Combinatorics, Probability and Computing (2017) 26, 367–405. © Cambridge University Press 2017
doi:10.1017/S0963548316000444

Stability and Turán Numbers of a Class of Hypergraphs via Lagrangians

AXEL BRANDT[1†], DAVID IRWIN[2] and TAO JIANG[3‡]

[1]Department of Mathematics and Computer Science, Davidson College, Davidson, NC 28035, USA
(e-mail: axbrandt@davidson.edu)
[2]Department of Mathematics, Ohio State University, Columbus, OH 43210, USA
(e-mail: irwin.315@osu.edu)
[3]Department of Mathematics, Miami University, Oxford, OH 45056, USA
(e-mail: jiangt@miamioh.edu)

Received 11 October 2016; revised 12 November 2016

Given a family of r-uniform hypergraphs $\mathcal{F}$ (or r-graphs for brevity), the Turán number $\mathrm{ex}(n,\mathcal{F})$ of $\mathcal{F}$ is the maximum number of edges in an r-graph on n vertices that does not contain any member of $\mathcal{F}$. A pair $\{u,v\}$ is *covered* in a hypergraph G if some edge of G contains $\{u,v\}$. Given an r-graph F and a positive integer $p \geqslant n(F)$, where $n(F)$ denotes the number of vertices in F, let H_p^F denote the r-graph obtained as follows. Label the vertices of F as $v_1,\dots,v_{n(F)}$. Add new vertices $v_{n(F)+1},\dots,v_p$. For each pair of vertices v_i,v_j not covered in F, add a set $B_{i,j}$ of $r-2$ new vertices and the edge $\{v_i,v_j\}\cup B_{i,j}$, where the $B_{i,j}$ are pairwise disjoint over all such pairs $\{i,j\}$. We call H_p^F *the expanded p-clique with an embedded* F. For a relatively large family of F, we show that for all sufficiently large n, $\mathrm{ex}(n,H_p^F) = |T_r(n,p-1)|$, where $T_r(n,p-1)$ is the balanced complete $(p-1)$-partite r-graph on n vertices. We also establish structural stability of near-extremal graphs. Our results generalize or strengthen several earlier results and provide a class of hypergraphs for which the Turán number is exactly determined (for large n).

2010 *Mathematics subject classification*: Primary 05C65
Secondary 05C35

1. Introduction

Given an r-uniform hypergraph (or r-graph for brevity) G, we use $n(G)$ and $|G|$ to denote the number of vertices and number of edges in G, respectively. Given a family of r-graphs $\mathcal{F}$, the Turán number $\mathrm{ex}(n,\mathcal{F})$ of $\mathcal{F}$ is the maximum number of edges in an r-graph on n vertices that does not contain any member of $\mathcal{F}$. The Turán density $\pi(\mathcal{F})$ of $\mathcal{F}$ is defined to be $\lim_{n\to\infty} \mathrm{ex}(n,\mathcal{F})/\binom{n}{r}$; such a limit is known to exist. Determining Turán numbers of graphs and

† Research supported in part by NSF GK-12 Transforming Experiences Grant DGE-0742434.
‡ Research supported in part by National Science Foundation grant DMS-1400249.

hypergraphs is one of the central problems in extremal combinatorics. For $r = 2$, the problem was asymptotically solved for all non-bipartite graphs in the form of the Erdős–Stone–Simonovits theorem, which states that if $\mathcal{F}$ is a family of graphs and the minimum chromatic number among all members is $p \geqslant 3$, then

$$\pi(\mathcal{F}) = \frac{p-2}{p-1}.$$

For $r \geqslant 3$, not too much is known. There are very few exact or asymptotic results. For a recent survey on hypergraph Turán numbers, the reader is referred to the survey by Keevash [13]. In this paper, we build on earlier works by Sidorenko [24], Pikhurko [20, 21], Mubayi [15] and Mubayi and Pikhurko [16] to obtain a general theorem that determines the exact Turán numbers of a class of hypergraphs for all sufficiently large n. Our main theorems substantially generalize or strengthen several earlier results.

2. History

2.1. Cancellative hypergraphs

The study of Turán numbers dates back to Mantel's theorem, which states that $\mathrm{ex}(n, K_3) = \lfloor n/2 \rfloor \cdot \lceil n/2 \rceil$. Katona [12] suggests an extension of the problem to hypergraphs. An r-graph G is called *cancellative* if, for any three edges A, B, C satisfying $A \cup B = A \cup C$, we have $B = C$. Equivalently, G is cancellative if it does not contain three distinct members A, B, C such that one contains the symmetric difference of the other two. When $r = 2$ the condition is equivalent to saying that G is triangle-free. Katona asked to determine the largest size of a cancellative 3-graph on n vertices. The problem was solved by Bollobás [2], who showed that for all n, the largest size of a cancellative 3-graph on n vertices is the balanced complete 3-partite 3-graph on n vertices. Keevash and Mubayi [14] gave a new proof of Bollobás's result and established stability of near-extremal graphs, showing that all cancellative 3-graphs on n with close to the maximum number of edges must be structurally close to the complete balanced 3-partite 3-graph. Bollobás [2] conjectured that for all $r \geqslant 4$, the largest cancellative r-graph on n vertices is the balanced complete r-partite r-graph on n vertices. This was proved to be true for $r = 4$ by Sidorenko [24]. However, Shearer [22] gave counter-examples showing that the conjecture is false for $r \geqslant 10$.

2.2. Generalized triangles

Frankl and Füredi [7, 8] considered a strengthening of cancellative r-graphs. For each $r \geqslant 2$, let $\sum_r$ consist of all r-graphs with three edges D_1, D_2, D_3 such that $|D_1 \cap D_2| = r - 1$ and $D_1 \triangle D_2 \subseteq D_3$, where $D_1 \triangle D_2$ denotes the symmetric difference of D_1 and D_2. Let the *generalized triangle* T_k be the member of $\sum_r$ with edges

$$\{1, \ldots, r\}, \quad \{1, 2, \ldots, r-1, r+1\} \quad \text{and} \quad \{r, r+1, r+2, \ldots, 2r-1\}.$$

For sufficiently large n, Frankl and Füredi [7] showed that

$$\mathrm{ex}\Big(n, \sum\nolimits_3\Big) = \mathrm{ex}(n, T_3) = \left\lfloor \frac{n}{3} \right\rfloor \cdot \left\lfloor \frac{n+1}{3} \right\rfloor \cdot \left\lfloor \frac{n+2}{3} \right\rfloor,$$

with the extremal graph being the balanced 3-partite 3-graph on n vertices. In [8], Frankl and Füredi determined the exact value of $\mathrm{ex}(n, \sum_5)$ for all n divisible by 11 and the exact value of

$\mathrm{ex}(n,\Sigma_6)$ for all n divisible by 12. For these n, the extremal graphs are blow-ups of the unique $(11,5,4)$ and $(12,6,5)$ Steiner systems. Frankl and Füredi [8] conjectured that for all $r \geqslant 4$, if $n \geqslant n_0(r)$ is sufficiently large then $\mathrm{ex}(n,\Sigma_r) = \mathrm{ex}(n,T_r)$. Pikhurko [20] proved the conjecture for $r = 4$, showing that

$$\mathrm{ex}\left(n,\sum\nolimits_4\right) = \mathrm{ex}(n,T_4) = \left\lfloor \frac{n}{4} \right\rfloor \cdot \left\lfloor \frac{n+1}{4} \right\rfloor \cdot \left\lfloor \frac{n+2}{4} \right\rfloor \cdot \left\lfloor \frac{n+3}{4} \right\rfloor,$$

with the balanced complete 4-partite 4-graph on n vertices being the unique extremal graph. Recently, Norin and Yepremyan [18] proved Frankl and Füredi's conjecture for $r = 5$ and $r = 6$.

2.3. Expanded cliques and generalized fans

Given a hypergraph H and a pair $\{x,y\}$ of vertices in H, we say that $\{x,y\}$ is *covered* in H if some edge in H contains both x and y. Let $T_r(n,\ell)$ denote the complete ℓ-partite r-graph on n vertices where no two parts differ by more than one in size. Mubayi [15] considered the Turán problem for the following family of r-graphs. For all $p \geqslant r \geqslant 2$, let $\mathcal{K}_p^r$ denote the family of r-graphs H that contains a set C of p vertices such that every pair in C is covered in H. Let H_p^r denote the unique member of $\mathcal{K}_p^r$ with edge set

$$\left\{\{i,j\} \cup B_{i,j} : \{i,j\} \in \binom{[p]}{2}\right\},$$

where the $B_{i,j}$ are pairwise disjoint $(r-2)$-sets outside $[p]$. We call H_p^r the *r-uniform expanded p-clique*. For all n,p,r, Mubayi [15] showed that $\mathrm{ex}(n,\mathcal{K}_p^r) = |T_r(n,p-1)|$ with the unique extremal graph being $T_r(n,p-1)$. Mubayi further established structural stability of near-extremal $\mathcal{K}_p^r$-free graphs. Using this stability property, Pikhurko [21] later strengthened Mubayi's result to show that $\mathrm{ex}(n,H_p^r) = |T_r(n,p-1)|$ for all sufficiently large n.

Mubayi and Pikhurko [16] considered the Turán problem for so-called generalized fans. Let Fan^r be the r-graph comprising $r+1$ edges $e_1,\dots,e_r,e$ such that $e_i \cap e_j = \{x\}$ for all $i \neq j$, where $x \notin e$, and $|e_i \cap e| = 1$ for all i. Note that Fan^2 is precisely a triangle. Mubayi and Pikhurko showed that for all $r \geqslant 3$ and all sufficiently large n,

$$\mathrm{ex}(n,Fan^r) = |T_r(n,r)| = \prod_{i=1}^{r} \left\lfloor \frac{n+i-1}{r} \right\rfloor.$$

3. The general problem on $\mathcal{K}_p^F$ and H_p^F

The problems mentioned in the previous section can be generalized as follows, as discussed in Keevash [13]. Let $r \geqslant 3$. Let F be an r-graph. Let $p \geqslant n(F)$. Let $\mathcal{K}_p^F$ denote the family of r-graphs H that contains a set C of p vertices, called the *core*, such that the subgraph of H induced by C contains a copy of F and such that every pair in C is covered in H. Let H_p^F be the member of $\mathcal{K}_p^F$ obtained as follows. We label the vertices of F as $v_1,\dots,v_{n(F)}$. Add new vertices $v_{n(F)+1},\dots,v_p$. Let $C = \{v_1,\dots,v_p\}$. For each pair of vertices $v_i,v_j \in C$ not covered in F, we add a set $B_{i,j}$ of $r-2$ new vertices and the edge $\{v_i,v_j\} \cup B_{i,j}$, where the $B_{i,j}$ are pairwise disjoint over all such pairs $\{i,j\}$. We call H_p^F an *expanded p-clique with an embedded F*. We call C the *core* of H_p^F.

Using this notation, we can describe the families of graphs considered in the last section as follows. Let L denote the r-graph on $r+1$ vertices consisting of two edges sharing $r-1$ vertices.

Then $\mathcal{K}_{r+1}^L = \Sigma_r$ and $H_{r+1}^L = T_r$, the generalized triangle. If F is the r-uniform empty graph then $\mathcal{K}_p^F = \mathcal{K}_p^r$ and $H_p^F = H_p^r$, the r-uniform expanded p-clique. Let e denote a single r-set, then $H_{r+1}^e = Fan^r$, the r-uniform generalized fan. More recently, Hefetz and Keevash [10] studied $\mathrm{ex}(n, H_6^{M_2})$ and determined its exact value for all large n (together with stability), where M_2 consists of two disjoint triples.

Our main results in this paper show that for a large family of F, $\mathrm{ex}(n, \mathcal{K}_p^F) = \mathrm{ex}(n, H_p^F) = |T_r(n, p-1)|$ for all sufficiently large n, and that all near-extremal n-vertex H_p^F-free r-graphs G must be structurally close to $T_r(n, p-1)$. In fact, we establish this stability of near-extremal graphs first and then use stability to obtain the exact value of $\mathrm{ex}(n, H_p^F)$. See Theorems 6.2 and 6.3 for a detailed description of our main results. Let us point out that for this family of r-graphs F, $\pi(n, H_p^F)$ can already be determined using the hypergraph Lagrangian notion (see Theorem 5.7, which was given in [13]). However, the main challenge is to establish stability for near-extremal graphs and to apply such stability to establish the exact value of $\mathrm{ex}(n, H_p^F)$ for the single graph H_p^F for all sufficiently large n. This is also a feature of some earlier works such as those on expanded cliques and generalized fans, where determining $\pi(H_p^F)$ is quite straightforward and most of the work goes into establishing stability and determining the exact value of $\mathrm{ex}(n, H_p^F)$.

From our general theorems, some earlier results follow as corollaries. The contribution of the two main theorems also lies in the fact that, under suitable assumptions about F, they reduce the determination of the exact value of $\mathrm{ex}(n, H_p^F)$ for large n, together with structural stability of near-extremal graphs, to determining (or just bounding) the Lagrangian density $\pi_\lambda(F)$. Our main method is the stability method used by Pikhurko [20], which was influenced by some earlier work by Sidorenko [24].

4. Notations and definitions

Before introducing our main results, we give some notations and definitions that will be used throughout the paper. Given a hypergraph G and a set S of vertices, the *link graph* of S in G, denoted by $\mathcal{L}_G(S)$, is the hypergraph with edge set $\{f : f \subseteq V(G) \setminus S, f \cup S \in G\}$. We write $\mathcal{L}_G(u)$ for $\mathcal{L}_G(\{u\})$. The *degree* of S in G, denoted by $d_G(S)$, is the number of edges of G that contain S, that is, $d_G(S) = |\mathcal{L}_G(S)|$. We denote the minimum vertex degree of G by $\delta(G)$ and the number of vertices of G by $n(G)$.

Let $p \geqslant 1$ be an integer. The *p-shadow* of G, denoted by $\partial_p(G)$, is the set of p-sets that are contained in edges of G, that is, $\partial_p(G) = \{f : |f| = p, f \subseteq e \text{ for some } e \in G\}$. Let $m \geqslant r \geqslant 1$ be positive integers. Let $[m]_r$ denote the falling factorial $m(m-1)\cdots m(m-r+1)$.

A hypergraph *G covers pairs* if every pair of its vertices is contained in some edge. If G is a hypergraph and S is a set of vertices in G, then $G[S]$ denotes the subgraph of G induced by S.

5. Hypergraph Lagrangians and Lagrangian density

In order to describe our results, we need the notion of Lagrangians for hypergraphs. To motivate the notion of hypergraph Lagrangians, we first review the usual hypergraph symmetrization process and some of its properties. Two vertices u, v in a hypergraph H are *non-adjacent* if $\{u, v\}$ is not covered in H. Given a hypergraph H and two non-adjacent vertices u and v in H, *symmetrizing v to u* is the operation that removes all the edges of H containing v and replaces them with $\{v \cup D : D \in \mathcal{L}_H(u)\}$. In other words, we make v a clone of u. The following property is implicit in [13]. We re-establish it for completeness.

Proposition 5.1. *Let p, r be positive integers, where $p \geqslant r+1$. Let F be an r-graph with $n(F) \leqslant p$ and let G be an r-graph that is $\mathcal{K}_p^F$-free. Let u, v be two non-adjacent vertices in G. Let G' be obtained from G by symmetrizing v to u. Then G' is also $\mathcal{K}_p^F$-free.*

Proof. First note that u, v have codegree 0 in G'. Suppose for contradiction that G' contains a member H of $\mathcal{K}_p^F$ with C being its core. Since u, v have codegree 0 in G' and every pair in C is covered in $H \subseteq G'$, C contains at most one of u and v. For each $e \in H$, if $v \notin e$ let $f(e) = e$ and if $v \in e$ let $f(e) = (e \setminus \{v\}) \cup \{u\}$. Let $L = \{f(e) : e \in H\}$. Then $L \subseteq G$ and L is a member of $\mathcal{K}_p^F$ with either C (if $v \notin C$) or $(C \setminus \{v\}) \cup \{u\}$ (if $v \in C$) being the core. This contradicts G being $\mathcal{K}_p^F$-free. □

Given an r-graph G and two non-adjacent vertices u, v, if $\mathcal{L}_G(u) = \mathcal{L}_G(v)$ then we say that u and v are *equivalent*, and write $u \sim v$. Otherwise we say that u, v are *non-equivalent*. Note that $\sim$ is an equivalence relation on $V(G)$. The *equivalence class* of a vertex v consists of all the vertices that are equivalent to v.

Algorithm 5.2 (symmetrization without cleaning). Let G be an r-graph. We perform the following as long as G contains two non-adjacent non-equivalent vertices. Let u, v be two such vertices where $d(u) \geqslant d(v)$; we symmetrize each vertex in the equivalence class of v to u. We terminate the process when there exist no more non-adjacent non-equivalent pairs.

Note that the algorithm always terminates since the number of equivalence classes strictly decreases after each step that can be performed.

As usual, if $V_1, \ldots, V_s$ are disjoint sets of vertices then

$$\Pi_{i=1}^s V_i = V_1 \times V_2 \times \cdots \times V_s = \{(x_1, x_2, \ldots, x_s) : x_i \in V_i \text{ for all } i = 1, \ldots, s\}.$$

We will abuse notation and use $\Pi_{i=1}^s V_i$ also to denote the set of the corresponding unordered s-sets. If L is a hypergraph on $[m]$, then a *blowup* of L is a hypergraph G whose vertex set can be partitioned into $V_1, \ldots, V_m$ such that

$$E(G) = \bigcup_{e \in E(L)} \prod_{i \in e} V_i.$$

The next proposition follows immediately from the algorithm.

Proposition 5.3. *Let G be an r-graph and let G^* be the graph obtained at the end of the symmetrization process applied to G. Then we have the following.*

(i) $|G| \leqslant |G^*|$.
(ii) *Let S consist of one vertex from each equivalence class of G^* under $\sim$. Then $G^*[S]$ covers pairs and G^* is a blowup of $G^*[S]$.*

Let G be an r-graph on $[n]$. A *weight function*, or *weight assignment*, f on G is a mapping from $V(G)$ to $[0,\infty)$. We say that f is a 1-*sum weight assignment* if $\sum_{v\in V(G)} f(v) = 1$. For every edge e in G, define $f(e) = \prod_{v\in e} f(v)$ and call it the *weight* of e. We may describe f using the vector $\tilde{x} = (x_1,\dots,x_n)$, where $x_i = f(i)$ for each $i \in [n]$. Now, define a polynomial in the variables $\tilde{x} = (x_1,\dots,x_n)$ by

$$p_G(\tilde{x}) = p_G(x_1,\dots,x_n) = r!\cdot \sum_{e\in E(G)} \prod_{i\in e} x_i.$$

We define the *Lagrangian* of G to be

$$\lambda(G) = \max\left\{ p_G(x_1,\dots,x_n) : x_i \geqslant 0 \text{ for all } i = 1,\dots,n,\ \sum_{i=1}^{n} x_i = 1 \right\}.$$

A 1-sum weight assignment $\tilde{x}$ on G with $p_G(\tilde{x}) = \lambda(G)$ is called an *optimal weight assignment* on G. Given an r-graph F, we define the *Lagrangian density* $\pi_\lambda(F)$ of F to be

$$\pi_\lambda(F) = \sup\{\lambda(G) : F \not\subseteq G\}. \tag{5.1}$$

Note that our definition of the Lagrangian follows that of Sidorenko [24], and differs from the definition given by Keevash [13] by a factor of $r!$. The following proposition follows immediately from the definition of $\pi_\lambda(F)$.

Proposition 5.4. *Let F be an r-graph. Let L be an F-free r-graph. Let G be an r-graph on $[n]$ that is a blowup of L. Then $|G| \leqslant \pi_\lambda(F)n^r/r!$.*

Proof. Suppose $V(L) = [s]$ and let $V_1,\dots,V_s$ be the partition of $V(G)$ with V_i corresponding to i. For each $i \in [s]$, let $x_i = |V_i|/n$. Let $\tilde{x} = (x_1,\dots,x_n)$. Then $x_i \geqslant 0$ for all $i \in [s]$, and $\sum_{i=1}^{s} x_i = 1$. Since G is a blowup of L, we have

$$|G| = \sum_{e\in L}\prod_{i\in e}|V_i| = n^r \sum_{e\in L}\prod_{i\in e} x_i = \frac{n^r}{r!}\cdot p_L(\tilde{x}) \leqslant \frac{n^r}{r!}\pi_\lambda(F),$$

where the last inequality follows from the definition of $\pi_\lambda(F)$ and the fact that L is F-free. $\square$

Given r-graphs F and G we say $f : V(F) \to V(G)$ is a *homomorphism* if it preserves edges, that is, for every $e \in E(F)$ we have $f(e) \in E(G)$. We say that G is *F-hom-free* if there is no homomorphism from F to G. The following proposition is given in the first few paragraphs of [13, Section 3].

Proposition 5.5 ([13]). *If F,G are r-graphs and G is F-hom-free, then $\pi(F) \geqslant \lambda(G)$. Moreover, $\pi(F) = \sup\{\lambda(G) : G$ is F-hom-free$\}$.*

The next proposition is implicit in the proof of [13, Theorem 3.1].

Proposition 5.6 ([13]). *If F is an r-graph that covers pairs, then $\pi(F) = \pi_\lambda(F)$.*

Proof. Clearly, if G is an r-graph that is F-hom-free then it is also F-free. We claim that the converse is also true in this case. Let G be F-free. If there were a homomorphism f from F to

G, then the fact that every two vertices in F lie in an edge of F, and that f preserves edges, forces f to be injective, contradicting G being F-free. So, G is F-hom-free. Now, by (5.1) and Proposition 5.5,

$$\pi_\lambda(F) = \sup\{\lambda(G) : F \not\subseteq G\} = \sup\{\lambda(G) : G \text{ is } F\text{-hom-free}\} = \pi(F). \qquad \square$$

As mentioned in the Introduction, the notion of hypergraph Lagrangians already yields the following tight bounds on $\mathrm{ex}(n, \mathcal{K}^F_{m+1})$ for certain r-graphs F. We describe the bounds in the following theorem, which is a more specific version of Theorem 3.1 of [13]. We give a proof using our language.

Theorem 5.7 ([13]). *Let F be an r-graph with $n(F) \leqslant m+1$. Suppose that $\pi_\lambda(F) \leqslant [m]_r/m^r$. Then for every n we have*

$$\mathrm{ex}(n, \mathcal{K}^F_{m+1}) \leqslant \frac{[m]_r}{m^r} \cdot \frac{n^r}{r!}.$$

Equality holds if r divides n. In particular,

$$\pi(\mathcal{K}^F_{m+1}) = \frac{[m]_r}{m^r}.$$

Proof. If L is a member of $\mathcal{K}^F_{m+1}$ with core C, then $\partial_2(L)$ contains an $(m+1)$-clique since every pair in C is covered in L. Since $\partial_2(T^r_m(n))$ does not contain an $(m+1)$-clique, then $L \not\subseteq T^r_m(n)$. Hence $T^r_m(n)$ is $\mathcal{K}^F_{m+1}$-free and $\mathrm{ex}(n, \mathcal{K}^F_{m+1}) \geqslant e(T^r_m(n))$. Since

$$\lim_{n\to\infty} |T^r_m(n)| \Big/ \binom{n}{r} = \frac{[m]_r}{m^r},$$

we have

$$\pi(\mathcal{K}^F_{m+1}) \geqslant \frac{[m]_r}{m^r}.$$

Next, let G be a $\mathcal{K}^F_{m+1}$-free r-graph on $[n]$. Let G^* be the final graph obtained at the end of the symmetrization process applied to G. By Propositions 5.1 and 5.3, G^* is $\mathcal{K}^F_{m+1}$-free and $e(G^*) \geqslant e(G)$. Let S consist of one vertex from each equivalence class of G^*. By Proposition 5.3, $G^*[S]$ covers pairs and G^* is a blowup of $G^*[S]$. If $F \subseteq G^*[S]$, then since $G^*[S]$ covers pairs, G^* (in fact, $G^*[S]$) contains a member of $\mathcal{K}^F_{m+1}$, a contradiction. Hence $F \not\subseteq G^*[S]$.

By Lemma 5.4, we have

$$|G| \leqslant |G^*| \leqslant \pi_\lambda(F)\frac{n^r}{r!} \leqslant \frac{[m]_r}{m^r} \cdot \frac{n^r}{r!}.$$

Since this holds for every $\mathcal{K}^F_{m+1}$-free G on $[n]$, we have

$$\mathrm{ex}(n, \mathcal{K}^F_{m+1}) \leqslant \frac{[m]_r}{m^r} \cdot \frac{n^r}{r!}.$$

Note that when r divides n,

$$|T^r_m(n)| = \frac{[m]_r}{m^r} \cdot \frac{n^r}{r!}.$$

Hence

$$\mathrm{ex}(n, \mathcal{K}^F_{m+1}) = \frac{[m]_r}{m^r} \cdot \frac{n^r}{r!}$$

in this case. Finally, a straightforward calculation shows that

$$\pi(\mathcal{K}^F_{m+1}) = \lim_{n\to\infty} \mathrm{ex}(n, \mathcal{K}^F_{m+1}) \Big/ \binom{n}{r} \leqslant \frac{[m]_r}{r!}.$$

Hence

$$\pi(\mathcal{K}^F_{m+1}) = \frac{[m]_r}{m^r}. \qquad \square$$

6. Main results

6.1. Main theorems

As mentioned in the Introduction, our main results determine the exact value of $\mathrm{ex}(n, H^F_{m+1})$ for certain r-graphs F for sufficiently large n and establish stability of near-extremal H^F_{m+1}-free graphs.

Definition 6.1. Let $m, r \geqslant 2$ be positive integers. Let F be an r-graph on at most $m+1$ vertices with $\pi_\lambda(F) \leqslant [m]_r/m^r$. We say that $\mathcal{K}^F_{m+1}$ is *m-stable* if, for every real $\varepsilon > 0$, there are a real $\delta_1 > 0$ and an integer n_1 such that, if G is a $\mathcal{K}^F_{m+1}$-free r-graph with $n \geqslant n_1$ vertices and more than

$$\left(\frac{[m]_r}{m^r} - \delta_1\right)\binom{n}{r}$$

edges, then G can be made m-partite by deleting at most εn vertices.

Theorem 6.2 (stability). *Let m, r be positive integers. Let F be an r-graph that either has at most m vertices or has $m+1$ vertices one of which has degree 1. If $\pi_\lambda(F) < [m]_r/m^r$, then $\mathcal{K}^F_{m+1}$ is m-stable.*

Theorem 6.3 (stability to exactness). *Let F be an r-graph that either has at most m vertices or has $m+1$ vertices one of which has degree 1. If $\mathcal{K}^F_{m+1}$ is m-stable, then there exists an integer n_2 such that, for all $n \geqslant n_2$, $\mathrm{ex}(n, H^F_{m+1}) = |T_r(n,m)|$. Also, $T_r(n,m)$ is the unique extremal graph.*

Theorems 6.2 and 6.3 immediately imply our main theorem.

Theorem 6.4 (main theorem). *Let m, r be positive integers. Let F be an r-graph that either has at most m vertices or has $m+1$ vertices one of which has degree 1. Suppose either $\pi_\lambda(F) < [m]_r/m^r$ or $\pi_\lambda(F) = [m]_r/m^r$ and $\mathcal{K}^F_{m+1}$ is m-stable. Then there exists a positive integer n_3 such that for all $n \geqslant n_3$ we have $\mathrm{ex}(n, H^F_{m+1}) = |T_r(n,m)|$. Also, $T_r(n,m)$ is the unique extremal graph.*

By Proposition 5.6, we have the following corollary.

Corollary 6.5. *Let m,r be positive integers. Let F be an r-graph that either has at most m vertices or has $m+1$ vertices one of which has degree 1. Suppose F covers pairs. Suppose either $\pi(F) < [m]_r/m^r$ or $\pi(F) = [m]_r/m^r$ and $\mathcal{K}^F_{m+1}$ is m-stable. Then there exists a positive integer n_4 such that for all $n \geqslant n_4$ we have $\mathrm{ex}(n, H^F_{m+1}) = |T_r(n,m)|$. Also, $T_r(n,m)$ is the unique extremal graph.*

To introduce our next main theorem, we need a definition. Given a 2-graph G and an integer $r \geqslant 2$, the $(r-2)$*-fold enlargement* of G is an r-graph F obtained by taking an $(r-2)$-set D that is vertex-disjoint from G and letting $F = \{e \cup D : e \in G\}$.

Define the function

$$f_r(x) = \frac{\prod_{i=1}^{r-1}(x+i-2)}{(x+r-3)^r}.$$

Note that $f_r(x) > 0$ on $[0,\infty)$ and $\lim_{x\to\infty} f_r(x) = 0$. Let M_r denote the last (*i.e.* rightmost) maximum of the function f_r on the interval $[2,\infty)$. As pointed out in [24], M_r is non-decreasing in r, and can be specifically calculated. For instance, $M_2 = M_3 = 2$, $M_4 = 2+\sqrt{3}$. Also, we will define $M_1 = 2$. The well-known Erdős–Sós conjecture says that if T is a k-vertex tree or forest, then $\mathrm{ex}(n,T) \leqslant n(k-2)/2$. The conjecture has been verified for many families of trees. The conjecture has also been verified when k is large [1]. The following theorem was proved by Sidorenko [24].

Theorem 6.6 ([24]). *Let $r,k \geqslant 2$ be integers where $k \geqslant M_r$. Let T be a tree on k vertices that satisfies the Erdő–Sós conjecture. Let F be the $(r-2)$-fold enlargement of T. Then*

$$\pi(\mathcal{K}^F_{k+r-2}) = \pi_\lambda(F) = \frac{[k+r-3]_r}{(k+r-3)^r} = (k-2)f_r(k).$$

In fact, Sidorenko's arguments showed that

$$\mathrm{ex}(n, \mathcal{K}^F_{r+k-2}) \leqslant \frac{[k+r-3]_r}{(k+r-3)^r}\frac{n^r}{r!},$$

where equality is attained if $r+k-3$ divides n. However, no structural stability of near-extremal families was established and neither was the exact value of $\mathrm{ex}(n, H^F_{r+k-2})$ determined. Recall that H^F_{k+r-2} is a specific member of the family $\mathcal{K}^F_{k+r-2}$. We strengthen Sidorenko's result by establishing structural stability of near-extremal $\mathcal{K}^F_{k+r-2}$-free families and then using this stability to establish the exact value of $\mathrm{ex}(n, H^F_{k+r-2})$ for all sufficiently large n. The $k=2$ case is trivial. We henceforth assume $k \geqslant 3$.

Theorem 6.7 (stability of enlarged trees). *Let $k \geqslant 3, r \geqslant 2$ be integers, where $k \geqslant M_r$. Let T be a k-vertex tree that satisfies the Erdős–Sós conjecture. Let F be the $(r-2)$-fold enlargement of T. Then $\mathcal{K}^F_{k+r-2}$ is $(k+r-3)$-stable.*

Theorems 6.7 and 6.3 together imply the following theorem.

Theorem 6.8 (exact result on enlarged trees). *Let $k \geqslant 3, r \geqslant 2$ be integers, where $k \geqslant M_r$. Let T be a k-vertex tree that satisfies the Erdős–Sós conjecture. Let F be the $(r-2)$-fold enlargement of T. There exists a positive integer n_5 such that for all $n \geqslant n_5$ we have*

$$\mathrm{ex}(n, H^F_{r+k-2}) = |T_r(n, r+k-3)|.$$

Also, $T_r(n,m)$ is the unique extremal graph.

To show that $\mathcal{K}^F_{k+r-2}$ is $(k+r-3)$-stable, we first establish some useful properties of the Lagrangian function of a 2-graph not containing a given tree T. These properties (see Section 10) may be of independent interest.

For the rest of the paper, we prove Theorems 6.2, 6.3 and 6.7.

6.2. Applications of the main theorems and related remarks

We deduce a few earlier results using our main theorems, and give a few new examples.

Corollary 6.9 ([21]). *Let*

$$H^r_p = \left\{ \{i,j\} \cup B_{i,j} : \{i,j\} \in \binom{[p]}{2} \right\},$$

where $B_{i,j}$ are pairwise disjoint $(r-2)$-sets outside $[p]$. Then $\mathrm{ex}(n, H^r_p) = |T_r(n, p-1)|$ for all sufficiently large n.

Proof. Let F denote the r-uniform empty graph. Then

$$\pi_\lambda(F) = 0 < \frac{[p-1]_r}{(p-1)^r}.$$

By Theorem 6.4, $\mathrm{ex}(n, H^r_p) = \mathrm{ex}(n, H^F_p) = |T_r(n, p-1)|$ for all sufficiently large n. □

Corollary 6.10 ([16]). *Let $r \geqslant 3$. Let Fan^r consist of $r+1$ edges $e_1, \dots, e_r, e$ such that $e_i \cap e_j = \{x\}$ for all $i \neq j$, where $x \notin e$, and $|e_i \cap e| = 1$ for all i. Then $\mathrm{ex}(n, Fan^r) = |T_r(n,r)|$ for sufficiently large n.*

Proof. Note that $Fan^r = H^F_{r+1}$, where F consists of a single edge. Clearly, $\pi_\lambda(F) = 0 < [r]_r / r^r$. By Theorem 6.4, $\mathrm{ex}(n, Fan^r) = |T_r(n,r)|$ for all sufficiently large n. □

Corollary 6.11 ([7]). *Let $T_3 = \{\{1,2,3\},\{1,2,4\},\{3,4,5\}\}$. Then $\mathrm{ex}(n, T_3) = |T_3(n,3)|$ for all sufficiently large n.*

Proof. Let $F = \{\{1,2,3\},\{1,2,4\}\}$. Then F is the 1-fold enlargement of the 2-tree $K_{1,2}$, which is known to satisfy the Erdős–Sós conjecture. Let $k=3$. Recall that $M_3 = 2$ (see the discussion before Theorem 6.6). So $k \geqslant M_3$. Note that $T_3 = H^F_4$. By Theorem 6.8, $\mathrm{ex}(n, T_3) = |T_3(n,3)|$ for all sufficiently large n. □

Note, however, that Theorem 6.6 cannot be applied to T_r, for $r \geqslant 4$, where $T_r = \{\{1,2,\dots,r\}, \{2,3,\dots,r+1\}, \{r, r+1, r+2, \dots, 2r-1\}\}$, since $M_r > 3$ for $r \geqslant 4$. In fact, for $r = 5,6$ we saw in the Introduction that the extremal graph is no longer the Turán graph $T_r(n,r)$. Rather, the extremal graphs are blowups of certain designs.

Theorems 6.4 and 6.8 give rise to numerous new graphs whose Turán number is exactly determined, with the Turán graph being the unique extremal graph. Indeed, by Corollary 6.5, if F covers pairs, $n(F) \leqslant m$, and $\pi(F) < [m]_r/m^r$, then $\mathrm{ex}(n, H_{m+1}^F) = |T_r(n,m)|$ for all sufficiently large n. So we can construct as many such examples as we want. Using Theorem 6.8, we can construct such examples very easily. For instance, let

$$F = \{\{1,2,m+1\},\{2,3,m+1\},\ldots,\{m-1,m,m+1\}\},$$

where $m \geqslant 3$. Then F is the 1-fold enlargement of the path P_m. By Theorem 6.8, $\mathrm{ex}(n, H_{m+1}^F) = T_3(n,m)$ for all sufficiently large n.

In general, to apply Theorem 6.4, we need a bound on $\pi_\lambda(F)$, rather than $\pi(F)$. If $\pi_\lambda(F) < [m]_r/m^r$, then we get stability and the exact result right away. If $\pi_\lambda(F) = [m]_r/m^r$, then establishing stability (if it is applicable) can take more work. We also want to point out that determining $\pi_\lambda(F)$ (rather than just bounding it) is generally open and should provide a rich collection of problems for further study. We give a few recent examples for $\pi_\lambda(F)$. Let M_t^r denote the r-uniform matching of t edges (*i.e.* t disjoint edges). Hefetz and Keevash [10] showed that $\pi_\lambda(M_2^3) \leqslant \frac{12}{25}$ and that

$$\mathrm{ex}(n, H_6^{M_2^3}) = |T_3(n,5)|, \quad \text{for all sufficiently large } n.$$

Jiang, Peng and Wu [11] later gave a short new proof of this result. They also generalized the result to show that, for all $t \geqslant 2$,

$$\pi_\lambda(M_t^3) = \frac{[3t-1]_3}{[3t-1]^3},$$

and that

$$\lambda(G) < \frac{[3t-1]_3}{[3t-1]^3} - c$$

for some small positive real c for all M_t^3-free graph $G \neq K_{3t-1}^3$. Using this, one can show that $H_{3t}^{M_t^3}$ is $(3t-1)$-stable and hence, by Theorem 6.4,

$$\mathrm{ex}(n, H_{3t}^{M_t^3}) = |T_3(n,3t-1)|$$

for all sufficiently large n. Jiang, Peng and Wu [11] also determined $\pi_\lambda(F)$ for a few other r-graphs F, and applied Theorem 6.4 to obtain the exact value of $\mathrm{ex}(n, H_{m+1}^F)$ for all sufficiently large n.

The survey by Keevash [13] included a list of r-graphs whose Turán number is determined exactly (for large n). That list was short. Our results give a rather large infinite family of r-graphs of the form H_{m+1}^F whose extremal graph turns out to be $T_r(n,m)$. To the best of our knowledge, we are not aware of r-graphs not of the form H_{m+1}^F whose extremal graph is the Turán graph $T_r(n,m)$. It would be an interesting problem to characterize r-graphs H for which the extremal graph for the Turán number $\mathrm{ex}(n,H)$ is $T_r(n,m)$.

7. Reduction from H^F_{m+1}-free graphs to $\mathcal{K}^F_{m+1}$-free graphs

In this short section, we establish the fact that every H^F_{m+1}-free r-graph on $[n]$ can be made $\mathcal{K}^F_{m+1}$-free by removing $O(n^{r-1})$ edges. In particular, this implies that to establish stability of near-extremal H^F_{m+1}-free graphs it suffices to establish stability of near-extremal $\mathcal{K}^F_{m+1}$-free graphs.

We need the following result by Frankl on the Turán number of a matching. As is well known, for sufficiently large n, the Turán number $\mathrm{ex}(n, M_{s+1})$ of an r-uniform matching M_{s+1} of size $s+1$ is

$$\binom{n}{r} - \binom{n-s}{r},$$

as was shown by Erdős [4]. However, for our purposes we will use the following slightly weaker but simpler bound which applies to all n.

Lemma 7.1 ([6]). *If H is an r-graph on $[n]$ that contains no $(s+1)$-matching, then $|H| \leqslant s\binom{n}{r-1}$.*

In fact, Frankl [6] showed that if H is an r-graph that has no $(s+1)$-matching, then $|H| \leqslant s|\partial_{r-1}(H)|$. For an integer $s \geqslant 2$, an *s-sunflower* with kernel D is a collection of s distinct sets $A_1, \dots, A_s$ such that, for all $i, j \in [s]$, $i \neq j$, $A_i \cap A_j = D$. Given an r-graph G and a set D, define the *kernel degree* of D in G, denoted by $d^*_G(D)$, to be

$$d^*_G(D) = \max\{s : G \text{ contains an } s\text{-sunflower with kernel } D\}.$$

Lemma 7.2. *Given an r-graph G on $[n]$ and integers $p, d > 0$, where $d < r$, there exists a subgraph G' of G with*

$$|G'| \geqslant |G| - p\binom{n}{d}\binom{n}{r-d-1}$$

such that, for every d-set D in $[n]$, if $d_{G'}(D) > 0$ then $d^*_{G'}(D) > p$.

Proof. Starting with G, as long as there exists a d-set D of vertices such that the degree of D in the remaining graph is non-zero but is at most $p\binom{n}{r-d-1}$, we remove all the edges containing D. Let G' denote the final remaining subgraph of G. Then

$$|G'| \geqslant |G| - p\binom{n}{r-d-1}\binom{n}{d}.$$

It is possible that G' is empty. If G' is non-empty, then for every d-set D that has non-zero degree in $\mathcal{G}'$, we have

$$|\mathcal{L}_{G'}(D)| = d_{G'}(D) > p\binom{n}{r-d-1}.$$

Since $\mathcal{L}_{G'}(D)$ is an $(r-d)$-graph on $[n]$, by Lemma 7.1, it contains a $(p+1)$-matching. Hence, $d^*_{G'}(D) > p$. ☐

Lemma 7.3. *Let $p = n(H^F_{m+1})$. If G is an H^F_{m+1}-free graph on $[n]$, then G contains a $\mathcal{K}^F_{m+1}$-free subgraph G' with*

$$|G'| \geqslant |G| - p\binom{n}{r-3}\binom{n}{2}.$$

In particular, $\pi(H^F_{m+1}) = \pi(\mathcal{K}^F_{m+1})$.

Proof. Let G be the given H^F_{m+1}-free graph on $[n]$. By Lemma 7.2, G contains a subgraph G' with

$$|G'| \geqslant |G| - p\binom{n}{r-3}\binom{n}{2}$$

such that for every pair $\{a,b\}$ of vertices, if $d_{G'}(\{a,b\}) > 0$ then $d^*_{G'}(\{a,b\}) > p$. We show that G' is $\mathcal{K}^F_{m+1}$-free. Suppose for contradiction that G' contains a member H of $\mathcal{K}^F_{m+1}$. Let C denote the core of L. Then $H[C]$ contains a copy of F. Let $\{x,y\}$ be any pair in C that is uncovered by F. By definition, $\{x,y\}$ is covered by some edge of H and hence by some edge of G'. So $d_{G'}(\{x,y\}) \neq 0$ and thus $d^*_{G'}(\{x,y\}) > p$. So G' contains a $(p+1)$-sunflower $\mathcal{S}$ with kernel $\{x,y\}$. Since $p = n(H^F_{m+1}) \geqslant |C|$, we can find an edge e of $\mathcal{S}$ containing $\{x,y\}$ that intersects C only in $\{x,y\}$. We can continue the process and cover each uncovered pair $\{a,b\}$ in C using an edge that intersects the current partial copy H' of H^F_{m+1} only in a and b. We can do so since $\{a,b\}$ is the kernel of a $(p+1)$-sunflower and H' has at most p vertices. Thus we can find a copy of H^F_{m+1} in G', and thus in G, contradicting our assumption that G is H^F_{m+1}-free. Hence G' is $\mathcal{K}^F_{m+1}$-free and

$$|G| \leqslant \mathrm{ex}(n, \mathcal{K}^F_{m+1}) + p\binom{n}{r-3}\binom{n}{2}.$$

Since

$$\mathrm{ex}(n, \mathcal{K}^F_{m+1}) \leqslant \mathrm{ex}(n, H^F_{m+1}) \leqslant \mathrm{ex}(n, \mathcal{K}^F_{m+1}) + p\binom{n}{r-3}\binom{n}{2},$$

we have $\pi(H^F_{m+1}) = \pi(\mathcal{K}^F_{m+1})$. □

8. Stability of near-extremal families and proof of Theorem 6.2

We use Pikhurko's approach [20] to establish stability of near-extremal families. First we describe a modified symmetrization algorithm used in [20] (and in [10], [18]) that is key to the approach. Compared to the usual symmetrization algorithm (Algorithm 5.2), the modified algorithm has an extra cleaning stage in each iteration. In the algorithm, at any stage, when we discuss the equivalence class of a vertex, this refers to the equivalence class under $\sim$ that we defined earlier. We always automatically readjust equivalence classes after we apply an operation to a graph. Given an r-graph L and a real α with $0 < \alpha \leqslant 1$, we say that L is *α-dense* if L has minimum degree at least $\alpha\binom{n(L)-1}{r-1}$.

Algorithm 8.1 (symmetrization and cleaning with threshold α).
Input: An r-graph G.
Output: An r-graph G^*.

Initiation: Let $G_0 = H_0 = G$. Set $i = 0$.
Iteration: For each vertex u in H_i, let $A_i(u)$ denote the equivalence class that u is in. If either H_i is empty or H_i contains no two non-adjacent non-equivalent vertices, then let $G^* = H_i$ and terminate. Otherwise, let u, v be two non-adjacent non-equivalent vertices in H_i, where $d_{H_i}(u) \geqslant d_{H_i}(v)$. We symmetrize each vertex in $A_i(v)$ to u. Let G_{i+1} denote the resulting graph. Note that after the symmetrization, the equivalence classes may change in G_{i+1}. But they are still well-defined. If G_{i+1} has minimum degree at least $\alpha\binom{n(G_{i+1})-1}{r-1}$, that is, if G_{i+1} is α-dense, then let $H_{i+1} = G_{i+1}$. Otherwise we let $L = G_{i+1}$ and repeat the following: let z be any vertex of minimum degree in L. We redefine $L = L - z$ unless in forming G_{i+1} from H_i we symmetrized the equivalence class of some vertex v in H_i to some vertex in the equivalence class of z in H_i. In that case, we redefine $L = L - v$ instead. We repeat the process until L becomes either α-dense or empty. Let $H_{i+1} = L$. We call the process of forming H_{i+1} from G_{i+1} 'cleaning'. Let Z_{i+1} denote the set of vertices removed, so that $H_{i+1} = G_{i+1} - Z_{i+1}$. By our definition, if H_{i+1} is non-empty then it is α-dense.

Let us now give an overview of how the approach roughly works. The statement we establish will be more general than that of Theorem 6.2. Let F be an r-graph with $\pi_\lambda(F) \leqslant [m]_r/m^r$ such that either $n(F) \leqslant m$ or $n(F) = m+1$ and F contains a degree 1 vertex. We take an n-vertex $\mathcal{K}^F_{m+1}$-free graph G with $|G| \sim |T_r(n,m)|$ and wish to show that under certain conditions G can be made m-partite by deleting $o(n)$ vertices. We choose an appropriately small real $\gamma > 0$ and apply Algorithm 8.1 with threshold $[m]_r/m^r - \gamma$ to G to obtain G^*. The fact that G is $\mathcal{K}^F_{m+1}$-free and that $|G| \sim |T_r(n,m)|$ easily guarantee that $n(G^*) = n - o(n)$. The condition we need now is that G^* has a set W of size $n - o(n)$ such that $G^*[W]$ is m-partite (if we allow an empty part in an m-partition). This is readily guaranteed if $\pi_\lambda(F) < [m]_r/m^r$. If $\pi_\lambda(F) = [m]_r/m^r$ then we will need the existence of W as a given condition in our statement. (Later, we will see that if F is the $(r-2)$-fold enlargement of a tree T that satisfies the Erdős–Sós conjecture, we can guarantee the existence of W even when $\pi_\lambda(F) = [m]_r/m^r$.)

The key argument we want to make now is that if $G_0 = G, G_1, G_2, \ldots, G_s = G^*$ denotes the sequence of the graphs we obtain in the execution of Algorithm 8.1, then in fact for each $i = 0, 1 \ldots, s$, $G_i[W]$ is m-partite. In particular, $G[W]$ is m-partite. Since $|W| = n - o(n)$, this means that G itself can be made m-partite by deleting $o(n)$ vertices.

To show that for each $i = 0, 1, \ldots, s$, $G_i[W]$ is m-partite, we use reverse induction on i. By our earlier discussion, $G_s[W] = G^*[W]$ is m-partite. This forms the basis step. Now assume that $G_{i+1}[W]$ is m-partite; we wish to show that $G_i[W]$ is also m-partite. To accomplish this suppose that $(A^{i+1}_1, \ldots, A^{i+1}_m)$ is an m-partition of $G_{i+1}[W]$, and suppose that in forming G_{i+1} from G_i we symmetrize the equivalence class C_v of some vertex v to the equivalence class C_u of some vertex u and perform the cleaning afterwards with the given threshold. After excluding some peripheral cases, we may assume that both u, v are in W and without loss of generality that $u, v \in A^{i+1}_1$. Let $W' = W \setminus C_v$. Note that since $G_i[W'] = G_{i+1}[W']$, $G_i[W']$ is m-partite with an m-partition $(A^{i+1}_1 \setminus C_v, A^{i+1}_2, \ldots, A^{i+1}_m)$. For convenience, let $U_1 = A^{i+1}_1 \setminus C_v, U_2 = A^{i+1}_2, \ldots, U_m = A^{i+1}_m$. Let E_v be the set of edges in G_i that contain a vertex in C_v. Now, our main goal is to show that there exists $j \in [m]$, such that for each $e \in E_v$, e intersects U_ℓ in at most one vertex for each $\ell \in [m]$ and that $e \cap U_j = \emptyset$. Clearly, by the definition of C_v, $|e \cap C_v| = 1$. Now, we can extend the m-partition $(U_1, \ldots, U_m)$ of $G_i[W']$ to an m-partition of $G_i[W]$ by replacing U_j with $U_j \cup (C_v \cap W)$.

To establish the existence of such an index $j \in [m]$, we will make full use of the fact that the condition on the density of G_{i+1} forces $G_{i+1}[W']$ to be an almost complete m-partite r-graph. In such an almost complete environment, any edge not in $G_{i+1}[W']$ that contains two vertices from the same U_ℓ would force the occurrence of a member of $\mathcal{K}^F_{m+1}$, a contradiction. This shows that each $e \in E_v$ intersects each U_ℓ in at most one vertex. Since $m \geqslant r$ and e intersects E_v, e misses some U_ℓ. In essence, we will let U_j be the part of $G_{i+1}[W']$ that the largest number of edges in E_v miss. But the actual choice of U_j will be slightly more technical than given here. Once U_j is chosen, we will make use of the almost-completeness of $G_{i+1}[W']$ to show that in fact all edges in E_v must miss U_j.

Now we present the details. First, in Section 8.1, we develop a series of lemmas that essentially say that $G_i[W]$ is an almost complete m-partite r-graph with almost equal parts for each i. After the lemmas we will present our main stability theorem, Theorem 8.7, in Section 8.2.

8.1. Lemmas on the structure of the graphs obtained in Algorithm 8.1

First let us mention a routine fact, which is established in [15] and can be verified straightforwardly.

Lemma 8.2 (Claim 1 in [15]). *For any integers $m \geqslant r \geqslant 2$ and real $\gamma > 0$, there exist a real $\beta = \beta(\varepsilon) > 0$ and an integer M_1 such that, for any m-partite r-graph G of order $n \geqslant M_1$ and size at least $([m]_r/m^r - \beta)\binom{n}{r}$, the number of vertices in each part is between $(n/m) - \varepsilon n$ and $(n/m) + \varepsilon n$.*

We may assume that ε is sufficiently small. First, we choose small positive reals

$$1 \gg c_2 \gg c_1 \gg \gamma_0 > 0,$$

and an integer n_1. Our first condition on n_1 is that $n_1 \geqslant M_1$, where M_1 is given in Lemma 8.2, and that n_1 satisfies (8.1) given below. Other conditions on n_1 will be stated implicitly throughout the proofs. We now describe the conditions on the constants as follows. First we choose c_1 to be sufficiently small and n_1 sufficiently large such that for all $N \geqslant n_1$ we have

$$\left(\frac{[m]_r}{m^r} - c_1\right)\binom{N-1}{r-1} \geqslant \left(\frac{[m]_r}{m^r} - 2c_1\right)\frac{N^{r-1}}{(r-1)!} = \binom{m-1}{r-1}\left(\frac{N}{m}\right)^{r-1} - 2c_1\frac{N^{r-1}}{(r-1)!}. \tag{8.1}$$

Next, subject to (8.1), we choose c_1, c_2 to be sufficiently small and n_1 sufficiently large such that, for $N \geqslant n_1$,

$$\left(\frac{[m]_r}{m^r} - c_1\right)\binom{N-1}{r-1} - \binom{m-2}{r-1}\left(\frac{N}{m} + c_2 N\right)^{r-1} > \frac{1}{2}\binom{m-2}{r-2}\left(\frac{N}{m}\right)^{r-1} > \frac{1}{2m^{r-1}}N^{r-1}. \tag{8.2}$$

Such choices exist by (8.1) and the fact that

$$\binom{m-1}{r-1} - \binom{m-2}{r-1} = \binom{m-2}{r-2}.$$

In addition, we can make our choice of c_1, c_2 solely dependent on m and r. Now, subject to (8.1) and (8.2), we choose c_1, c_2, γ_0 to satisfy

$$c_2 < \frac{1}{6m}\left(\frac{1}{10m}\right)^{m-1}, \quad c_2 < \frac{1}{(2m)^{mr}}, \quad c_1 < \min\left\{\frac{c_2}{8}, \beta\left(\frac{c_2}{2m}\right)\right\}, \quad \gamma_0 + 4\gamma_0(r-1) < c_1, \tag{8.3}$$

where the function β is defined as in Lemma 8.2. Note that all c_1, c_2, γ_0 can be defined to be solely dependent on m and r.

Now, let $\gamma < \gamma_0$ be given. Choose $\delta > 0$ to be sufficiently small that

$$\frac{\gamma - \delta}{\gamma + \delta} \geqslant 1 - \gamma. \tag{8.4}$$

Let $n_0 = 2n_1$. Let G be a $\mathcal{K}_{m+1}^F$-free graph on $[n]$, where $n \geqslant n_0$, such that

$$|G| > \left(\frac{[m]_r}{m^r} - \delta\right)\binom{n}{r}.$$

Let G^* be the final graph obtained by applying Algorithm 8.1 to G with threshold $[m]_r/m^r - \gamma$. Suppose the algorithm terminates after s steps. So, $G^* = G_s$.

Lemma 8.3. *Let $Z^* = \bigcup_{i=1}^s Z_i$, that is, Z^* is the set of vertices removed by Algorithm 8.1 with threshold $[m]_r/m^r - \gamma$. Then $|Z^*| < \gamma n$. Hence, $n(G^*) \geqslant (1-\gamma)n$ and G^* is $([m]_r/m^r - \gamma)$-dense.*

Proof. Let $p = |Z^*|$. Let $\alpha = [m]_r/m^r$. By the algorithm, when symmetrizing, the number of edges does not decrease. When deleting a vertex, the number of edges we lose is at most $(\alpha - \gamma)\binom{x-1}{r-1}$, where x is the number of vertices remaining in the graph before the deletion of that vertex. Hence

$$\begin{aligned} |G_s| &\geqslant |G| - (\alpha - \gamma)\sum_{i=1}^{p}\binom{n-i}{r-1} \\ &\geqslant (\alpha - \delta)\binom{n}{r} - (\alpha - \gamma)\left[\binom{n}{r} - \binom{n-p}{r}\right]. \end{aligned}$$

Since symmetrizing preserves $\mathcal{K}_{m+1}^F$-freeness and deletion of vertices certainly also does, G_s is $\mathcal{K}_{m+1}^F$-free. By Theorem 5.7,

$$|G_s| \leqslant \alpha\frac{(n-p)^r}{r!} < (\alpha + \delta)\binom{n-p}{r},$$

for sufficiently large n. Hence we have

$$(\alpha + \delta)\binom{n-p}{r} \geqslant (\alpha - \delta)\binom{n}{r} - (\alpha - \gamma)\left[\binom{n}{r} - \binom{n-p}{r}\right].$$

This yields

$$(\gamma + \delta)\binom{n-p}{r} \geqslant (\gamma - \delta)\binom{n}{r}.$$

Hence

$$\left(\frac{n-p}{n}\right)^r \geqslant \binom{n-p}{r} \Big/ \binom{n}{r} \geqslant \frac{\gamma-\delta}{\gamma+\delta} \geqslant 1-\gamma,$$

where the last inequality holds by (8.4). Hence

$$1-\frac{p}{n} \geqslant (1-\gamma)^{1/r} \geqslant 1-\gamma.$$

So $p \leqslant \gamma n$. Hence $n(G^*) \geqslant (1-\gamma)n$. Since the algorithm terminates with a non-empty G^*, G^* is $([m]_r/m^r - \gamma)$-dense. □

Suppose now that there exists $W \subseteq V(G^*)$ with $|W| \geqslant (1-\gamma_0)|V(G^*)|$ such that W is the union of at most m equivalence classes of G^*. Let $N = |W|$. Then

$$N \geqslant (1-\gamma_0)(1-\gamma)n \geqslant n - 2\gamma_0 n. \tag{8.5}$$

Since $n \geqslant n_0$, certainly $N \geqslant n/2 \geqslant n_1$.

Lemma 8.4. *For each $i \in [s]$, we have*

$$\delta(G_i[W]) = \delta(H_i[W]) \geqslant \left(\frac{[m]_r}{m^r} - c_1\right)\binom{N-1}{r-1}.$$

Hence, in particular,

$$|G_i[W]| = |H_i[W]| \geqslant \left(\frac{[m]_r}{m^r} - c_1\right)\binom{N}{r}.$$

Proof. Note that for every $i \in [s]$, $G_i[W] = H_i[W]$, since $H_i = G_i - Z_i$ and $Z_i \subseteq Z^* \subseteq [n] \setminus W$. For convenience, let $\alpha = [m]_r/m^r$. Let $i \in [s]$. By the algorithm, H_i is $(\alpha-\gamma)$-dense, that is,

$$\delta(H_i) \geqslant (\alpha-\gamma)\binom{n(H_i)-1}{r-1}.$$

For each vertex x in W, by (8.5), there are at most $2\gamma_0 n\binom{n(H_i)-2}{r-2}$ edges of H_i that contain x and a vertex outside W. For each $i \in [s]$, since $n(H_i) \geqslant |W| > (1-2\gamma_0)n$, we have

$$n \leqslant \frac{1}{1-2\gamma_0} n(H_i) \leqslant 2(n(H_i)-1).$$

Then

$$\begin{aligned}
\delta(H_i[W]) &\geqslant (\alpha-\gamma)\binom{n(H_i)-1}{r-1} - 2\gamma_0 n\binom{n(H_i)-2}{r-2} \\
&\geqslant (\alpha-\gamma_0)\binom{n(H_i)-1}{r-1} - 4\gamma_0(n(H_i)-1)\binom{n(H_i)-2}{r-2} \\
&= (\alpha-\gamma_0-4\gamma_0(r-1))\binom{n(H_i)-1}{r-1} \geqslant (\alpha-c_1)\binom{N-1}{r-1},
\end{aligned}$$

where the last inequality follows from (8.3). □

Next, we develop a routine but useful lemma on near-complete m-partite r-graphs. Given an m-partite r-graph L with parts $A_1,\dots,A_m$, a *transversal* is a set S of vertices consisting of one vertex from each part. The transversal S is *complete* if it induces a complete r-graph on S. A transversal that is not complete is called *non-complete*.

Lemma 8.5. *Let L be an m-partite r-graph on $N \geqslant n_0$ vertices, where*

$$\delta(L) \geqslant \left(\frac{[m]_r}{m^r} - c_1\right)\binom{N-1}{r-1}.$$

Let $A_1,\dots,A_m$ be an m-partition of L. Then

(i) *for each $j \in [m]$, $||A_j| - N/m| < c_2 N$,*

(ii) *the number of non-complete transversals is at most $c_2 N^m$,*

(iii) *the number of non-complete transversals containing any one vertex is at most $c_2 N^{m-1}$.*

Proof. Since

$$\delta(L) \geqslant \left(\frac{[m]_r}{m^r} - c_1\right)\binom{N-1}{r-1},$$

we have

$$|L| \geqslant \left(\frac{[m]_r}{m^r} - c_1\right)\binom{N}{r}.$$

Since L is m-partite on $N \geqslant M_1$ vertices and $c_1 < \beta(c_2/2m)$, by Lemma 8.2,

$$\left| |A_j| - \frac{N}{m} \right| < \left(\frac{c_2}{2m}\right) N < c_2 N, \quad \text{for all } j \in [m]. \tag{8.6}$$

Hence item (i) holds. Let K denote the complete m-partite r-graph with parts $A_1,\dots,A_m$. Then

$$|K| \leqslant |T_r(N,m)| \leqslant \left(\frac{[m]_r}{m^r} + c_1\right)\binom{N}{r}, \quad \text{for sufficiently large } N.$$

Since

$$|L| \geqslant \left(\frac{[m]_r}{m^r} - c_1\right)\binom{N}{r},$$

we have

$$|K \setminus L| < 2c_1 \binom{N}{r}.$$

Each non-complete transversal must contain a member of $K \setminus L$. On the other hand, for a fixed member of $K \setminus L$, there are at most $(\max_j |A_j|)^{m-r} \leqslant (2N/m)^{m-r}$ transversals that contain it. So the number of non-complete transversals is at most

$$2c_1 \binom{N}{r}\left(\frac{2N}{m}\right)^{m-r} < 2c_1 N^m \leqslant c_2 N^m.$$

This proves item (ii). It remains to prove item (iii). Let x be any vertex. Without loss of generality, suppose $x \in A_1$. Let L_x denote the link graph of x in L. Let K_x denote the complete $(m-1)$-partite

$(r-1)$-graph with parts $A_2, \ldots, A_m$. A complete $(m-1)$-partite $(r-1)$-graph K' with $\lfloor N/m \rfloor$ vertices in each part has at most

$$\frac{[m]_r}{m^r}\binom{N-1}{r-1}$$

edges. Since

$$\left| |A_j| - \frac{N}{m} \right| < \left(\frac{c_2}{2m}\right)N, \quad \text{for } j = 2, \ldots, m,$$

we can delete at most $(c_2/2)N$ vertices from K_x to obtain a subgraph of K'. Hence,

$$|K_x| \leqslant |K'| + \left(\frac{c_2}{2}\right)N^{r-1} < \frac{[m]_r}{m^r}\binom{N-1}{r-1} + \left(\frac{c_2}{2}\right)N^{r-1}.$$

Since

$$|L_x| \geqslant \left(\frac{[m]_r}{m^r} - c_1\right)\binom{N-1}{r-1},$$

we have

$$|K_x \setminus L_x| < \left(c_1 + \frac{c_2}{2}\right)N^{r-1} < \left(\frac{3c_2}{4}\right)N^{r-1}.$$

Let T denote the collection of non-complete transversals that contain x. Every member of T must contain either an edge $e \in K \setminus L$ where $x \notin e$ or an edge $\{x\} \cup f \in K \setminus L$ where $f \in K_x \setminus L_x$. The number of members of T of the former type is at most

$$|K \setminus L| \cdot \left(\max_j |A_j|\right)^{m-1-r} \leqslant 2c_1\binom{N}{r}\left(\frac{2N}{m}\right)^{m-1-r} < 2c_1 N^{m-1} \leqslant \left(\frac{c_2}{4}\right)N^{m-1}.$$

The number of members of T_2 of the latter type is at most

$$|K_x \setminus L_x| \cdot \left(\max_j |A_j|\right)^{m-r} \leqslant \left(\frac{3c_2}{4}\right)N^{r-1}\left(\frac{2N}{m}\right)^{m-r} < \left(\frac{3c_2}{4}\right)N^{m-1}.$$

Hence $|T| \leqslant c_2 N^{m-1}$. □

Let us also include a fact that will be used in the proof of Theorem 6.2.

Lemma 8.6. *Let F be an r-graph such that either $n(F) \leqslant m$ or $n(F) = m+1$ and F contains a degree 1 vertex. If L is an r-graph obtained from the complete r-graph K on $[m]$ by duplicating vertex 1 into $1'$ and adding an edge e covering $\{1, 1'\}$, then L contains a member of $\mathcal{K}^F_{m+1}$.*

Proof. Let $C = [m] \cup \{1'\}$. Whether $n(F) \leqslant m$ or $n(F) = m+1$ and F contains a vertex of degree 1, it is easy to see that $L[C]$ contains F and that all pairs in C are covered in L. So L contains a member of $\mathcal{K}^F_{m+1}$. □

8.2. General stability theorem and proof of Theorem 6.2

We now present a stability theorem that is more general than Theorem 6.2. This theorem will be useful for establishing stability for H^F_{m+1} even when $\pi_\lambda(F) = [m]_r/m^r$ (whereas in Theorem 6.2, we assume $\pi_\lambda(F) < [m]_r/m^r$). Since we want the theorem to be as widely applicable as possible, the statements are rather technical.

Theorem 8.7. *Let $m \geqslant r \geqslant 2$ be integers. Let F be an r-graph with $\pi_\lambda(F) \leqslant [m]_r/m^r$ such that either $n(F) \leqslant m$ or $n(F) = m+1$ and F contains a vertex of degree 1. There exists a real $\gamma_0 = \gamma_0(m,r) > 0$ such that, for every positive real $\gamma < \gamma_0$, there exist a real $\delta > 0$ and an integer n_0 such that the following is true for all $n \geqslant n_0$. Let G be a $\mathcal{K}^F_{m+1}$-free r-graph on $[n]$ with $|G| > ([m]_r/m^r - \delta)\binom{n}{r}$. Let G^* be the final graph produced by Algorithm 8.1 with threshold $[m]_r/m^r - \gamma$. Then $n(G^*) \geqslant (1-\gamma)n$ and G^* is $([m]_r/m^r - \gamma)$-dense. Further, if there is a set $W \subseteq V(G^*)$ with $|W| \geqslant (1-\gamma_0)|V(G^*)|$ such that W is the union of a collection of at most m equivalence classes of G^*, then $G[W]$ is m-partite.*

Proof. By Lemma 8.3, $n(G^*) \geqslant (1-\gamma)n$ and G^* is $([m]_r/m^r - \gamma)$-dense. By Lemma 5.1, symmetrizing preserves $\mathcal{K}^F_{m+1}$-freeness. Deletion of vertices certainly also does. So G^* is $\mathcal{K}^F_{m+1}$-free. Next, we want to prove that $G[W]$ is m-partite. To do that, we use reverse induction on i to prove that for every $i \in [s]$, $G_i[W]$ is m-partite. By our assumption, $G_s[W]$ is m-partite. This establishes the basis step. Let $i < s$. Assume that $G_{i+1}[W]$ is m-partite; we prove that $G_i[W]$ must also be m-partite. As before, let $N = |W|$. Let $A^{i+1}_1, \ldots, A^{i+1}_m$ be an m-partition of $G_{i+1}[W]$. By Lemma 8.4, we have

$$\delta(G_{i+1}[W]) = \delta(H_{i+1}[W]) \geqslant \left(\frac{[m]_r}{m^r} - c_1\right)\binom{N-1}{r-1} \quad \text{and}$$
$$|G_{i+1}[W]| = |H_{i+1}[W]| \geqslant \left(\frac{[m]_r}{m^r} - c_1\right)\binom{N}{r}.$$

By Lemma 8.5,

$$\left| |A^{i+1}_j| - \frac{N}{m} \right| < c_2 N, \quad \text{for all } j \in [m].$$

In particular, we may assume that $N/2m \leqslant |A^{i+1}_j| \leqslant 2N/m$ for all $j \in [m]$. Let K_{i+1} denote the complete m-partite graph on W with parts $A^{i+1}_1, \ldots, A^{i+1}_m$.

Suppose that in forming G_{i+1} from H_i we symmetrized the equivalence class C_v of v in H_i to some vertex u in H_i. If none of C_v is in W, then $G_i[W] = G_{i+1}[W]$ and there is nothing to prove. So we may assume that $C_v \cap W \neq \emptyset$. Since all the vertices in C_v are the same, we assume that $v \in C_v \cap W$. By our algorithm this means $u \in W$ as well. Indeed, by construction, since we symmetrized C_v to u, in the subsequent cleaning steps of our algorithm, u would be removed only if all of C_v were removed. Also, from step $i+1$ forward, u and v always lie in the same equivalence class. Since W is the union of equivalence classes of G_s and $v \in W$, we should have $u \in W$ as well. Without loss of generality, suppose $u \in A^{i+1}_1$. Let $U_1 = A^{i+1}_1 \setminus C_v$ and $W' = W \setminus C_v$. For each $j = 2, \ldots, m$, let $U_j = A^{i+1}_j$. Then $U_1, \ldots, U_m$ is an m-partition of $G_{i+1}[W']$ and also note that $G_{i+1}[W'] = H_i[W'] = G_i[W']$. Let E_v be the set of edges of $H_i[W]$ that contains v. Let $E'_v = \{e \setminus \{v\} : e \in E_v\}$. Then $|E'_v| = |E_v|$. By Lemma 8.4,

$$|E_v| = |E'_v| \geqslant \left(\frac{[m]_r}{m^r} - c_1\right)\binom{N-1}{r-1}. \tag{8.7}$$

Claim 1. For every $e \in E_v$, and for every $j \in [m]$, we have $|e \cap U_j| \leqslant 1$.

Proof of Claim 1. First we show that for every $e \in E_v$, $|e \cap U_1| \leqslant 1$. Suppose for contradiction that there exists $e \in E_v$ with $|e \cap U_1| \geqslant 2$. Let $a, b \in e \cap U_1$. Let $\mathcal{S}$ be the collection of all $(m-1)$-sets S obtained by selecting one vertex from U_ℓ for each $\ell \in [m] \setminus \{1\}$. Then

$$|\mathcal{S}| \geqslant \left(\frac{N}{2m}\right)^{m-1} > 2c_2 N^{m-1}, \tag{8.8}$$

where the last inequality follows from (8.3). For each $S \in \mathcal{S}$, note that $S \cup \{a\}$ and $S \cup \{b\}$ are both transversals in $G_{i+1}[W]$ (relative to $A_1^{i+1}, \ldots, A_m^{i+1}$). By Lemma 8.5 there are at most $2c_2 N^{m-1}$ non-complete transversals in $G_{i+1}[W]$ containing either a or b. Hence, by (8.8) there exists $S \in \mathcal{S}$ such that $S_1 = S \cup \{a\}$ and $S_2 = S \cup \{b\}$ are complete transversals in $G_{i+1}[W]$. That is, S_1 and S_2 both induce complete r-graphs in $G_{i+1}[W]$. Since $S_1, S_2 \subseteq W'$ and $G_i[W'] = H_i[W'] = G_{i+1}[W']$, S_1 and S_2 both induce complete r-graphs in $H_i[W]$ as well. By Lemma 8.6, the union of these two complete r-graphs plus e contains a member of $\mathcal{K}_{m+1}^F$ in $H_i[W]$, a contradiction. Hence, for all $e \in E_v$, $|e \cap U_1| \leqslant 1$.

Next, let $j \in [m] \setminus \{1\}$. Suppose there exists $e \in E_v$ such that $|e \cap U_j| \geqslant 2$. If $|U_1| \geqslant N/10m$, then we argue as above, the only difference being that we replace (8.8) with $|\mathcal{S}| \geqslant (N/10m)^{m-1} > 2c_2 N^{m-1}$, which still holds by (8.3). Hence, we may assume that $|U_1| < N/10m$.

Since $A_1^{i+1} = U_1 \cup (C_v \cap W)$ and $|A_1^{i+1}| \geqslant N/2m$, we have $|C_v \cap W| \geqslant 0.4(N/m)$. Let $a, b \in e \cap U_j$. Recall that $u \in U_1$. Suppose first that the number of non-complete transversals in $G_{i+1}[W]$ containing both u and a is at least $3mc_2 N^{m-2}$. Then, since all of $C_v \cap W$ is symmetrized to u in forming G_{i+1} from H_i and $C_v \cap W$ and u are both in A_1^{i+1}, the number of non-complete transversals in $G^{i+1}[W]$ that contain a is at least

$$3mc_2 N^{m-2} |C_v \cap W| \geqslant 3mc_2 \frac{0.4}{m} N^{m-1} > c_2 N^{m-1},$$

contradicting Lemma 8.5. Hence, the number of non-complete transversals containing both u and a is at most $3mc_2 N^{m-2}$. Similarly the number of non-complete transversals containing both u and b is at most $3mc_2 N^{m-2}$. Let $\mathcal{S}$ be the collection of $(m-2)$-sets S obtained by selecting one vertex from U_j for each $j \in [m] \setminus \{1, j\}$. Then

$$|\mathcal{S}| \geqslant \left(\frac{N}{2m}\right)^{m-2} > 6mc_2 N^{m-2}, \tag{8.9}$$

where the last inequality follows from (8.3). For each $S \in \mathcal{S}$, $S_1 = S \cup \{u, a\}$ is a transversal in $G_{i+1}[W]$ containing both u and a, and $S_2 = S \cup \{u, b\}$ is a transversal in $G_{i+1}[W]$ containing both u and b. By (8.9), there exists $S \in \mathcal{S}$ such that both S_1 and S_2 are complete transversals in $G_{i+1}[W]$. As before they both induce complete r-graphs in $H_i[W]$ as well. Their union together with e now contains a member of $\mathcal{K}_{m+1}^F$, a contradiction. Hence, for all $e \in E_v$, $j \in [m]$, $|e \cap U_j| \leqslant 1$. $\square$

By Claim 1, for all $f \in E'_v$ and for all $j \in [m]$, $|f \cap U_j| \leqslant 1$. So each member f of E'_v intersects some $r-1$ parts among $U_1, \ldots, U_m$. By an averaging argument, there exist some $r-1$ parts $U_{j_1}, \ldots, U_{j_{r-1}}$ such that at least $|E'_v| / \binom{m}{r-1}$ members of E'_v intersect these $r-1$ parts and no other parts. Let $J = \{j_1, \dots, j_{r-1}\}$. Let

$$E_J = \{f \in E'_v : \forall j \in J, f \cap U_j \neq \emptyset\}. \tag{8.10}$$

By our discussion,

$$|E_J| \geqslant |E'_v| \Big/ \binom{m}{r-1} \geqslant \left(\frac{[m]_r}{m^r} - c_1\right)\binom{N-1}{r-1} \Big/ \binom{m}{r-1} > \frac{1}{(2m)^r} N^{r-1}, \tag{8.11}$$

for sufficiently large $N \geqslant n_1$.

Let

$$I = \left\{ i \in [m] : |\partial_1(E'_v) \cap U_i| \geqslant \frac{1}{(2m)^r} N \right\}.$$

By (8.11) and the definition of I, we have $J \subseteq I$. First, suppose that $|I| \leqslant m-2$. By our earlier discussion, for all $i \in I \subseteq [m]$, $|U_i| \leqslant N/m + c_2 N$. Also, for each $i \notin I$, the number of members of E'_v intersecting U_i is trivially at most

$$\frac{1}{(2m)^r} N \cdot N^{r-2} = \frac{1}{(2m)^r} N^{r-1}.$$

Hence, by (8.2) (with room to spare),

$$|E'_v| \leqslant \binom{m-2}{r-1}\left(\frac{N}{m} + c_2 N\right)^{r-1} + m\left[\frac{1}{(2m)^r} N^{r-1}\right] < \left(\frac{[m]_r}{m^r} - c_1\right)\binom{N-1}{r-1},$$

contradicting (8.7). Hence

$$|I| \geqslant m-1.$$

If $|I| = m-1$, then let $k \in [m] \setminus I$. If $|I| = m$, then let $k \in I \setminus J$.

Claim 2. For all $e \in E_v$, we have $e \cap U_k = \emptyset$.

Proof of Claim 2. Suppose for contradiction that there exists $e \in E_v$ containing a vertex $y \in U_k$. Let $\mathcal{T}$ be the collection of $(m-r)$-sets T obtained by selecting one vertex from $\partial_1(E'_v) \cap U_\ell$ for each $\ell \in [m] \setminus (J \cup \{k\}) \subseteq I$. For each $T \in \mathcal{T}$ and $f \in E_J$, $T' = T \cup f \cup \{y\}$ is a transversal in $G_{i+1}[W]$ containing y. The number of different T' is at least

$$|E_J|\left[\frac{1}{(2m)^r} N\right]^{m-r} \geqslant \frac{N^{r-1}}{(2m)^r} \cdot \left[\frac{N}{(2m)^r}\right]^{m-r} = \frac{N^{m-1}}{(2m)^{r(m-r)+r}} > c_2 N^{m-1},$$

where the last inequality follows from (8.3). By Lemma 8.5, the number of non-complete transversals in $G_{i+1}[W]$ containing y is less than $c_2 N^{m-1}$. So there exist $T \in \mathcal{T}, f \in E_J$ such that $T' = T \cup f \cup \{y\}$ is a complete transversal in $G_{i+1}[W]$. As before, T' also induces a complete r-graph in $H_i[W]$. Now we can find a member of $\mathcal{K}^F_{m+1}$ in $H_i[W]$ as follows. Let $C' = \{v\} \cup T'$. If $n(F) \leqslant m$, then we map F into T'. If $n(F) = m+1$ and z is a degree 1 vertex in F, then we map F into C' with z mapped to v. Such mappings exist since $v \cup f \in H_i[W]$ and T' is complete in $H_i[W]$. It remains to check that all pairs in C' are covered in $H_i[W]$. Pairs not containing v are covered since $H_i[T']$ is complete. Pairs of the form $\{v, a\}$ where $a \in f$ are covered by $\{v\} \cup f$. The pair $\{v, y\}$ is covered by e. The remaining pairs have the form $\{v, b\}$, where $b \in T$. By our definition of T, $b \in \partial_1(E'_v)$. Hence there exist an edge in E_v that contains v and b. We have thus shown that $H_i[W]$ contains a member of $\mathcal{K}^F_{m+1}$. This contradicts $H_i[W]$ being $\mathcal{K}^F_{m+1}$-free. $\square$

We have thus far shown that each edge in E_v intersects each U_j in at most one vertex and intersects U_k in no vertex. Since all the vertices in $S_v \cap W$ behave the same as v in $H_i[W]$, we see that $H_i[W]$ is m-partite with an m-partition $U'_1, \dots, U'_m$, where $U'_j = U_j$ for each $j \in [m] \setminus \{k\}$ and $U'_k = U_k \cup (S_v \cap W)$. This completes the induction and the proof of Theorem 8.7. $\square$

Now we can prove Theorem 6.2; namely, we show that if F is an r-graph with $\pi_\lambda(F) < [m]_r/m^r$ such that either $n(F) \leqslant m$ or $n(F) = m+1$ and F contains a vertex of degree 1, then $\mathcal{K}^F_{m+1}$ is m-stable.

Proof of Theorem 6.2. Let $\varepsilon > 0$ be given. We may assume that ε is sufficiently small that $\varepsilon < \gamma_0$, where γ_0 is given in Theorem 8.7. Let $\beta = [m]_r/m^r - \pi_\lambda(F)$. Let $\gamma = \min\{\varepsilon, \beta/3r\}$. Let δ, n_0 be the constants guaranteed by Theorem 8.7 for the above-defined γ. Let $\delta_1 = \min\{\delta, \beta/3\}$. Let $n_1 \geqslant n_0$ be sufficiently large that for $n \geqslant n_1$ we have

$$\left(\frac{[m]_r}{m^r} - \delta_1 - \gamma r\right)\binom{n}{r} > \left(\frac{[m]_r}{m^r} - \frac{2\beta}{3}\right)\binom{n}{r} > \pi_\lambda(F)\frac{n^r}{r!}.$$

Let G be a $\mathcal{K}^F_{m+1}$-free graph of order $n \geqslant n_1$ and size more than $([m]_r/m^r - \delta_1)\binom{n}{r}$. Let G^* be the final graph produced by applying Algorithm 8.1 to G with threshold $[m]_r/m^r - \gamma$. By Theorem 8.7, $n(G^*) \geqslant (1-\gamma)n \geqslant (1-\varepsilon)n$. Since G^* is the final graph produced by Algorithm 8.1, if S consists of one vertex from each equivalence class of G^* then $G^*[S]$ covers pairs and G^* is a blowup of $G^*[S]$. If $|S| \leqslant m$, then $W = V(G^*)$ is the union of at most m equivalence classes of G^*. By Theorem 8.7, $G[W]$ is m-partite. So G can be made m-partite by deleting at most εn vertices and we are done.

We henceforth assume that $|S| \geqslant m+1$. If $F \subseteq G^*[S]$, then since $G^*[S]$ covers pairs we can find a member of $\mathcal{K}^F_{m+1}$ in $G^*[S]$ by using any $(m+1)$-set that contains a copy of F as the core, contradicting G^* being $\mathcal{K}^F_{m+1}$-free. Hence $G^*[S]$ is F-free. In producing G^* from G, observe that each time we symmetrize, the number of edges does not decrease. Since at most γn vertices are deleted in the process,

$$|G^*| > |G| - \gamma n\binom{n-1}{r-1} \geqslant \left(\frac{[m]_r}{m^r} - \delta_1 - \gamma r\right)\binom{n}{r} > \pi_\lambda(F)\frac{n^r}{r!},$$

contradicting Lemma 5.4. $\square$

9. Establishing exactness from stability

In this section we prove Theorem 6.3. Let F be an r-graph such that either $n(F) \leqslant m$ or $n(F) = m+1$ and F contains a vertex of degree 1. We prove that if $\mathcal{K}^F_{m+1}$ is m-stable then $\mathrm{ex}(n, H^F_{m+1}) = |T_r(n,m)|$ for sufficiently large n.

Proof of Theorem 6.3. First we define a few constants. Let

$$c_1 = \frac{1}{(2m)^{m^3}}, \quad c_2 = \frac{c_1}{2mr}, \quad c_3 = \frac{c_1}{r}, \quad c_4 = \frac{c_1}{2r\binom{m-1}{r-2}}, \quad c_5 = \left(\frac{c_4}{2m}\right)^{m^3}. \tag{9.1}$$

Let

$$\varepsilon = \min\left\{\frac{c_5}{2r}, \frac{c_1 c_2}{2r^2}\right\}. \tag{9.2}$$

Since $\mathcal{K}_{m+1}^F$ is m-stable, by Definition 6.1, there exist a real $\delta_1 > 0$ and a positive integer n_1 such that, for all $n \geqslant n_1$, if G is a $\mathcal{K}_{m+1}^F$-free r-graph on $[n]$ with $|G| > ([m]_r/m^r - \delta_1)\binom{n}{r}$ edges, then G can be made m-partite by deleting at most εn vertices. By further reducing δ_1 if needed, we may assume that $\delta_1 \leqslant \varepsilon$. Let n_2 be sufficiently large that $n_2 \geqslant n_1$ and that every $n \geqslant n_2$ satisfies various inequalities involving n that we will specify throughout the proof. Let G now be a maximum H_{m+1}^F-free graph on $[n]$ with $n \geqslant n_2$. Since $T_r(n,m)$ is H_{m+1}^F-free, we have

$$|G| \geqslant |T_r(n,m)|. \tag{9.3}$$

Let $p = n(\mathcal{K}_{m+1}^F)$. By Lemma 7.3, G contains a subgraph G' with

$$|G'| \geqslant |G| - p\binom{n}{r-3}\binom{n}{2}$$

such that G' is $\mathcal{K}_{m+1}^F$-free. For sufficiently large $n \geqslant n_2$ we have

$$|G'| > \left(\frac{[m]_r}{m^r} - \delta_1\right)\binom{n}{r}.$$

Since $\mathcal{K}_{m+1}^F$ is m-stable and $n(G') \geqslant n_1$, G' can be made m-partite by deleting at most εn vertices. Hence, in particular, G' contains an m-partite subgraph with at least

$$|G'| - \varepsilon n^r \geqslant |G| - p\binom{n}{r-3}\binom{n}{2} - \varepsilon n^r \geqslant |G| - 2\varepsilon n^r$$

edges (assuming that n_2 is sufficiently large). Among all m-partitions of $[n]$, let $V_1 \cup \ldots \cup V_m$ be an m-partition of $[n]$ that maximizes

$$\phi = \sum_{e \in G} |\{i \in [m] : e \cap V_i \neq \emptyset\}|. \tag{9.4}$$

Let K be the complete m-partite r-graph on $[n]$ with parts $V_1, \ldots, V_m$. By the definition of ϕ, we have $\phi \geqslant r|G \cap K|$. By the choice of K, we have $|G \cap K| \geqslant |G| - 2\varepsilon n^r$ and thus $\phi \geqslant r(|G| - 2\varepsilon n^r)$. On the other hand, $\phi \leqslant r|G| - |G \setminus K|$, since each edge of $G \setminus K$ contributes at most $r-1$ to ϕ. It follows that

$$|G \setminus K| \leqslant 2r\varepsilon n^r. \tag{9.5}$$

We call an edge e on $[n]$ *crossing* if it contains at most one vertex of each V_i, that is, if $e \in K$. Let

$$M = K \setminus G \quad \text{and} \quad B = G \setminus K.$$

We call edges in M *missing edges*. We call edges in B *bad edges*. Since $|G| \geqslant |T_r(n,m)| \geqslant |K|$, we have $|B| \geqslant |M|$. By (9.5), we then have

$$|M| \leqslant |B| \leqslant 2r\varepsilon n^r. \tag{9.6}$$

Our goal for the rest of the proof is to show that in fact $B = \emptyset$, from which we would have $|G| \leqslant |K| \leqslant |T_r(n,m)|$, which would complete our proof. For the rest of the proof, we suppose $B \neq \emptyset$ and will derive a contradiction.

First, note that

$$|K| \geqslant |K \cap G| = |G| - |G \setminus K| \geqslant |T_r(n,m)| - 2r\varepsilon n^r, \tag{9.7}$$

for sufficiently large n. For sufficiently large n, this implies

$$0.9\frac{n}{m} \leqslant |V_i| \leqslant 1.1\frac{n}{m}, \quad \text{for all } i \in [m]. \tag{9.8}$$

Let $q = \binom{m+1}{2} + r$. Let $K_m^r(q)$ denote the complete m-partite r-graph with q vertices in each part. Let $A_1, \dots, A_m$ denote the m parts.

Claim 3. If u, v are two vertices in some part of a copy L of $K_m^r(q)$ in G, then $d_G(\{u,v\}) = 0$.

Proof of Claim 3. Without loss of generality, suppose u, v lie in A_1. Suppose for contradiction that u, v lie in some edge e of G. Let C denote the core of H_{m+1}^F. By our assumption about F, there exists $z \in C$ such that z lies in 0 or 1 edge of F. Since $m+1 \geqslant r$, there exists $y \in C \setminus \{z\}$ such that $d_F(\{y,z\}) = 0$. We can obtain a copy H_{m+1}^F in G by mapping y, z to u, v, respectively, and the other vertices of C into $A_2, \dots, A_m$, one into each part. It remains to cover the pairs in C that are uncovered by F. The pair $\{x,y\}$ is covered by e. Since each part of L still has at least $\binom{m+1}{2}$ vertices outside e, it is easy to cover all such pairs so that the covering edges are pairwise disjoint outside C. This contradicts G being H_{m+1}^F-free. □

Claim 4. Let $e \in B$ and suppose $|e \cap V_i| \geqslant 2$. Let $u, v \in e \cap V_i$. Then either $d_M(u) \geqslant c_1 n^{r-1}$ or $d_M(v) \geqslant c_1 n^{r-1}$.

Proof of Claim 4. Without loss of generality, suppose $i = 1$. Let S be a set of mq vertices obtained by selecting u, v, $q-2$ vertices from $V_1 \setminus e$ and q vertices from each of $V_2 \setminus e, \dots, V_m \setminus e$. By Claim 3, $K[S] \not\subseteq G[S]$. Hence, for each such S, there exists $f \in K[S] \setminus G[S] \subseteq M$. There are at least $(n/2m)^{mq-2}$ choices of S. Suppose first that for at least half of the choices of S, the corresponding f is disjoint from $\{u,v\}$. For each fixed $f \in M$, there are at most n^{mq-2-r} choices of S for which $f \in K[S] \setminus G[S]$. Hence, using the definition of c_1 in (9.1) and the definition of ε in (9.2), we have

$$|M| \geqslant (1/2)(n/2m)^{mq-2}/n^{mq-2-r} \geqslant c_1 n^r > 2r\varepsilon n^r,$$

contradicting (9.6). Next, suppose for at least half of the choices of S, $K[S] \setminus G[S]$, f intersects $\{u,v\}$. Without loss of generality, suppose for at least $1/4$ of the choices of S, f contains u. Since each such f is contained in at most n^{mq-1-r} many S, there are at least

$$(1/4)(n/2m)^{mq-2}/n^{mq-1-r} \geqslant c_1 n^{r-1}$$

edges of M that contain u. $\square$

Let

$$W = \{w : d_M(w) \geqslant c_1 n^{r-1}\}.$$

By the definition of W and (9.6), we have

$$c_1 n^{r-1}|W| \leqslant \sum_{x\in[n]} d_M(x) = r|M| \leqslant 2r^2\varepsilon n^r.$$

Hence

$$|W| \leqslant \frac{2r^2\varepsilon}{c_1} n < c_2 n, \tag{9.9}$$

where the last inequality follows from (9.1) and (9.2). We call a pair $\{u,v\}$ of vertices in G a *bad pair* if there exist $e \in B$ and $i \in [m]$ such that $u,v \in e \cap V_i$. By Claim 4, each bad pair must contain an element of W. Hence, by (9.9), we have the following claim.

Claim 5. The number of bad pairs in G is at most $c_2 n^2$.

We call vertices in W *defect vertices*. By Claim 4, each $e \in B$ contains a defect vertex in a part V_i where $|e \cap V_i| \geqslant 2$. We pick such a defect vertex, denote it by $c(e)$ and call it the *centre* of e. (The choices for $c(e)$ may not be unique, but we will fix one.) For each defect vertex w, let $b(w)$ denote the number of edges $e \in B$ such that $c(e) = w$. By our discussion above and (9.6), we have

$$\sum_{w\in W} b(w) \geqslant |B| \geqslant |M| = \frac{1}{r}\sum_{x\in[n]} d_M(x) \geqslant \frac{1}{r}\sum_{w\in W} d_M(w).$$

Hence there exists $w_0 \in W$ such that

$$b(w_0) \geqslant \frac{1}{r} d_M(w_0) \geqslant \frac{c_1}{r} n^{r-1} = c_3 n^{r-1}. \tag{9.10}$$

Without loss of generality, may assume that $w_0 \in V_1$. Let

$$L = \{e \setminus w_0 : e \in B, c(e) = w_0\}.$$

By definition, any $e \in B$ with $c(e) = w_0$ contains at least one other vertex of V_1. Hence

$$f \cap V_1 \neq \emptyset, \quad \text{for all } f \in L. \tag{9.11}$$

By Claim 5, for each $i \in [m]$, the number of members of L that contain two or more vertices of V_i is at most $c_2 n^2 \cdot n^{r-3} = c_2 n^{r-1}$. Hence, using the definition of the constants given in (9.1), the number of members of L that contains at most one vertex from each V_i is at least

$$c_3 n^{r-1} - m c_2 n^{r-1} = \left(\frac{c_1}{r} - m\frac{c_1}{2mr}\right) n^{r-1} = \frac{c_1}{2r} n^{r-1} = c_4 \binom{m-1}{r-2} n^{r-1}.$$

Each such member of L contains a vertex of V_1 by (9.11). By the pigeonhole principle, for some collection of $r-2$ parts outside V_1, without loss of generality, say $V_2,\dots,V_{r-1}$, there are at least c_4n^{r-1} members of L that contain exactly one vertex of each of $V_1,V_2,\dots,V_{r-1}$. Let L' denote the collection of these members. Then L' is an $(r-1)$-partite $(r-1)$-graph with an $(r-1)$-partition $U_1,\dots,U_{r-1}$, where $U_i = V(L')\cap V_i$ for all $i\in[r-1]$, and

$$|L'| \geqslant c_4n^{r-1}. \tag{9.12}$$

Recall that $p = n(H^F_{m+1})$. Assuming n is sufficiently large, by Lemma 7.2, L' has a subgraph L'' with

$$|L''| \geqslant |L'| - pn\binom{n}{r-3} \geqslant \frac{1}{2}c_4n^{r-1} \tag{9.13}$$

such that for each vertex x with $d_{L''}(x) > 0$ we have $d^*_{L''}(x) > p$. Let us remove isolated vertices from L''. Then the condition implies that

$$d^*_G(\{w_0,x\}) \geqslant p+1, \quad \text{for all } x\in V(L''). \tag{9.14}$$

For each $i = r,\dots,m$, let

$$D_i = \left\{x\in V_i : d_G(\{x,w_0\}) > p\binom{n-3}{r-3}\right\}.$$

By the definition of D_i and Lemma 7.1, we have

$$d^*_G(\{w_0,x\}) \geqslant p+1, \quad \text{for all } i = r+1,\dots,m,\ x\in D_i. \tag{9.15}$$

Claim 6. For each $i = r,\dots,m$, $|D_i| \geqslant \frac{1}{2}c_4n$.

Proof of Claim 6. Suppose for contradiction that for some i, $|D_i| < \frac{1}{2}c_4n$. Without loss of generality, suppose that $|D_m| < \frac{1}{2}c_4n$. Let us consider a different m-partition of $V(G)$ by moving w_0 from V_1 to V_m. Let us consider the change to the value of ϕ, defined in (9.4). The only edges e of G whose contributions to ϕ are decreased by the move are those satisfying $e\cap V_1 = \{w_0\}$ and $e\cap V_m \neq \emptyset$. The decrease is 1 per edge. We can bound the number of such edges as follows. The number of edges of G containing w_0 and a vertex of D_m is at most

$$|D_m|\cdot n^{r-2} < \frac{1}{2}c_4n^{r-1}.$$

For each vertex x of $V_m\setminus D_m$ we have

$$d_G(\{x,w_0\}) < p\binom{n-1}{r-3}.$$

Hence the number of edges of G containing w_0 and a vertex of $V_m\setminus D_m$ is at most

$$n\cdot p\binom{n-1}{r-3} < \frac{1}{4}c_4n^{r-1}, \quad \text{for sufficiently large } n.$$

Hence there are fewer than

$$\frac{1}{2}c_4 n^{r-1} + \frac{1}{4}c_4 n^{r-1} < c_4 n^{r-1}$$

such edges.

On the other hand, the contribution to ϕ from each in $\{w_0 \cup f : f \in L'\}$ is increased by 1 by moving w_0 to V_m. By (9.12), $|L'| \geqslant c_4 n^{r-1}$. Hence $V_1 \backslash, V_2, \dots, V_{m-1}, V_{m+1} \cup w_0$ is a partition that has a higher ϕ-value than $V_1, \dots, V_m$, contradicting our choice of $V_1, \dots, V_m$. □

We are now ready to complete the proof of the theorem. Let $\mathcal{S}$ be the collection of all mq-sets S obtained by picking the vertex set of an edge f of L'', then picking $q-1$ vertices from $V_i \setminus f$, for each $i \in [r-1]$, and then picking q vertices from each of $D_r, \dots, D_m$ By (9.1), (9.8), (9.13) and Claim 6,

$$|\mathcal{S}| \geqslant \frac{1}{2}c_4 n^{r-1} \cdot \left(\frac{n}{2m}\right)^{(r-1)(q-1)} \cdot \left(\frac{1}{2}c_4 n\right)^{(m-r+1)q} \geqslant c_5 n^{mq}.$$

Claim 7. For each $S \in \mathcal{S}$, $K[S] \nsubseteq G[S]$.

Proof of Claim 7. Suppose for contradiction that $K[S] \subseteq G[S]$. Let C denote the core of H^F_{m+1}. By our assumption, C contains a vertex z that lies in 0 or 1 edge of F. Let f be a member of L'' contained in S. Let $S' \subseteq S$ contain $V(f)$ and one vertex from each of $V_r, \dots, V_m$. By our assumption, $G[S']$ is complete. Thus $G[S' \cup w_0]$ contains a copy F' of F, where w_0 plays the role of z and if z has degree 1 in F then $w_0 \cup f$ plays the role of the unique edge of F containing z. With $C = S' \cup \{w_0\}$, we can obtain a copy of H^F_{m+1} in G as follows. It suffices to cover the pairs $\{a,b\}$ in C that are uncovered by F' using edges that intersect C only in a,b and are pairwise disjoint outside C. Let $\{a,b\}$ be such pair. If $a,b \neq w_0$, then we can use an edge in $G[S] \supseteq K[S]$ to cover $\{a,b\}$ as in the proof of Claim 3. To cover a pair of the form $\{w_0, a\}$, we use (9.14) if $a \in V(f)$ or (9.15) if $a \in S' \setminus V(f)$. Hence $H^F_{m+1} \subseteq G$, a contradiction. □

Now, for each $S \in \mathcal{S}$, by Claim 7, $K[S]$ contains a member of $K \setminus G$. On the other hand, each member of $K \setminus G$ trivially is contained in at most n^{mq-r} different S. Hence,

$$|K \setminus G| \geqslant |\mathcal{S}|/n^{mq-r} \geqslant c_5 n^{mq}/n^{mq-r} = c_5 n^r \geqslant 2r\varepsilon n^r,$$

contradicting (9.6). The contradiction completes our proof of Theorem 6.3. □

10. Stability of expanded cliques with embedded enlarged trees

In this section we prove Theorem 6.7. The main work in proving Theorem 6.7 is to establish stability-type properties of Lagrangian functions of r-graphs that do not contain $(r-2)$-enlargement of a 2-tree. These properties may be of independent interest. Given an r-graph G on $[t]$ and variables $\tilde{x} = (x_1, \dots, x_t)$, recall that

$$p_G(\tilde{x}) = r! \cdot \sum_{e \in G} \prod_{i \in e} x_i,$$

and that $\lambda(G)$ is the maximum $p_G(\tilde{x})$ over all 1-sum weight assignments $\tilde{x}$. For each $i \in [t]$, let

$$\lambda_i = \frac{\partial(p_G(\tilde{x}))}{\partial(x_i)}.$$

Then it is straightforward to verify that

$$\lambda_i = r! \cdot \sum_{f \in \mathcal{L}_G(i)} \prod_{j \in f} x_j \quad \text{and} \quad p_G(\tilde{x}) = \frac{1}{r} \sum_{i=1}^{n} \lambda_i x_i. \tag{10.1}$$

By (10.1), we have

$$p_G(\tilde{x}) \leqslant \frac{1}{r} \max_i \lambda_i, \quad \text{and thus } \max_i \lambda_i \geqslant r \cdot p_G(\tilde{x}). \tag{10.2}$$

The following lemma will be useful for our analysis.

Lemma 10.1. *Let $\eta > 0$ be a real. Let G be an r-graph on $[t]$ and $\tilde{x} = (x_1, \dots, x_t)$ a 1-sum weight assignment on G with $p_G(x) \geqslant \lambda(G) - \eta$. Then there exists $i \in [t]$ such that $x_i \geqslant \max_{j \in [t]} x_j - 2r!\sqrt{\eta}$ and $\lambda_i \geqslant \max_{j \in [t]} \lambda_j - 2r!\sqrt{\eta}$.*

Proof. Suppose $x_a = \max_j x_j$ and $\lambda_b = \max_j \lambda_j$. If $a = b$ then the claim holds with $i = a = b$. So assume $a \neq b$. If $\lambda_a > \lambda_b - 2r!\sqrt{\eta}$, then the claim holds with $i = a$. Similarly, if $x_b > x_a - 2r!\sqrt{\eta}$, then the claim holds with $i = b$. So we may assume that $\lambda_a < \lambda_b - 2r!\sqrt{\eta}$ and $x_b < x_a - r!\sqrt{\eta}$. That is, $x_a - x_b > 2r!\sqrt{\eta}$ and $\lambda_b - \lambda_a > 2r!\sqrt{\eta}$.

Let

$$w_a = \sum_{e \in \mathcal{L}(a) \setminus \mathcal{L}(b)} \prod_{i \in e} x_i, \quad w_b = \sum_{e \in \mathcal{L}(b) \setminus \mathcal{L}(a)} \prod_{i \in e} x_i \quad \text{and} \quad w^* = \sum_{e \in \mathcal{L}(\{a,b\})} \prod_{i \in e} x_i.$$

It is easy to see that $0 \leqslant w_a, w_b, w^* \leqslant 1$. Note that $\lambda_a = r!(w_a + x_b w^*)$ and $\lambda_b = r!(w_b + x_a w^*)$. Hence,

$$\lambda_b - \lambda_a = r![w_b - w_a + (x_a - x_b)w^*].$$

Let

$$d = \frac{1}{2r!}(\lambda_b - \lambda_a).$$

Consider a new weight assignment $\tilde{y} = (y_1, \dots, y_n)$ defined by letting $y_a = x_a - d$, $y_b = x_b + d$ and, for all $i \in [t] \setminus \{a, b\}$, $y_i = x_i$. Then

$$\begin{aligned}
p_G(\tilde{y}) - p_G(\tilde{x}) &= r![(x_a - d)(x_b + d) - x_a x_b)w^* - dw_a + dw_b]. \\
&= r![d((x_a - x_b)w^* + (w_b - w_a)] - d^2 w^*] \\
&\geqslant r!d[((x_a - x_b)w^* + (w_b - w_a)] - d] \\
&= d[\lambda_b - \lambda_a - r!d] = \frac{1}{2} d(\lambda_b - \lambda_a) = \frac{1}{4r!}(\lambda_b - \lambda_a)^2 > \eta,
\end{aligned}$$

contradicting our assumption that $p_G(\tilde{x}) \geqslant \lambda(G) - \eta$. $\square$

Given $0 < \beta \leqslant 1$ and an r-graph G on $[t]$, let

$$\lambda_\beta(G) = \max\left\{ p_G(x_1,\dots,x_t) : \forall i \in [t]\, x_i \geqslant 0,\ \sum_i x_i = 1, \max_i x_i = \beta \right\}.$$

Clearly,

$$\lambda(G) = \max_{0<\beta\leqslant 1} \lambda_\beta(G).$$

The following lemma played a crucial role in Sidorenko's arguments.

Lemma 10.2 (Lemma 3.3 of [24]). *Let $k, r \geqslant 2$ be integers, where $k \geqslant M_{r-1}$. Let $0 < \beta \leqslant 1$ be a real. Let T be a k-vertex tree that satisfies the Erdős–Sós conjecture. Let T be the $(r-2)$-fold enlargement of T. If G is an F-free r-graph then $\lambda_\beta(G) \leqslant (k-2) f_r(z)$, where*

$$z = \max\left\{ \frac{1}{\beta} - r + 3, k \right\}.$$

Let us also mention a useful fact about the function $f_r(x)$, namely

$$\frac{f_r(x)}{f_{r-1}(x)} = \left(\frac{x+r-4}{x+r-3} \right)^{r-1}. \tag{10.3}$$

Further, let us recall the well-known fact that

$$\pi_\lambda(K_k) = \pi(K_k) = \frac{k-2}{k-1}$$

(see *e.g.* [5] and [17]).

Lemma 10.3. *Let $k \geqslant 3$. Let T be a k-vertex tree that satisfies the Erdős–Sós conjecture and let F be the $(r-2)$-enlargement of T, where $r \geqslant 2$ and $k \geqslant M_r$. For every real $\alpha > 0$, there exists a real $\gamma = \gamma(\alpha) > 0$ such that if G is an F-free r-graph on $[t]$ and $\tilde{x} = (x_1, \dots, x_t)$ a 1-sum weight assignment on G with $p_G(\tilde{x}) \geqslant (k-2) f_r(k) - \gamma$, then there exists $i \in [t]$ such that*

(i)

$$\left| x_i - \frac{1}{k+r-3} \right| < \alpha.$$

(ii)

$$\left| \sum_{j \in V(\mathcal{L}_G(i))} x_j - \frac{k+r-4}{k+r-3} \right| < \alpha.$$

(iii)

$$\lambda_i > r(k-2) f_r(k) - \alpha.$$

Proof. Let $\alpha > 0$ be given. Choose a sufficiently small real $d > 0$ such that

$$\frac{1}{k+r-3+d} > \frac{1}{k+r-3} - \frac{\alpha}{2}.$$

Since $k > M_r$ and M_r is the rightmost local maximum of $f_r(x)$, $f_r(x)$ is strictly decreasing on $[k,\infty)$. Choose $\gamma > 0$ to be sufficiently small that

$$\gamma < f_r(k) - f_r(k+d), \quad 3r!\sqrt{\gamma} < \frac{\alpha}{2} \quad \text{and} \quad \left(1 - \frac{3r!\sqrt{\gamma}}{r(k-2)f_r(k)}\right)^{1/(r-1)} > 1-\alpha. \tag{10.4}$$

Let $\beta = \max_j x_j$ and $\lambda_{\max} = \max_j \lambda_j$. If $1/\beta - r + 3 > k + d$ and by Lemma 10.2,

$$p_G(\tilde{x}) \leqslant \lambda_\beta(G) \leqslant (k-2)f_r(k+d) < (k-2)[f_r(k)-\gamma] < (k-2)f_r(k) - \gamma,$$

contradicting our assumption. Hence $1/\beta - r + 3 \leqslant k + d$. Solving for β, by our choice of d we have that

$$\beta \geqslant \frac{1}{k+r-3+d} > \frac{1}{k+r-3} - \frac{\alpha}{2}.$$

By Lemma 10.1, for some $i \in [n]$, say $i = 1$, we have

$$x_1 \geqslant \beta - 2r!\sqrt{\gamma} \quad \text{and} \quad \lambda_1 \geqslant \lambda_{\max} - 2r!\sqrt{\gamma}. \tag{10.5}$$

Since $3r!\sqrt{\gamma} < \alpha/2$ by our choice of γ,

$$x_1 \geqslant \frac{1}{k+r-3} - \frac{\alpha}{2} - 2r!\sqrt{\gamma} \geqslant \frac{1}{k+r-3} - \alpha. \tag{10.6}$$

By (10.2), (10.5) and our assumption that $p_G(\tilde{x}) \geqslant (k-2)f_r(k) - \gamma$,

$$\begin{aligned} \lambda_1 &\geqslant \lambda_{\max} - 2r!\sqrt{\gamma} \geqslant r((k-2)f_r(k) - \gamma) - 2r!\sqrt{\gamma} \\ &> r(k-2)f_r(k) - 3r!\sqrt{\gamma} \geqslant r(k-2)f_r(k) - \alpha. \end{aligned} \tag{10.7}$$

This proves item (iii). Next, we prove that

$$\sum_{j \in V(\mathcal{L}_G(i))} x_i > \frac{k+r-4}{k+r-3} - \alpha.$$

If $r = 2$ then $\lambda_1 = 2\sum_{j \in N_G(1)} x_j$ and hence

$$\sum_{j \in V(\mathcal{L}_G(1))} x_j = \frac{\lambda_1}{2} > (k-2)f_2(k) - \frac{\alpha}{2} \geqslant \frac{k-2}{k-1} - \alpha.$$

We henceforth assume that $r \geqslant 3$. Let $s = \sum_{j \in \mathcal{L}_G(1)} x_j$. For each $j \in V(\mathcal{L}_G(1))$ let $y_j = x_j/s$. Then $\sum_{j \in V(\mathcal{L}_G(1)} y_j = 1$. Let $\tilde{y}$ denote the 1-sum weight assignment on $\mathcal{L}_G(1)$ defined by the y_j. Let F' denote the $(r-3)$-enlargement of T. Since G is F-free, $\mathcal{L}_G(1)$ is F'-free. Since F' is the $(r-3)$-enlargement of T, where T is a k-vertex tree satisfying the Erdős–Sós conjecture and $k \geqslant M_r \geqslant M_{r-1}$, by Theorem 6.6, then $\pi_\lambda(F') \leqslant (k-2)f_{r-1}(k)$. Hence,

$$\lambda_1 = r! \sum_{e \in \mathcal{L}_G(1)} \prod_{j \in e} x_j = rs^{r-1} \cdot (r-1)! \sum_{e \in \mathcal{L}_G(1)} \prod_{j \in e} y_j \leqslant rs^{r-1} \cdot (k-2)f_{r-1}(k). \tag{10.8}$$

By (10.7) and (10.8), we have

$$rs^{r-1} \cdot (k-2)f_{r-1}(k) > r(k-2)f_r(k)\left(1 - \frac{3r!\sqrt{\gamma}}{r(k-2)f_r(k)}\right).$$

Using (10.3), we have

$$s^{r-1} > \left(\frac{k+r-4}{k+r-3}\right)^{r-1}\left(1-\frac{3r!\sqrt{\gamma}}{r(k-2)f_r(k)}\right).$$

Hence, by our choice of γ given in (10.4), we have

$$\begin{aligned}\sum_{j\in V(\mathcal{L}_G(1))} x_j = s &> \frac{k+r-4}{k+r-3}\left(1-\frac{3r!\sqrt{\gamma}}{r(k-2)f_r(k)}\right)^{1/(r-1)}\\ &> \frac{k+r-4}{k+r-3}(1-\alpha) > \frac{k+r-4}{k+r-3}-\alpha. \end{aligned} \tag{10.9}$$

Now, (10.6) and (10.9) together prove items (ii) and (iii). □

We need another lemma from [24]. Given a graph G, let

$$d(G) = \max_{H\subseteq G}\frac{2e(H)}{n(H)}.$$

So, $d(G)$ is the maximum average degree of a subgraph of G over all subgraphs of G.

Lemma 10.4 ([24] Theorem 2.4). *Let G be a graph on $[t]$. Let $\tilde{y} = (y_1,\dots,y_t)$ be a weight assignment on G where $\max_i y_i = 1$. Then*

$$\frac{p_G(\tilde{y})}{\sum_{i=1}^t y_i} \leqslant d(G).$$

Corollary 10.5. *Let G be a graph on $[t]$. Let $\tilde{x} = (x_1,\dots,x_t)$ be a 1-sum weight assignment on G where $\max_i x_i = \beta$. Then $p_G(\tilde{x}) \leqslant \beta d(G)$.*

Proof. For each $i\in[t]$, let $y_i = x_i/\beta$. Then $\max_i y_i = 1$ and $\sum_i y_i = 1/\beta$. Using Lemma 10.4, we have

$$p_G(\tilde{x}) = \beta^2 p_G(\tilde{y}) \leqslant \beta^2 d(G)\sum_i y_i = \beta d(G).$$ □

Lemma 10.6. *Let H be a graph with average degree d and let G be obtained from H by adding a new vertex and making it adjacent to all of $V(H)$. Then G has average degree at least $d+1$.*

Proof. Suppose H has p vertices. Clearly $p\geqslant d+1$. We have $n(G)=p+1$ and $e(G)=pd/2+p$. So G has average degree

$$\frac{2e(G)}{n(G)} = \frac{pd+2p}{p+1} \geqslant \frac{pd+p+d+1}{p+1} = d+1.$$ □

Lemma 10.7. *Let d be a positive integer. Let $0<\varepsilon<1$ be a real. There exists a real $\delta_d(\varepsilon)>0$ such that if G a graph on $[t]$ with $d(G)\leqslant d$ and $\tilde{x} = (x_1,\dots x_t)$ is a 1-sum weight assignment on*

G with

$$p_G(\tilde{x}) \geqslant \frac{d}{d+1} - \delta_d,$$

then there exists $I \subseteq [t]$ *with* $|I| \leqslant d+1$ *such that* $\sum_{i \in I} x_i \geqslant 1-\varepsilon$.

Proof. We use induction on d. For the basis step, let $d = 1$. Let $\delta_1(\varepsilon) = \varepsilon/2$. Suppose $d(G) \leqslant 1$ and $p_G(\tilde{x}) \geqslant \frac{1}{2} - \delta_1$. By our assumption about G, each non-trivial component of G is a single edge. Suppose that there are s non-trivial components with total vertex weights $w_1, \ldots, w_s$, respectively, where, without loss of generality, suppose $w_1 = \max_j w_j$. We have

$$\frac{1}{2} - \delta_1(\varepsilon) \leqslant p_G(x) \leqslant 2! \sum_{i=1}^{s} \frac{w_i^2}{4} \leqslant \frac{1}{2} w_1 (w_1 + w_2 + \ldots + w_s) \leqslant \frac{1}{2} w_1.$$

Hence $w_1 \geqslant 1 - 2\delta_1(\varepsilon) = 1 - \varepsilon$, implying that there exists $I \subseteq [t], |I| = 2$ with $\sum_{i \in I} x_i \geqslant 1 - \varepsilon$.

For the induction step, let $d \geqslant 2$. Choose a small real α such that

$$0 < \alpha < \frac{\varepsilon}{4}$$

and

$$\left(\frac{d}{d+1} + \alpha\right)^{-2} \cdot \left(\frac{d(d-1)}{(d+1)^2} - 7\alpha\right) > \frac{d-1}{d} - \delta_{d-1}\left(\frac{\varepsilon}{2}\right). \tag{10.10}$$

Choose $0 < \delta_d < \alpha - 2\alpha^2$ to be sufficiently small that

$$\delta_d + 2r! \sqrt{\delta_d} < \alpha.$$

Suppose

$$p_G(\tilde{x}) \geqslant \frac{d}{d+1} - \delta_d.$$

By Lemma 10.1, there exists $i \in [t]$, say $i = 1$, such that

$$x_1 \geqslant \max_i x_i - 2r! \sqrt{\delta_d} \quad \text{and} \quad \lambda_1 \geqslant \max_i \lambda_i - 2r! \sqrt{\delta_d}.$$

By (10.2), we have

$$\max_i \lambda_i \geqslant 2 p_G(\tilde{x}) \geqslant \frac{2d}{d+1} - 2\delta_d$$

and hence

$$\lambda_1 \geqslant \frac{2d}{d+1} - 2\delta_d - 2r! \sqrt{\delta_d}.$$

Since $\lambda_1 = 2 \sum_{j \in N_G(1)} x_j$, this also yields

$$\sum_{j \in N_G(1)} x_j \geqslant \frac{d}{d+1} - \delta_d - r! \sqrt{\delta_d} \geqslant \frac{d}{d+1} - \alpha. \tag{10.11}$$

Hence

$$x_1 \leqslant \frac{1}{d+1} + \alpha.$$

Let $\beta = \max_i x_i$. By Corollary 10.5, we have

$$\frac{d}{d+1} - \delta_d \leqslant p_G(\tilde{x}) \leqslant \beta d.$$

Hence,

$$\beta \geqslant \frac{1}{d+1} - \frac{\delta_d}{d+1}$$

and thus

$$x_1 \geqslant \beta - 2r!\sqrt{\delta_d} \geqslant \frac{1}{d+1} - \frac{\delta_d}{d+1} - 2r!\sqrt{\delta_d} \geqslant \frac{1}{d+1} - \alpha. \tag{10.12}$$

Let $N = N_G(1)$ and $\overline{N} = [t] \setminus (N \cup \{1\})$. By (10.11) and (10.12),

$$\sum_{j \in \overline{N}} x_j < 2\alpha. \tag{10.13}$$

Since 1 is adjacent to all of N and $d(G) \leqslant d$, by Lemma 10.6, $d(G[N]) \leqslant d-1$. Let $s = \sum_{j \in N} x_j$. By (10.11) and (10.12), we have

$$\frac{d}{d+1} - \alpha \leqslant s \leqslant \frac{d}{d+1} + \alpha.$$

Since $\lambda_1 = 2s$, we also have

$$2\left(\frac{d}{d+1} - \alpha\right) \leqslant \lambda_1 \leqslant 2\left(\frac{d}{d+1} + \alpha\right).$$

For each $j \in N$ let $y_j = x_j/s$. Then $\sum_{j \in N} y_j = 1$. Let $\tilde{y}$ denote the 1-sum weight assignment on $G[N]$ given by the y_j. Using the upper bounds on λ_1, x_1, and (10.13), we have

$$\begin{aligned} s^2 p_{G[N]}(\tilde{y}) = 2 \sum_{\{i,j\} \in G[N]} x_i x_j &\geqslant p_G(\tilde{x}) - \lambda_1 x_1 - 2 \sum_{j \in [t]} x_j \sum_{j \in \overline{N}} x_j \\ &\geqslant \left(\frac{d}{d+1} - \delta_d\right) - 2\left(\frac{d}{d+1} + \alpha\right)\left(\frac{1}{d+1} + \alpha\right) - 4\alpha \\ &= \frac{d(d-1)}{(d+1)^2} - \delta_d - 6\alpha - 2\alpha^2 \\ &\geqslant \frac{d(d-1)}{(d+1)^2} - 7\alpha. \end{aligned}$$

Since

$$s \leqslant \frac{d}{d+1} + \alpha,$$

this yields

$$\left(\frac{d}{d+1} + \alpha\right)^2 p_{G[N]}(\tilde{y}) \geqslant \frac{d(d-1)}{(d+1)^2} - 7\alpha.$$

By (10.10), this yields

$$p_{G[N]}(\tilde{y}) \geqslant \frac{d-1}{d} - \delta_{d-1}\left(\frac{\varepsilon}{2}\right). \tag{10.14}$$

Since $d(G[N]) \leqslant d-1$, by (10.14) and the induction hypothesis, there exists $J \subseteq N$ with $|J| \leqslant d$ such that $\sum_{j\in J} y_j \geqslant 1-\varepsilon/2$. Hence

$$\sum_{j\in J} x_j \geqslant s\sum_{j\in J} y_j \geqslant s\left(1-\frac{\varepsilon}{2}\right) \geqslant \left(\frac{d}{d+1}-\alpha\right)\left(1-\frac{\varepsilon}{2}\right) > \frac{d}{d+1}-\alpha-\frac{\varepsilon}{2}. \tag{10.15}$$

Let $I = J \cup \{1\}$. By (10.12) and (10.15), we have

$$\sum_{i\in I} x_i \geqslant \frac{1}{d+1}-\alpha+\frac{d}{d+1}-\alpha-\frac{\varepsilon}{2} \geqslant 1-\varepsilon.$$

This completes the induction step and the proof. □

In the next lemma, we use Lemma 10.7 to establish stability of the Lagrangian function of T-free graphs where T is a tree that satisfies the Erdős–Sós conjecture. Note that such stability obviously does not exist for the Lagrangian function of K_k-free graphs. Thus we consider any complete $(k-1)$-partite graph G, which is K_k-free. Any weight assignment $\tilde{x}$ on G in which the total vertex weight on each part is $1/(k-1)$ satisfies

$$p_G(\tilde{x}) = \frac{k-2}{k-1} = \pi_\lambda(K_k),$$

though any two such $\tilde{x}$ can be very different.

Lemma 10.8. *Let $k \geqslant 3$ be an integer. Let T be a k-vertex tree that satisfies the Erdős–Sós conjecture. Let $0<\varepsilon<1$ be a real. There exists a real $\delta_k(\varepsilon)>0$ such that the following is true: If G is a T-free graph on $[t]$ and $\tilde{x}=(x_1,\dots x_t)$ is a 1-sum weight assignment on G such that*

$$p_G(\tilde{x}) \geqslant \frac{k-2}{k-1}-\delta_k,$$

then there exists $I \subseteq [t]$ with $|I| \leqslant k-1$ such that $\sum_{i\in I} x_i \geqslant 1-\varepsilon$.

Proof. Since T satisfies the Erdős–Sós conjecture and G is T-free, we have $d(G) \leqslant k-2$. The lemma follows from Lemma 10.7 with $d=k-2$. □

We can now use Lemma 10.8 to establish stability of the Lagrangian function of an r-graph not containing a given enlarged tree.

Lemma 10.9. *Let $k \geqslant 3, r \geqslant 2$ be integers where $k \geqslant M_r$. Let T be a k-vertex tree that satisfies the Erdős–Sós conjecture. Let F be the $(r-2)$-fold enlargement of T. Let $\varepsilon>0$ be any real. There exists a real $\hat{\delta}_r=\hat{\delta}_r(\varepsilon)>0$ such that the following holds. Let G be a F-free r-graph on $[t]$ and let $\tilde{x}=(x_1,\dots,x_t)$ be a 1-sum weight assignment on G such that*

$$p_G(\tilde{x}) \geqslant \frac{(k+r-3)_r}{(k+r-3)^r}-\hat{\delta}_r.$$

Then there exists $I \subseteq [t]$ with $|I| \leqslant r+k-3$ such that $\sum_{i\in I} x_i \geqslant 1-\varepsilon$.

Proof. We use induction on r. The basis step $r=2$ was established by Lemma 10.8. For the induction step, let $r \geqslant 3$. Let $\varepsilon > 0$ be given. Let α be a real such that

$$0 < \alpha < \frac{\varepsilon}{4},$$

and

$$r^{-1}\left(\frac{k+r-4}{k+r-3}+\alpha\right)^{-(r-1)} \cdot \left[r\frac{(k+r-4)_{r-1}}{(k+r-3)^{r-1}} - \alpha\right] > \frac{(k+r-4)_{r-1}}{(k+r-4)^{r-1}} - \hat{\delta}_{r-1}\left(\frac{\varepsilon}{2}\right). \qquad (10.16)$$

Let $\hat{\delta}_r = \gamma(\alpha)$, where the function γ is given in Lemma 10.3. Suppose $\tilde{x}$ is a 1-sum weight assignment on G with

$$p_G(\tilde{x}) \geqslant \frac{(k+r-3)_r}{(k+r-3)^r} - \hat{\delta}_r.$$

By Lemma 10.3, there exists $i \in [t]$, say $i=1$, such that with $s = \sum_{j\in V(\mathcal{L}_G(1))} x_j$ we have

$$\left|x_1 - \frac{1}{k+r-3}\right| < \alpha, \quad \left|s - \frac{k+r-4}{k+r-3}\right| < \alpha \quad \text{and} \quad \lambda_1 > r\frac{(k+r-3)_r}{(k+r-3)^r} - \alpha. \qquad (10.17)$$

For each $j \in V(\mathcal{L}_G(1))$ let $y_j = x_j/s$. Then $\sum_{j\in V(\mathcal{L}_G(1))} y_j = 1$. Let $\tilde{y}$ denote the 1-sum weight function on $\mathcal{L}_G(1)$ defined by the y_j. As usual we have

$$\lambda_1 = r!\cdot \sum_{e\in\mathcal{L}_G(1)}\prod_{j\in e} x_j = rs^{r-1}(r-1)!\cdot \sum_{e\in\mathcal{L}_G(1)}\prod_{j\in e} y_j \leqslant rs^{r-1} p_{\mathcal{L}_G(1)}(\tilde{y}).$$

Hence, by (10.17), we have

$$r\left(\frac{k+r-4}{k+r-3}+\alpha\right)^{r-1} p_{\mathcal{L}_G(1)}(\tilde{y}) > r\frac{(k+r-3)_r}{(k+r-3)^r} - \alpha = r\frac{(k+r-4)_{r-1}}{(k+r-3)^{r-1}} - \alpha.$$

By (10.16), this yields

$$p_{\mathcal{L}_G(1)}(\tilde{y}) > \frac{(k+r-4)_{r-1}}{(k+r-4)^{r-1}} - \hat{\delta}_{r-1}\left(\frac{\varepsilon}{2}\right) = (k-2)f_{r-1}(k) - \hat{\delta}_{r-1}\left(\frac{\varepsilon}{2}\right). \qquad (10.18)$$

Let F' denote the $(r-3)$-enlargement of T. Since G is T-free, clearly $\mathcal{L}_G(1)$ is T'-free. Since F' is the $(r-3)$-enlargement of T, where T is k-vertex tree satisfying the Erdős–Sós conjecture, and $k \geqslant M_r \geqslant M_{r-1}$, by (10.18) and the induction hypothesis, there exists $J \subseteq V(\mathcal{L}_G(1))$ such that $\sum_{j\in J} y_j \geqslant 1 - \varepsilon/2$. Hence

$$\sum_{j\in J} x_j \geqslant s\left(1-\frac{\varepsilon}{2}\right) \geqslant \left(\frac{k+r-4}{k+r-3}-\alpha\right)\left(1-\frac{\varepsilon}{2}\right) \geqslant \frac{k+r-4}{k+r-3} - \alpha - \frac{\varepsilon}{2}. \qquad (10.19)$$

Let $I = J \cup \{1\}$. By (10.17) and (10.19) we have

$$\sum_{i\in I} x_i \geqslant \frac{1}{k+r-3} - \alpha + \frac{k+r-4}{k+r-3} - \alpha - \frac{\varepsilon}{2} \geqslant 1 - \varepsilon.$$

This completes the induction and the proof. $\square$

Proof of Theorem 6.7. Let $\varepsilon > 0$ be given. We may assume ε to be sufficiently small that $\varepsilon < \gamma_0$, where γ_0 is defined in Theorem 8.7. Let

$$\gamma = \min\left\{\frac{1}{2}\hat{\delta}_r\left(\frac{\varepsilon}{2}\right), \frac{\varepsilon}{2}\right\},$$

where $\hat{\delta}_r$ is given in Lemma 10.9. By our definition, $\gamma < \gamma_0$. Let δ and n_0 be the constants guaranteed by Theorem 8.7 for the above-defined γ. Let n be sufficiently large that $n \geqslant n_0$ and that n satisfies some other inequalities given below. Let G be any $\mathcal{K}^F_{k+r-2}$-free r-graph on $[n]$ with

$$|G| > \left(\frac{(k+r-3)_r}{(k+r-3)^r} - \delta\right)\binom{n}{r}.$$

Let G^* be the final graph produced by Algorithm 8.1 with threshold

$$\frac{(k+r-3)_r}{(k+r-3)^r} - \gamma.$$

By Theorem 6.2, G^* is $\mathcal{K}^F_{k+r-2}$-free, $n(G^*) \geqslant (1-\gamma)n$, and

$$G^* \text{ is } \left(\frac{(k+r-3)_r}{(k+r-3)^r} - \gamma\right)\text{-dense.}$$

Let $N = n(G^*)$. Since G^* is $\left(\frac{(k+r-3)_r}{(k+r-3)^r} - \gamma\right)$-dense, we have

$$|G^*| \geqslant \left(\frac{(k+r-3)_r}{(k+r-3)^r} - \gamma\right)\binom{N}{r}. \tag{10.20}$$

Suppose G has s equivalence classes $A_1, \ldots, A_s$. Let S consist of one vertex from each equivalence class of G^*. Without loss of generality, suppose $S = [s]$. Then $G^*[S]$ covers pairs and G^* is a blowup of $G^*[S]$. For each $i \in [s]$, let $x_i = |A_i|/N$. Then $\sum_i x_i = 1$. So $\tilde{x} = (x_1, \ldots, x_s)$ is a 1-sum weight assignment on $G^*[S]$. Also,

$$p_G(\tilde{x}) = r! \sum_{e \in G} \prod_{i \in e} x_i = \frac{r!}{N^r}|G^*| \geqslant \frac{(k+r-3)_r}{(k+r-3)^r} - 2\gamma \geqslant \frac{(k+r-3)_r}{(k+r-3)^r} - \hat{\delta}_r\left(\frac{\varepsilon}{2}\right),$$

where the two inequalities follow from (10.20), our definition of γ, and the assumption that N is sufficiently large. By Lemma 10.9, there exists $I \subseteq [s]$, where $|I| \leqslant k+r-3$ such that $\sum_{i \in I} x_i \geqslant 1 - \varepsilon/2$. Let $W = \bigcup_{i \in I} A_i$. Then

$$|W| \geqslant \left(1 - \frac{\varepsilon}{2}\right)N \geqslant (1-\gamma_0)N.$$

By Theorem 8.7, $G[W]$ is $(k+r-3)$-partite. Since

$$|W| \geqslant \left(1 - \frac{\varepsilon}{2}\right)N \geqslant \left(1 - \frac{\varepsilon}{2}\right)(1-\gamma)n \geqslant \left(1 - \frac{\varepsilon}{2}\right)^2 n > n - \varepsilon n,$$

G can be made $(k+r-3)$-partite by deleting at most εn vertices. So, $\mathcal{K}^F_{k+r-2}$ is $(k+r-3)$-stable. $\square$

Note added in proof

Around the time of our submission of the paper, we learned that S. Norin and L. Yepremyan [19] had independently obtained the main results in this paper, using a very different stability method that they developed in [18]. See also [26] for a detailed description of their method.

Acknowledgements

We thank the referee for his/her thoughtful comments, which helped to improve our presentation.

References

[1] Ajtai, M., Komlós, J., Simonovits, M. and Szemerédi, E. Unpublished. The solution of the Erdos-Sos conjecture for large trees, in preparation.
[2] Bollobás, B. (1974) Three-graphs without two triples whose symmetric difference is contained in a third. *Discrete Math.* **8** 21–24.
[3] Erdős, P. (1964) On extremal problems of graphs and generalized graphs. *Israel J. Math.* **2** 183–190.
[4] Erdős, P. (1965) A problem on independent r-tuples. *Ann. Univ. Sci. Budapest* **8** 93–95.
[5] Erdős, P. and Simonovits, M. (1966) A limit theorem in graph theory. *Studia. Sci. Math. Hungar.* **1** 51–57.
[6] Frankl, P. (2013) Improved bounds for Erdős' matching conjecture. *J. Combin. Theory Ser. A* **120** 1068–1072.
[7] Frankl, P. and Füredi, Z. (1983) A new generalization of the Erdős–Ko–Rado theorem. *Combinatorica* **3** 341–349.
[8] Frankl, P. and Füredi, Z. (1989) Extremal problems whose solutions are the blowups of the small Witt-designs. *J. Combin. Theory Ser. A* **52** 129–147.
[9] Füredi, Z., Pikhurko, O. and Simonovits, M. (2005) On triple systems with independent neighbourhoods. *Combin. Probab. Comput.* **14** 795–813.
[10] Hefetz, D. and Keevash, P. (2013) A hypergraph Turán theorem via Lagrangians of intersecting families. *J. Combin. Theory Ser. A* **120** 2020–2038.
[11] Jiang, T., Peng, Y. and Wu, B. Turán numbers of extensions of some sparse hypergraphs via Lagrangians, in preparation. `arXiv:1609.08983`
[12] Katona, G. O. H. (1974) Extremal problems for hypergraphs. In *Combinatorics* (M. Hall Jr and J. H. van Lint, eds), Vol. 16 of the NATO Advanced Study Institutes series, Springer pp. 215–244.
[13] Keevash, P. (2011) *Hypergraph Turán Problems*, Surveys in Combinatorics, Cambridge University Press, pp. 83–140.
[14] Keevash, P. and Mubayi, D. (2004) Stability theorems for cancellative hypergraphs. *J. Combin. Theory Ser. B* **92** 163–175.
[15] Mubayi, D. (2006) A hypergraph extension of Turán's theorem. *J. Combin. Theory Ser. B* **96** 122–134.
[16] Mubayi, D. and Pikhurko, O. (2007) A new generalization of Mantel's theorem to k-graphs. *J. Combin. Theory Ser. B* **97** 669–678.
[17] Motzkin, T. S. and Straus, E. G. (1965) Maxima for graphs and a new proof of a theorem of Turán. *Canad. J. Math.* **17** 533–540.
[18] Norin, S. and Yepremyan, L. (2017) Turán number of generalized triangles. *J. Combin. Theory Ser. A* **146** 312–343.
[19] Norin, S. and Yepremyan, L. Turán numbers of extensions, in preparation. `arXiv:1510.04689`
[20] Pikhurko, O. (2008) An exact Turán result for the generalized triangle. *Combinatorica* **28** 187–208.
[21] Pikhurko, O. (2013) Exact computation of the hypergraph Turán function for expanded complete 2-graphs. *J. Combin. Theory Ser. B* **103** 220–225.
[22] Shearer, J. (1996) A new construction for cancellative families of sets. *Electron. J. Combin.* **3** R15.

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

[24] Sidorenko, A. F. (1989) Asymptotic solution for a new class of forbidden r-graphs. *Combinatorica* **9** 207–215.

[25] Simonovits, M. (1968) A method for solving extremal problems in graph theory. In *Theory of Graphs: Proc. Coll. Tihany, Hungary* (P. Erdős and G. Katona, eds), Academic Press, pp. 279–319.

[26] Yepremyan, L. (2016) PhD dissertation, McGill University.