim 2.9. *We have* $|\mathcal{L}_v| \geq (1 - 4r^2\epsilon^{1/2})d(v)$ *for every* $v \in E$.

Proof of Claim 2.9. First we show that for every $v \in E$

$$\sum_{S\in L(v)} \sigma(S) \geq (1 - 4r^2\epsilon) d(v)\left(\hat{\sigma} - d(v)\right). \tag{9}$$

Suppose that (9) is not true and let $u \in E$ be a counterexample. Then

$$\begin{aligned}\sum_{v\in E}\sum_{S\in L(v)} \sigma(S) &= \sum_{S\in L(u)} \sigma(S) + \sum_{v\in E\setminus\{u\}}\sum_{S\in L(v)} \sigma(S) \\ &\leq (1 - 4r^2\epsilon) d(u)\left(\hat{\sigma} - d(u)\right) + \sum_{v\in E\setminus\{u\}} d(v)\left(\hat{\sigma} - d(v)\right) \\ &\leq (1 - 2r\epsilon)\sum_{v\in E} d(v)\left(\hat{\sigma} - d(v)\right) \\ &\quad + 2r\epsilon\sum_{v\in E} d(v)\left(\hat{\sigma} - d(v)\right) - 4r^2\epsilon d(u)\left(\hat{\sigma} - d(u)\right).\end{aligned}$$

Due to Claim 2.8, it is easy to see that $\sum_{v\in E} d(v)\left(\hat{\sigma} - d(v)\right) < 2rd(u)\left(\hat{\sigma} - d(u)\right)$. Therefore, by Jensen's inequality,

$$\begin{aligned}\sum_{v\in E}\sum_{S\in L(v)} \sigma(S) &\leq (1 - 2r\epsilon)\sum_{v\in E} d(v)\left(\hat{\sigma} - d(v)\right) \\ &\leq (1 - 2r\epsilon)\left(\sum_{v\in E} d(v)\right)\left(\hat{\sigma} - \frac{\sum_{v\in E} d(v)}{r}\right) = (1 - 2r\epsilon)\frac{r-1}{r}\hat{\sigma}^2.\end{aligned}$$

Then it follows from (5) that

$$
\begin{aligned}
|\mathcal{H}| &\le \frac{1}{r(r-1)^{\frac{1}{r-1}}}|\partial\mathcal{H}|^{\frac{r-2}{r-1}}\left(\sum_{v\in E}\sum_{S\in L(v)}\sigma(S)+\sum_{S\in\partial\mathcal{H}\setminus\bigcup_{v\in E}L(v)}\sigma(S)\right)^{\frac{1}{r-1}}\\
&\le \frac{1}{r(r-1)^{\frac{1}{r-1}}}|\partial\mathcal{H}|^{\frac{r-2}{r-1}}\left((1-2r\epsilon)\frac{r-1}{r}\hat{\sigma}^2+(|\partial\mathcal{H}|-\hat{\sigma})\,\hat{\sigma}\right)^{\frac{1}{r-1}}\\
&\le \frac{1}{r(r-1)^{\frac{1}{r-1}}}|\partial\mathcal{H}|^{\frac{r-2}{r-1}}\left(\left(|\partial\mathcal{H}|-\left(\frac{1}{r}+2(r-1)\epsilon\right)\hat{\sigma}\right)\hat{\sigma}\right)^{\frac{1}{r-1}}.
\end{aligned}
$$

Then, it follows from $\hat{\sigma}\le|\partial\mathcal{H}|$ (Claim 2.7) that

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

a contradiction. Therefore, (9) holds for every $v\in E$. Then, apply Lemma 2.4 with $V=L(v)$ and $f(A)=\sigma(A)$ for every $A\in L(v)$ we obtain

$$
\begin{aligned}
|L(v)\setminus\mathcal{L}_v| &\le \frac{(\hat{\sigma}-d(v))-\sum_{S\in L(v)}\sigma(S)/d(v)}{(\hat{\sigma}-d(v))-\left(1-\epsilon^{1/2}\right)(\hat{\sigma}-d(v))}\cdot d(v)\\
&\le \frac{(\hat{\sigma}-d(v))-(1-4r^2\epsilon)\,(\hat{\sigma}-d(v))}{\epsilon^{1/2}\,(\hat{\sigma}-d(v))}\cdot d(v)\le 4r^2\epsilon^{1/2}d(v).
\end{aligned}
$$

The completes the proof of Claim 2.9. ∎

Let

$$
\mathcal{G}=\left\{S\in\partial\mathcal{H}\colon\sigma(S)\ge\left(\frac{r-1}{r}-2r\epsilon^{1/2}\right)|\partial\mathcal{H}|\right\}.
$$

Claim 2.10. *We have* $|\mathcal{G}|\ge(1-8r^2\epsilon^{1/2})|\partial\mathcal{H}|$.

Proof of Claim 2.10. By Claims 2.8 and 2.7, for every $v\in E$ and $S\in\mathcal{L}_v$ we have

$$
\begin{aligned}
\sigma(S)\ge\left(1-\epsilon^{1/2}\right)(\hat{\sigma}-d(v)) &\ge \left(1-\epsilon^{1/2}\right)\left(\frac{r-1}{r}-2r\epsilon^{1/2}\right)\hat{\sigma}\\
&\ge \left(1-\epsilon^{1/2}\right)\left(\frac{r-1}{r}-2r\epsilon^{1/2}\right)(1-2r\epsilon)\,|\partial\mathcal{H}|\\
&\ge \left(\frac{r-1}{r}-2r\epsilon^{1/2}\right)|\partial\mathcal{H}|.
\end{aligned}
$$

On the other hand, by Claims 2.9 and 2.7,

$$
\sum_{v\in E}|\mathcal{L}_v|\ge\sum_{v\in E}(1-4r^2\epsilon^{1/2})d(v)=(1-4r^2\epsilon^{1/2})\hat{\sigma}\ge(1-8r^2\epsilon^{1/2})|\partial\mathcal{H}|.
$$

Therefore, by Lemma 2.1, $|\mathcal{G}|\ge\sum_{v\in E}|\mathcal{L}_v|\ge(1-8r^2\epsilon^{1/2})|\partial\mathcal{H}|$. ∎

Claim 2.11. *We have* $\Delta(\mathcal{H})\le\left(\frac{1}{r}+3r\epsilon^{1/2}\right)|\partial\mathcal{H}|$.

Proof of Claim 2.11. Suppose to the contrary that there exists a vertex $u \in V(\mathcal{H})$ with

$$d(u) > \left(\frac{1}{r} + 3r\epsilon^{1/2}\right)|\partial\mathcal{H}|.$$

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

$$\sigma(S) \le \hat{\sigma} - d(u) < |\partial\mathcal{H}| - \left(\frac{1}{r} + 3r\epsilon^{1/2}\right)|\partial\mathcal{H}| = \left(\frac{r-1}{r} - 3r\epsilon^{1/2}\right)|\partial\mathcal{H}|.$$

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

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

which contradicts Claim 2.10. ∎

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

Claim 2.12. *We have* $|U| \le \left(1 + 6r^3\epsilon^{1/2}\right) r^{\frac{r-2}{r-1}} |\partial\mathcal{H}|^{\frac{1}{r-1}}$.

Proof of Claim 2.12. First we show that for every $v \in U$,

$$d(v) \ge \left(\frac{1}{r} - 3r^2\epsilon^{1/2}\right)|\partial\mathcal{H}|. \tag{10}$$

Suppose that (10) is not true and let $u \in U$ be a counterexample. Then choose a set $S \in \mathcal{G}$ such that $u \in S$. By the definition of $\mathcal{G}$,

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

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

$$\begin{aligned} d(u') \ge \frac{\sigma(S) - d(u)}{r-2} &> \frac{\left((r-1)/r - 2r\epsilon^{1/2}\right)|\partial\mathcal{H}| - \left(1/r - 3r^2\epsilon^{1/2}\right)|\partial\mathcal{H}|}{r-2} \\ &> \left(\frac{1}{r} + 3r\epsilon^{1/2}\right)|\partial\mathcal{H}|, \end{aligned}$$

which contradicts Claim 2.11. Therefore, (10) holds for every $v \in U$, and it follows from $\sum_{v\in U} d(v) \le r|\mathcal{H}|$ and Theorem 1.3 that

$$|U| \le \frac{r|\mathcal{H}|}{\left(1/r - 3r^2\epsilon^{1/2}\right)|\partial\mathcal{H}|} \le \frac{r\left(|\partial\mathcal{H}|/r\right)^{\frac{r}{r-1}}}{\left(1/r - 3r^2\epsilon^{1/2}\right)|\partial\mathcal{H}|} < \left(1 + 6r^3\epsilon^{1/2}\right) r^{\frac{r-2}{r-1}} |\partial\mathcal{H}|^{\frac{1}{r-1}}.$$

∎

Claim 2.13. *We have* $|\mathcal{H}[U]| \ge \left(1 - 33r^4\epsilon^{1/2}\right)\left(\frac{|\partial\mathcal{H}|}{r}\right)^{\frac{r}{r-1}}$.

Proof of Claim 2.13. By (10) and Claim 2.10, for every $u \in U$ we have

$$\begin{aligned} d_{\mathcal{H}[U]}(u) \ge d_{\mathcal{H}}(u) - |\partial\mathcal{H} \setminus \mathcal{G}| &\ge \left(\frac{1}{r} - 3r^2\epsilon^{1/2}\right)|\partial\mathcal{H}| - 8r^2\epsilon^{1/2}|\partial\mathcal{H}| \\ &= \left(\frac{1}{r} - 11r^2\epsilon^{1/2}\right)|\partial\mathcal{H}|. \end{aligned}$$

For every $0 \le i \le r$ let $\mathcal{E}_i$ be the set of edges in $\mathcal{H}$ that have exactly i vertices in U and note that $\mathcal{E}_r = \mathcal{H}[U]$. Then by Claim 2.11 we have

$$\begin{aligned}\sum_{i\in[r-1]} i|\mathcal{E}_i| = \sum_{u\in U} d_{\mathcal{H}}(u) - r|\mathcal{E}_r| &= \sum_{u\in U} d_{\mathcal{H}}(u) - \sum_{u\in U} d_{\mathcal{H}[U]}(u) \\ &\le \sum_{u\in U} \left(\Delta(\mathcal{H}) - d_{\mathcal{H}[U]}(u)\right) \\ &\le \left(\left(\frac{1}{r} + 3r\epsilon^{1/2}\right)|\partial\mathcal{H}| - \left(\frac{1}{r} - 11r^2\epsilon^{1/2}\right)|\partial\mathcal{H}|\right)|U| \\ &\le 12r^2\epsilon^{1/2}|\partial\mathcal{H}||U|.\end{aligned}$$

It follows from Claim 2.12 that

$$\sum_{i\in[r-1]} i|\mathcal{E}_i| \le 12r^2\epsilon^{1/2}|\partial\mathcal{H}| \cdot \left(1+6r^2\epsilon^{1/2}\right) r^{\frac{r-2}{r-1}}|\partial\mathcal{H}|^{\frac{1}{r-1}} \le 24r^2\epsilon^{1/2}|\partial\mathcal{H}|^{\frac{r}{r-1}}.$$

On the other hand, by Theorem 1.3, $|\mathcal{E}_0| \le (|\partial\mathcal{H}| - |\mathcal{G}|)^{\frac{r}{r-1}} \le 8r^2\epsilon^{1/2}|\partial\mathcal{H}|^{\frac{r}{r-1}}$. Therefore,

$$\begin{aligned}|\mathcal{H}[U]| = |\mathcal{H}| - \sum_{i=0}^{r-1}|\mathcal{E}_i| &\ge (1-\epsilon)\left(\frac{|\partial\mathcal{H}|}{r}\right)^{\frac{r}{r-1}} - 24r^2\epsilon^{1/2}|\partial\mathcal{H}|^{\frac{r}{r-1}} - 8r^2\epsilon^{1/2}|\partial\mathcal{H}|^{\frac{r}{r-1}} \\ &\ge \left(1-33r^4\epsilon^{1/2}\right)\left(\frac{|\partial\mathcal{H}|}{r}\right)^{\frac{r}{r-1}}.\end{aligned}$$

∎

Claim 2.14. *We have* $|U| \ge \left(1-35r^4\epsilon^{1/2}\right) r^{\frac{r-2}{r-1}}|\partial\mathcal{H}|^{\frac{1}{r-1}}$.

Proof of Claim 2.14. It follows from Claims 2.11, 2.13, and $\sum_{u\in U} d_{\mathcal{H}}(u) \ge r|\mathcal{H}[U]|$ that

$$|U| \ge \frac{r|\mathcal{H}[U]|}{\Delta(\mathcal{H})} \ge \frac{r\left(1-33r^4\epsilon^{1/2}\right)(|\mathcal{H}|/r)^{\frac{r}{r-1}}}{\left(1/r+3r\epsilon^{1/2}\right)|\partial\mathcal{H}|} \ge \left(1-35r^4\epsilon^{1/2}\right) r^{\frac{r-2}{r-1}}|\partial\mathcal{H}|^{\frac{1}{r-1}}.$$

∎

Now Claims 2.12 and 2.14 and $|\partial\mathcal{H}| = x^{r-1}/r^{r-2}$ imply that $|U| = (1\pm\epsilon_1)x$. Claim 2.13 shows that

$$|\mathcal{H}[U]| \ge \left(1-33r^4\epsilon^{1/2}\right)\left(\frac{|\partial\mathcal{H}|}{r}\right)^{\frac{r}{r-1}} = \left(1-33r^4\epsilon^{1/2}\right)\left(\frac{x}{r}\right)^r \ge (1-\epsilon_1)\left(\frac{x}{r}\right)^r.$$

Together with Theorem 1.3 we obtain

$$|\partial\left(\mathcal{H}[U]\right)| \ge r|\mathcal{H}[U]|^{\frac{r-1}{r}} \ge (1-\epsilon_1)\frac{x^{r-1}}{r^{r-2}}.$$

∎

2.4 Proof of Lemma 2.6

Proof of Lemma 2.6. The proof if by induction on r. The case $r=2$ follows from the Erdős-Simonovits stability theorem [28] (for K_3-free graphs). So we may assume that

$r \geq 3$. Fix $r \geq 3$ and $\delta > 0$. Let $\epsilon > 0$ be sufficiently small, $x > 0$ be sufficiently large, and $\mathcal{H}$ be a cancellative r-graph satisfying assumptions in Lemma 2.6. Let

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

$$\widehat{V}_L = \left\{ v \in V(\mathcal{H}) \colon d(v) \geq \left(\frac{1}{r} - 3r^2 \epsilon^{1/2} \right) |\partial \mathcal{H}| \right\},$$

$V_S = V(\mathcal{H}) \setminus V_L$, and $\widehat{V}_S = V(\mathcal{H}) \setminus \widehat{V}_L$. It follows from the definition that for every $v \in V_S$,

$$|\partial L(v)| \geq \frac{(r-1)\,(d(v))^{\frac{r-2}{r-1}}}{(1-\epsilon^{1/2})^{\frac{r-2}{r-1}}}. \tag{11}$$

Claim 2.15. *We have* $|\widehat{V}_L| \geq \left(1 - 36r^4\epsilon^{1/2}\right) n$, *and hence* $|\widehat{V}_S| \leq 36r^4\epsilon^{1/2}n$.

Proof of Claim 2.15. Since $|\partial\mathcal{H}| = (1 \pm \epsilon)n^{r-1}/r^{r-2}$ and $|\mathcal{H}| \geq (1-\epsilon)(n/r)^r$, it follows from Claim 2.14 and (10) that there exists a set $U \subset V(\mathcal{H})$ of size at least

$$\left(1 - 35r^4\epsilon^{1/2}\right) r^{\frac{r-2}{r-1}} |\partial\mathcal{H}|^{\frac{1}{r-1}} \geq \left(1 - 35r^4\epsilon^{1/2}\right)(1-\epsilon)\, n \geq \left(1 - 36r^4\epsilon^{1/2}\right) n,$$

such that $d(v) \geq (1/r - 3r^2\epsilon^{1/2})|\partial\mathcal{H}|$ for every $v \in U$. Therefore, $|\widehat{V}_L| \geq |U| \geq \left(1 - 36r^4\epsilon^{1/2}\right) n$, and hence $|\widehat{V}_S| = n - |\widehat{V}_L| \leq 36r^4\epsilon^{1/2}n$. ∎

Claim 2.16. *We have* $|V_L| \geq (1 - 37r^4\epsilon^{1/2})n$.

Proof of Claim 2.16. It is easy to see that for every $v \in V(\mathcal{H})$ the $(r-1)$-graph $L(v)$ is also cancellative, so by Theorem 1.3, $d(v) \leq (|\partial L(v)|/(r-1))^{(r-1)/(r-2)}$. Therefore, by the definition of V_L,

$$\begin{aligned}
|\mathcal{H}| &= \frac{1}{r} \sum_{v \in V(\mathcal{H})} d(v) \\
&= \frac{1}{r} \left(\sum_{v \in V_L} (d(v))^{\frac{1}{r-1}} (d(v))^{\frac{r-2}{r-1}} + \sum_{v \in V_S} (d(v))^{\frac{1}{r-1}} (d(v))^{\frac{r-2}{r-1}} \right) \\
&\leq \frac{1}{r(r-1)} \left(\sum_{v \in V_L} (d(v))^{\frac{1}{r-1}} |\partial L(v)| + (1-\epsilon^{1/2})^{\frac{r-2}{r-1}} \sum_{v \in V_S} (d(v))^{\frac{1}{r-1}} |\partial L(v)| \right) \\
&= \frac{1}{r(r-1)} \sum_{v \in V(\mathcal{H})} (d(v))^{\frac{1}{r-1}} |\partial L(v)| - \frac{1 - (1-\epsilon^{1/2})^{\frac{r-2}{r-1}}}{r(r-1)} \sum_{v \in V_S} (d(v))^{\frac{1}{r-1}} |\partial L(v)|.
\end{aligned}$$

Together with (11) we obtain

$$\begin{aligned}
|\mathcal{H}| &\leq \frac{1}{r(r-1)} \sum_{v \in V(\mathcal{H})} (d(v))^{\frac{1}{r-1}} |\partial L(v)| - \frac{1 - (1-\epsilon^{1/2})^{\frac{r-2}{r-1}}}{(1-\epsilon^{1/2})^{\frac{r-2}{r-1}}} \frac{1}{r} \sum_{v \in V_S} d(v) \\
&\leq \frac{1}{r(r-1)} \sum_{v \in V(\mathcal{H})} (d(v))^{\frac{1}{r-1}} |\partial L(v)| - \frac{\epsilon^{1/2}}{2r} \sum_{v \in V_S} d(v).
\end{aligned}$$

So by (6),

$$|\mathcal{H}| \le \left(\frac{|\partial\mathcal{H}|}{r}\right)^{\frac{r}{r-1}} - \frac{\epsilon^{1/2}}{2r}\sum_{v\in V_S} d(v).$$

Then, it follows from $|\mathcal{H}| > (1-\epsilon)\left(|\partial\mathcal{H}|/r\right)^{r/(r-1)} \ge \left(|\partial\mathcal{H}|/r\right)^{\frac{r}{r-1}} - \epsilon|\partial\mathcal{H}|^{\frac{r}{r-1}}$ and Claim 2.15 that

$$\begin{aligned}\left(\frac{|\partial\mathcal{H}|}{r}\right)^{\frac{r}{r-1}} - \epsilon|\partial\mathcal{H}|^{\frac{r}{r-1}} &\le \left(\frac{|\partial\mathcal{H}|}{r}\right)^{\frac{r}{r-1}} - \frac{\epsilon^{1/2}}{2r}\sum_{v\in V_S} d(v)\\ &\le \left(\frac{|\partial\mathcal{H}|}{r}\right)^{\frac{r}{r-1}} - \frac{\epsilon^{1/2}}{2r}|V_S\setminus\widehat{V}_S|\left(\frac{1}{r}-3r^2\epsilon^{1/2}\right)|\partial\mathcal{H}|\\ &\le \left(\frac{|\partial\mathcal{H}|}{r}\right)^{\frac{r}{r-1}} - \frac{\epsilon^{1/2}}{3r^2}\left(|V_S|-36r^4\epsilon^{1/2}\right)|\partial\mathcal{H}|,\end{aligned}$$

which implies that $|V_S| \le 37r^4\epsilon^{1/2}n$. Therefore, $|V_L| = n - |V_S| \ge \left(1-37r^4\epsilon^{1/2}\right)n$. ∎

Claims 2.15 and 2.16 imply that $|V_L\cap\widehat{V}_L| > \left(1-73r^2\epsilon^{1/2}\right)n$. Then due to $|\mathcal{H}| \ge (1-\epsilon)(n/r)^r$, there exists an edge $\widehat{E}\in\mathcal{H}[V_L\cap\widehat{V}_L]$. By the definition of V_L and $\widehat{V}_L$, for every $v\in\widehat{E}$ we have

$$d(v) \ge (1-\epsilon^{1/2})\left(\frac{|\partial L(v)|}{r-1}\right)^{\frac{r-1}{r-2}}, \tag{12}$$

and

$$d(v) \ge \left(\frac{1}{r}-3r^2\epsilon^{1/2}\right)|\partial\mathcal{H}| \ge \left(\frac{1}{r}-3r^2\epsilon^{1/2}\right)(1-\epsilon)\frac{n^{r-1}}{r^{r-2}} \ge \left(1-4r^3\epsilon^{1/2}\right)\frac{n^{r-1}}{r^{r-1}}. \tag{13}$$

On the other hand, since $\sum_{v\in\widehat{E}} d(v) \le |\partial\mathcal{H}|$, (13) implies that for every $v\in\widehat{E}$,

$$d(v) \le |\partial\mathcal{H}| - (r-1)\left(1-4r^3\epsilon^{1/2}\right)\frac{n^{r-1}}{r^{r-1}} \le \left(1+4r^4\epsilon^{1/2}\right)\frac{n^{r-1}}{r^{r-1}}. \tag{14}$$

Notice that $L(v)$ is a cancellative $(r-1)$-graph. So (12) and Theorem 1.3 imply that

$$(1-\epsilon^{1/2})\left(\frac{|\partial L(v)|}{r-1}\right)^{\frac{r-1}{r-2}} \le |L(v)| \le \left(\frac{|\partial L(v)|}{r-1}\right)^{\frac{r-1}{r-2}}. \tag{15}$$

On the other hand, (13) and (14) give

$$\left(1-4r^3\epsilon^{1/2}\right)\frac{n^{r-1}}{r^{r-1}} \le |L(v)| \le \left(1+4r^4\epsilon^{1/2}\right)\frac{n^{r-1}}{r^{r-1}}. \tag{16}$$

Combining (15) with (16) we obtain

$$\left(1-5r^4\epsilon^{1/2}\right)(r-1)\left(\frac{n}{r}\right)^{r-2} \le |\partial L(v)| \le \left(1+5r^4\epsilon^{1/2}\right)(r-1)\left(\frac{n}{r}\right)^{r-2}. \tag{17}$$

Let x be the real number such that $|\partial L(v)| = x^{r-2}/(r-1)^{r-3}$, and for convenience let us assume that x is an integer. Then (17) implies that

$$\left(1-5r^4\epsilon^{1/2}\right)\frac{r-1}{r}n \le x \le \left(1+5r^4\epsilon^{1/2}\right)\frac{r-1}{r}n. \tag{18}$$

Now Lemma 2.5 applied to $L(v)$ implies that there exists a set $U_v \subset N(v) \subset V(\mathcal{H})$ of size $(1 \pm \epsilon_1)x$ (and to keep our calculations simple let us assume that $|U_v| = x$) such that

$$|L(v)[U_v]| \geq (1 - \epsilon_1)\,|L(v)| \geq (1 - 2\epsilon_1)\left(\frac{x}{r-1}\right)^{r-1}, \tag{19}$$

and

$$|\partial\,(L(v)[U_v])\,| \geq (1 - \epsilon_1)\,|\partial L(v)| \geq (1 - 2\epsilon_1)\,\frac{x^{r-2}}{(r-1)^{r-3}}, \tag{20}$$

where $\epsilon_1 = 35r^4\epsilon^{1/4}$ (the exponent $1/4$ is due to $\epsilon^{1/2}$ in the first inequality in (15)). On the other hand, it follows from (17) and (18) that

$$|\partial\,(L(v)[U_v])\,| \leq |\partial L(v)| \leq (1 + 2\epsilon_1)\,\frac{x^{r-2}}{(r-1)^{r-3}}. \tag{21}$$

By (19), (20), (21), and the induction hypothesis, there exists a partition $U_v = V_1 \cup \cdots \cup V_{r-1}$ such that all but at most $\delta_1 x^{r-1}$ edges in $L(v)[U_v]$ have exactly one vertex in each V_i, where $\delta_1 = \delta_1(r-1, 2\epsilon_1)$ is a sufficiently small constant guaranteed by Lemma 2.6. Let $L'(v) \subset L(v)$ be the collection of edges in $L(v)$ that have exactly one vertex in each V_i. Then by (16) and (19),

$$|L'(v)| \geq |L(v)[U_v]| - \delta_1 x^{r-1} \geq (1 - \epsilon_1)|L(v)| - \delta_1 x^{r-1} \geq (1 - \delta_2)\frac{n^{r-1}}{r^{r-1}}, \tag{22}$$

where $\delta_2 = 5r^3\epsilon^{1/2} + r^r\delta_1$.

Let

$$\mathcal{G} = \left\{A \in \partial\mathcal{H} \colon |N(A)| \geq \left(1 - \epsilon^{1/4}\right)\frac{n}{r}\right\}.$$

Claim 2.17. *We have* $|\mathcal{G}| \geq (1 - 15r^5\epsilon^{1/4})\frac{n^{r-1}}{r^{r-2}}$.

Proof of Claim 2.17. By Lemma 2.3 and (18), for every $v \in \widehat{E}$ and every $A \in L(v)$ we have

$$|N(A)| \leq |V(\mathcal{H}) \setminus N(v)| \leq |V(\mathcal{H}) \setminus U_v| \leq \left(1 + 5r^5\epsilon^{1/2}\right)\frac{n}{r}.$$

Therefore, by (14) and Lemma 2.1, all but at most

$$|\partial\mathcal{H}| - \sum_{v \in \widehat{E}} |L(v)| \leq (1 + \epsilon)\frac{n^{r-1}}{r^{r-2}} - r\left(1 - 4r^3\epsilon^{1/2}\right)\frac{n^{r-1}}{r^{r-1}} \leq 5r^4\epsilon^{1/2}\frac{n^{r-1}}{r^{r-1}}$$

edges $A \in \partial\mathcal{H}$ satisfy $N(A) \leq \left(1 + 5r^5\epsilon^{1/2}\right)n/r$. It follows that

$$\begin{aligned}
r|\mathcal{H}| &= \sum_{A \in \mathcal{G}} N(A) + \sum_{A \in \partial\mathcal{H} \setminus \mathcal{G}} N(A) \\
&\leq |\mathcal{G}|\left(1 + 5r^5\epsilon^{1/2}\right)\frac{n}{r} + 5r^4\epsilon^{1/2}\frac{n^{r-1}}{r^{r-1}} \cdot n + |\partial\mathcal{H} \setminus \mathcal{G}|\left(1 - \epsilon^{1/4}\right)\frac{n}{r} \\
&= |\partial\mathcal{H}|\left(1 - \epsilon^{1/4}\right)\frac{n}{r} + |\mathcal{G}|\left(\epsilon^{1/4} + 5r^5\epsilon^{1/2}\right)\frac{n}{r} + 5r^4\epsilon^{1/2}\frac{n^r}{r^{r-1}} \\
&\leq (1 + \epsilon)\frac{n^{r-1}}{r^{r-2}}\left(1 - \epsilon^{1/4}\right)\frac{n}{r} + |\mathcal{G}|\left(\epsilon^{1/4} + 5r^5\epsilon^{1/2}\right)\frac{n}{r} + 5r^4\epsilon^{1/2}\frac{n^r}{r^{r-1}} \\
&\leq \left(1 - \epsilon^{1/4} + \epsilon\right)\frac{n^r}{r^{r-1}} + |\mathcal{G}|\left(\epsilon^{1/4} + 5r^5\epsilon^{1/2}\right) + 5r^4\epsilon^{1/2}\frac{n^r}{r^{r-1}}.
\end{aligned}$$

Since $|\mathcal{H}| \geq (1-\epsilon)n^r/r^r$, the inequality above implies

$$\left(1-\epsilon^{1/4}+\epsilon\right)\frac{n^r}{r^{r-1}} + |\mathcal{G}|\left(\epsilon^{1/4}+5r^5\epsilon^{1/2}\right)\frac{n}{r} + 5r^4\epsilon^{1/2}\frac{n^r}{r^{r-1}} \leq (1-\epsilon)\frac{n^r}{r^{r-1}}.$$

Therefore,

$$|\mathcal{G}| \geq \frac{\epsilon^{1/4}-2\epsilon-5r^4\epsilon^{1/2}}{\epsilon^{1/4}+5r^5\epsilon^{1/2}}\frac{n^{r-1}}{r^{r-2}} \geq (1-15r^5\epsilon^{1/4})\frac{n^{r-1}}{r^{r-2}}.$$

∎

Now fix $v \in \widehat{E}$ and let $V_r = V(\mathcal{H}) \setminus U_v$. By Lemma 2.3, every edge $A \in L'(v)$ satisfies $N(A) \subset V_r$. By Claim 2.17 all but at most

$$|\partial\mathcal{H}| - |\mathcal{G}| \leq 16r^5\epsilon^{1/4}\frac{n^{r-1}}{r^{r-2}}$$

edges $A \in L'(v)$ satisfy $|N(A)| \geq (1-\epsilon^{1/4})n/r$. Therefore, by (22), the number of edges in $\mathcal{H}$ that have exactly one vertex in each V_i is at least

$$\begin{aligned}\left(|L'(v)| - 16r^5\epsilon^{1/4}\frac{n^{r-1}}{r^{r-2}}\right)(1-\epsilon^{1/4})\frac{n}{r} &\geq \left((1-\delta_2)\frac{n^{r-1}}{r^{r-1}} - 16r^5\epsilon^{1/4}\frac{n^{r-1}}{r^{r-2}}\right)(1-\epsilon^{1/4})\frac{n}{r}\\ &\geq (1-\delta_3)\left(\frac{n}{r}\right)^r,\end{aligned}$$

where $\delta_3 = \delta_2 + 17r^6\epsilon^{1/4}$ (we can choose $\epsilon > 0$ to be sufficiently small such that $\delta_3 \leq \delta$). This completes the proof of Lemma 2.6. ∎

3 Expansion of cliques

3.1 Preliminaries

For an r-graph $\mathcal{H}$ the *clique set* $\mathcal{K}_{\mathcal{H}}$ of $\mathcal{H}$ is

$$\mathcal{K}_{\mathcal{H}} = \left\{A \subset V(\mathcal{H})\colon\ (\partial_{r-2}\mathcal{H})[A] \cong K_{|A|}\right\}.$$

It was prove in [17] that

$$\sigma(S) \leq (\ell-r+1)|\partial\mathcal{H}|, \quad \forall S \in \mathcal{K}_{\mathcal{H}}.$$

Let $z = z(\mathcal{H}) \geq 0$ be the largest real number such that for all $R \in \mathcal{K}_{\mathcal{H}}$ with $|R| \leq \ell - 1$,

$$\sigma(R) \leq (\ell-r+1)\,|\partial\mathcal{H}| - (\ell-|R|)\,z.$$

The following inequalities can be found in [17, Section 5].

Lemma 3.1 ([17])**.** *Let $\mathcal{H}$ be a $\mathcal{K}_{\ell+1}^r$-free r-graph, and $R_0 \in \mathcal{K}_{\mathcal{H}}$ be a set of size at most $\ell-1$ with $\sigma(R_0) = (\ell-r+1)\,|\partial\mathcal{H}| - (\ell-|R|)\,z$, where $z = z(\mathcal{H}) \geq 0$ is defined as above. Then*

$$|\mathcal{H}| \leq \frac{\binom{\ell-1}{r-1}^{\frac{r-2}{r-1}}}{r\binom{\ell-1}{r-2}}(r-1)^{\frac{r-2}{r-1}}|\partial\mathcal{H}|^{\frac{r-2}{r-1}}\left(\sum_{E\in\partial\mathcal{H}}\sigma(E)\right)^{\frac{1}{r-1}}, \tag{23}$$

$$\sum_{E\in\partial\mathcal{H}}\sigma(E) \leq (\ell-r+1)\left(|\partial\mathcal{H}|-2z\right)|\partial\mathcal{H}| + z^2\ell - \frac{\left((\ell-r+1)|\partial\mathcal{H}|-z\ell\right)^2}{|R_0|}. \tag{24}$$

3.2 Proof of Theorem 1.7

The proof is similar to the proof of Theorem 1.4, but it is simpler because we just need to prove a similar result as Lemma 2.5 and then we can use Theorem 1.5 directly [1].

Proof of Theorem 1.7. Fix $\ell \geq r \geq 2$ and $\delta > 0$. Let $\epsilon > 0$ be a sufficiently constant and $x > 0$ be a sufficiently large real number. Let $\mathcal{H}$ be a $\mathcal{K}^r_{\ell+1}$-free r-graph satisfying the assumptions in Theorem 1.7. Notice that the inequality in (2) is equivalent to

$$|\mathcal{H}| \geq (1-\epsilon)\frac{\binom{\ell}{r}}{\binom{\ell}{r-1}^{\frac{r}{r-1}}}|\partial\mathcal{H}|^{\frac{r}{r-1}}. \tag{25}$$

Let $z = z(\mathcal{H})$ be the same as defined in Section 3.1, and let $R_0 \in \mathcal{K}_{\mathcal{H}}$ be a set of size at most $\ell - 1$ with

$$\sigma(R_0) = (\ell - r + 1)\,|\partial\mathcal{H}| - (\ell - |R_0|)\,z.$$

Claim 3.2. *We have*

$$\sum_{E\in\partial\mathcal{H}} \sigma(E) \leq (1 - r\epsilon)\frac{(\ell-r+1)(r-1)}{\ell}|\partial\mathcal{H}|^2. \tag{26}$$

Proof of Claim 3.2. Suppose to the contrary that (26) fails. Then by (23),

$$\begin{aligned}
|\mathcal{H}| &\leq \frac{\binom{\ell-1}{r-1}^{\frac{r-2}{r-1}}}{r\binom{\ell-1}{r-2}}(r-1)^{\frac{r-2}{r-1}}|\partial\mathcal{H}|^{\frac{r-2}{r-1}}\left(\sum_{E\in\partial\mathcal{H}}\sigma(E)\right)^{\frac{1}{r-1}} \\
&\leq (1-r\epsilon)^{\frac{1}{r-1}}\frac{\binom{\ell-1}{r-1}^{\frac{r-2}{r-1}}}{r\binom{\ell-1}{r-2}}(r-1)^{\frac{r-2}{r-1}}|\partial\mathcal{H}|^{\frac{r-2}{r-1}}\frac{(\ell-r+1)^{\frac{1}{r-1}}(r-1)^{\frac{1}{r-1}}}{\ell^{\frac{1}{r-1}}}|\partial\mathcal{H}|^{\frac{2}{r-1}} \\
&= (1-r\epsilon)^{\frac{1}{r-1}}\frac{\binom{\ell}{r}}{\binom{\ell}{r-1}^{\frac{r}{r-1}}}|\partial\mathcal{H}|^{\frac{r}{r-1}} < (1-\epsilon)\frac{\binom{\ell}{r}}{\binom{\ell}{r-1}^{\frac{r}{r-1}}}|\partial\mathcal{H}|^{\frac{r}{r-1}},
\end{aligned}$$

contradicting (25). ∎

Next, we show that z is close to $\frac{\ell-r+1}{\ell}|\partial\mathcal{H}|$.

Claim 3.3. *We have*

$$z = (1 \pm \ell r\epsilon^{1/2})\frac{\ell-r+1}{\ell}|\partial\mathcal{H}|. \tag{27}$$

Proof of Claim 3.3. Suppose to the contrary that (27) fails. Then by (24),

$$\begin{aligned}
\sum_{E\in\partial\mathcal{H}}\sigma(E) &\leq (\ell-r+1)\,(|\partial\mathcal{H}| - 2z)\,|\partial\mathcal{H}| + z^2\ell - \frac{((\ell-r+1)|\partial\mathcal{H}| - z\ell)}{|R_0|} \\
&= \frac{(\ell-r+1)(r-1)}{\ell}|\partial\mathcal{H}|^2 - \ell\left(\frac{\ell}{|R_0|} - 1\right)\left(z - \frac{\ell-r+1}{\ell}|\partial\mathcal{H}|\right)^2 \\
&\leq \frac{(\ell-r+1)(r-1)}{\ell}|\partial\mathcal{H}|^2 - \ell\left(\frac{\ell}{\ell-1} - 1\right)\left(\ell r\epsilon^{1/2}\cdot\frac{\ell-r+1}{\ell}|\partial\mathcal{H}|\right)^2 \\
&< (1-r\epsilon)\frac{(\ell-r+1)(r-1)}{\ell}|\partial\mathcal{H}|^2
\end{aligned}$$

[1] One could also use a similar inductive argument to prove a similar result as Lemma 2.6 to avoid using Theorem 1.5.

contradicting Claim 3.2. ∎

It follows from the definition of z and Claim 3.3 that for every $R \in \mathcal{K}_{\mathcal{H}}$ with $|R| \leq \ell - 1$,

$$\begin{aligned}\sigma(R) &\leq (\ell - r + 1)\,|\partial\mathcal{H}| - (\ell - |R|)\,z \\ &\leq (\ell - r + 1)\,|\partial\mathcal{H}| - (\ell - |R|)\,(1 - \ell r\epsilon^{1/2})\frac{\ell - r + 1}{\ell}|\partial\mathcal{H}| \\ &\leq (1 + \ell^2 r\epsilon^{1/2})\frac{(\ell - r + 1)|R|}{\ell}|\partial\mathcal{H}|.\end{aligned}$$

In particular, for every $v \in V(\mathcal{H})$ we have

$$d(v) = \sigma(\{v\}) \leq (1 + \ell^2 r\epsilon^{1/2})\frac{\ell - r + 1}{\ell}|\partial\mathcal{H}|, \tag{28}$$

and for every $E \in \partial\mathcal{H}$ we have

$$\sigma(E) \leq (1 + \ell^2 r\epsilon^{1/2})\frac{(\ell - r + 1)(r - 1)}{\ell}|\partial\mathcal{H}|. \tag{29}$$

Let

$$\mathcal{G} = \left\{E \in \partial\mathcal{H} \colon \sigma(E) \geq (1 - \epsilon^{1/4})\frac{(\ell - r + 1)(r - 1)}{\ell}|\partial\mathcal{H}|\right\}.$$

Claim 3.4. *We have* $|\mathcal{G}| \geq (1 - \ell^2 r\epsilon^{1/4})|\partial\mathcal{H}|$.

Proof of Claim 3.4. It follows from Lemma 2.4, Claim 3.2, and (29) that

$$\begin{aligned}|\partial\mathcal{H} \setminus \mathcal{G}| &\leq \frac{(1 + \ell^2 r\epsilon^{1/2})(\ell - r + 1)(r - 1)|\partial\mathcal{H}|/\ell - \sum_{E \in \partial\mathcal{H}}\sigma(E)/|\partial\mathcal{H}|}{(1 + \ell^2 r\epsilon^{1/2})(\ell - r + 1)(r - 1)|\partial\mathcal{H}|/\ell - (1 - \epsilon_4)(\ell - r + 1)(r - 1)|\partial\mathcal{H}|/\ell}|\partial\mathcal{H}| \\ &\leq \frac{\ell^2 r\epsilon^{1/2} + r\epsilon}{\ell^2 r\epsilon^{1/2} + \epsilon^{1/4}}|\partial\mathcal{H}| \leq \ell^2 r\epsilon^{1/4}|\partial\mathcal{H}|.\end{aligned}$$

∎

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

Claim 3.5. *For every* $u \in U$ *we have* $d(u) \geq (1 - 2\epsilon^{1/4})\frac{\ell - r + 1}{\ell}|\partial\mathcal{H}|$.

Proof of Claim 3.5. It follows from the definition of U that for every $u \in U$ there exists $E \in \mathcal{G}$ with $u \in E$. Then it follows from the definition of $\mathcal{G}$ that $\sigma(E) \geq (1 - \epsilon^{1/4})(\ell - r + 1)(r - 1)|\partial\mathcal{H}|/\ell$. So by (28)

$$\begin{aligned}d(u) &= \sigma(E) - \sum_{v \in E \setminus \{u\}} d_{\mathcal{H}}(v) \\ &\geq (1 - \epsilon^{1/4})\frac{(\ell - r + 1)(r - 1)}{\ell}|\partial\mathcal{H}| - (r - 2)(1 + \ell^2 r\epsilon^{1/2})\frac{\ell - r + 1}{\ell}|\partial\mathcal{H}| \\ &\geq (1 - 2\epsilon^{1/4})\frac{\ell - r + 1}{\ell}|\partial\mathcal{H}|.\end{aligned}$$

∎

Next we show an upper bound for $|U|$.

Claim 3.6. *We have* $|U| \le (1+4\epsilon^{1/4})\ell \left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{1/(r-1)}$.

Proof of Claim 3.6. Since $\sum_{u\in U} d(u) \le r|\mathcal{H}|$, by Claim 3.5 and Theorem 1.6,

$$|U| \le \frac{r|\mathcal{H}|}{(1-2\epsilon^{1/4})\frac{\ell-r+1}{\ell}|\partial\mathcal{H}|} \le \frac{r\binom{\ell}{r}\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{\frac{r}{r-1}}}{(1-2\epsilon^{1/4})\frac{\ell-r+1}{\ell}|\partial\mathcal{H}|}$$
$$= \frac{\frac{r}{\ell-r+1}\frac{\binom{\ell}{r}}{\binom{\ell}{r-1}}}{(1-2\epsilon^{1/4})}\ell\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{1/(r-1)} \le (1+4\epsilon^{1/4})\ell\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{1/(r-1)}.$$

Here we used the identity $\binom{\ell}{r}/\binom{\ell}{r-1} = (\ell-r+1)/r$. ∎

Claim 3.7. *We have* $|\mathcal{H}[U]| \ge (1-9\ell^{2r}r\epsilon^{1/4})\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{\frac{r}{r-1}}$.

Proof of Claim 3.7. By Claims 3.5 and 3.4, for every $u\in U$ we have

$$d_{\mathcal{H}[U]}(u) \ge d_{\mathcal{H}}(u) - |\partial\mathcal{H}\setminus\mathcal{G}| \ge (1-2\epsilon^{1/4})\frac{\ell-r+1}{\ell}|\partial\mathcal{H}| - \ell^2 r\epsilon^{1/4}|\partial\mathcal{H}|$$
$$\ge (1-3\ell^3 r\epsilon^{1/4})\frac{\ell-r+1}{\ell}|\partial\mathcal{H}|.$$

For every $0\le i\le r$ let $\mathcal{E}_i$ be the set of edges in $\mathcal{H}$ that have exactly i vertices in U and note that $\mathcal{E}_r = \mathcal{H}[U]$. Then by (28) and Claim 3.6,

$$\sum_{i\in[r-1]} i|\mathcal{E}_i| = \sum_{u\in U} d_{\mathcal{H}}(u) - r|\mathcal{E}_r| = \sum_{u\in U} d_{\mathcal{H}}(u) - \sum_{u\in U} d_{\mathcal{H}[U]}(u)$$
$$\le \sum_{u\in U}\big(\Delta(\mathcal{H}) - d_{\mathcal{H}[U]}(u)\big)$$
$$\le (\ell^2 r\epsilon^{1/2} + 3\ell^3 r\epsilon^{1/4})\frac{\ell-r+1}{\ell}|\partial\mathcal{H}||U|$$
$$\le 4\ell^3 r\epsilon^{1/4}\frac{\ell-r+1}{\ell}|\partial\mathcal{H}|\cdot(1+4\epsilon^{1/4})\ell\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{1/(r-1)}$$
$$\le 8\ell^3 r^2\binom{\ell}{r}\epsilon^{1/4}\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{\frac{r}{r-1}}.$$

On the other hand, by Claim 3.4 and Theorem 1.3,

$$|\mathcal{E}_0| \le \binom{\ell}{r}\left(\frac{|\partial\mathcal{H}| - |\mathcal{G}|}{\binom{\ell}{r-1}}\right)^{\frac{r}{r-1}} \le \ell^2 r\epsilon^{1/4}\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{\frac{r}{r-1}}.$$

Therefore,

$$|\mathcal{H}[U]| = |\mathcal{H}| - \sum_{i=0}^{r-1}|\mathcal{E}_i| \ge (1-\epsilon)\binom{\ell}{r}\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{\frac{r}{r-1}} - \left(8\ell^3 r^2\binom{\ell}{r}\epsilon^{1/4} + \ell^2 r\epsilon^{1/4}\right)\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{\frac{r}{r-1}}$$
$$\ge (1-9\ell^3 r^2\epsilon^{1/4})\binom{\ell}{r}\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{\frac{r}{r-1}}.$$

∎

Let $m = |U|$. Then it follows from Claim 3.6 and $|\partial\mathcal{H}| = \binom{\ell}{r-1}(x/\ell)^{r-1}$ that

$$m \le (1+4\epsilon^{1/4})\ell\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{1/(r-1)} \le (1+4\epsilon^{1/4})x.$$

Claim 3.7 implies that

$$\begin{aligned}|\mathcal{H}[U]| &\ge (1-9\ell^3r^2\epsilon^{1/4})\binom{\ell}{r}\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{\frac{r}{r-1}} \\ &\ge (1-9\ell^3r^2\epsilon^{1/4})\binom{\ell}{r}\left(\frac{x}{\ell}\right)^r \\ &\ge (1-9\ell^3r^2\epsilon^{1/4})\frac{1}{(1+4\epsilon^{1/4})^r}\binom{\ell}{r}\left(\frac{m}{\ell}\right)^r \ge (1-10\ell^3r^2\epsilon^{1/4})\binom{\ell}{r}\left(\frac{m}{\ell}\right)^r.\end{aligned}$$

Now Theorem 1.5 applied to $\mathcal{H}[U]$ implies that $\mathcal{H}[U]$ contains a subgraph $\mathcal{H}'$ of size at least

$$\begin{aligned}|\mathcal{H}[U]| - \delta_1 m^r \ge (1-10\ell^3r^2\epsilon^{1/4})\binom{\ell}{r}\left(\frac{m}{\ell}\right)^r - \delta_1 m^r &\ge |\mathcal{H}| - (10\ell^3r^2\epsilon^{1/4}+\delta_1)m^r \\ &\ge |\mathcal{H}| - 2(10\ell^3r^2\epsilon^{1/4}+\delta_1)x^r\end{aligned}$$

such that $\mathcal{H}'$ is a subgraph of $T_r(m,\ell)$. If $m \le \lceil x\rceil$, then let $V' = U$ and we are done. Otherwise let $V' \subset U$ be a subset of size $\lceil x\rceil$. Then due to $m \le (1+4\epsilon^{1/4})x$, the number of edges in $\mathcal{H}'[V']$ is at least $|\mathcal{H}'| - 4\epsilon^{1/4}x^r \ge |\mathcal{H}| - 2(10\ell^3r^2\epsilon^{1/4}+\delta_1)x^r - 4\epsilon^{1/4}x^r$ (we can choose $\epsilon > 0$ to be sufficiently small such that $2(10\ell^3r^2\epsilon^{1/4}+\delta_1)+4\epsilon^{1/4} \le \delta$). This completes the proof of Theorem 1.7. ∎

4 Concluding remarks

In this work we showed some extensions of Keevash's stability result of the Kruskal-Katona theorem to the classes of calcellative hypergraphs and hypergraphs without the expansion of cliques. In general, one could ask whether similar results hold for other $\mathcal{F}$-free hypergraphs. An classical example suggested by Sós is the Fano plane, which is the 3-graph on vertex set [7] with edge set

$$\{123, 345, 561, 174, 275, 376, 246\}.$$

The extremal properties of the Fano plane have been well studies by several authors (see e.g. [6, 13, 7, 2]). However, a Kruskal-Katona type result for the Fano plane is still not known.

It is interesting that the inequality in Theorem 1.3 is tight for every integer $r \ge 2$ while the maximum size of an n-vertex cancellative r-graph is still unknown (even asymptotically) for every $r \ge 5$. This suggests that one could prove a Kruskal-Katona type result for a family $\mathcal{F}$ whose Turán density is not known. A famous example is the complete 3-graph on four vertices K_4^3 (see [29]). It was shown in [17] that the Turán problem for K_4^3 does not have the stability property (assuming the famous Turán tetrahedron conjecture is true). So, it would be very interesting if the Kruskal-Katona type result for K_4^3 has the stability property. There are many other interesting cases one could consider for $\mathcal{F}$, and we refer the reader to the nice survey of Keevash [11] for more details.

It seems that Theorems 1.4 and 1.7 belong to a new type of stability results for $\mathcal{F}$-free hypergrahs, which are different from the stability results proved before (see e.g. [28, 1, 12, 7, 13, 21, 24, 8, 22, 5, 23, 3, 16, 18, 19]). It would be interesting to see whether there are any applications of this new type of stability results.

References

[1] B. Andrásfai, P. Erdős, and V. T. Sós. On the connection between chromatic number, maximal clique and minimal degree of a graph. *Discrete Math.*, 8:205–218, 1974.

[2] L. Bellmann and C. Reiher. Turán's theorem for the Fano plane. *Combinatorica*, 39(5):961–982, 2019.

[3] A. Bene Watts, S. Norin, and L. Yepremyan. A Turán theorem for extensions via an Erdős-Ko-Rado theorem for Lagrangians. *Combinatorica*, 39(5):1149–1171, 2019.

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

[5] A. Brandt, D. Irwin, and T. Jiang. Stability and Turán numbers of a class of hypergraphs via Lagrangians. *Combin. Probab. Comput.*, 26(3):367–405, 2017.

[6] D. De Caen and Z. Füredi. The maximum size of 3-uniform hypergraphs not containing a Fano plane. *J. Combin. Theory Ser. B*, 78(2):274–276, 2000.

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

[8] D. Hefetz and P. Keevash. A hypergraph Turán theorem via Lagrangians of intersecting families. *J. Combin. Theory Ser. A*, 120(8):2020–2038, 2013.

[9] G. Katona. A theorem of finite sets. In *Theory of graphs (Proc. Colloq., Tihany, 1966)*, pages 187–207, 1968.

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

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

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

[13] P. Keevash and B. Sudakov. The Turán number of the Fano plane. *Combinatorica*, 25(5):561–574, 2005.

[14] J. B. Kruskal. The number of simplices in a complex. In *Mathematical optimization techniques*, pages 251–278. Univ. of California Press, Berkeley, Calif., 1963.

[15] X. Liu. New short proofs to some stability theorems. *arXiv preprint arXiv:1903.01606*, 2019.

[16] X. Liu and D. Mubayi. A hypergraph Turán problem with no stability. accepted by Combinatorica.

[17] X. Liu and D. Mubayi. The feasible region of hypergraphs. *J. Comb. Theory, Ser. B*, 148:23–59, 2021.

[18] X. Liu, D. Mubayi, and C. Reiher. Hypergraphs with many extremal configurations. *arXiv preprint arXiv:2102.02103.* submitted.

[19] X. Liu, D. Mubayi, and C. Reiher. A unified approach to hypergraph stability. *arXiv preprint arXiv:2104.00167.* submitted.

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

[21] D. Mubayi. A hypergraph extension of Turán's theorem. *J. Combin. Theory, Ser. B*, 96(1):122–134, 2006.

[22] S. Norin and L. Yepremyan. Turán number of generalized triangles. *J. Combin. Theory Ser. A*, 146:312–343, 2017.

[23] S. Norin and L. Yepremyan. Turán numbers of extensions. *J. Combin. Theory Ser. A*, 155:476–492, 2018.

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

[25] O. Pikhurko. Exact computation of the hypergraph Turán function for expanded complete 2-graphs. *J. Combin. Theory, Ser. B*, 103(2):220–225, 2013.

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

[27] 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.

[28] 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.

[29] P. Turán. On an extermal problem in graph theory. *Mat. Fiz. Lapok*, 48:436–452, 1941.