On Possible Turán Densities

Oleg Pikhurko*
Mathematics Institute and DIMAP
University of Warwick
Coventry CV4 7AL, UK

Abstract

The *Turán density* $\pi(\mathcal{F})$ of a family $\mathcal{F}$ of k-graphs is the limit as $n \to \infty$ of the maximum edge density of an $\mathcal{F}$-free k-graph on n vertices. Let $\Pi_\infty^{(k)}$ consist of all possible Turán densities and let $\Pi_{\mathrm{fin}}^{(k)} \subseteq \Pi_\infty^{(k)}$ be the set of Turán densities of finite k-graph families.

Here we prove that $\Pi_{\mathrm{fin}}^{(k)}$ contains every density obtained from an arbitrary finite construction by optimally blowing it up and using recursion inside the specified set of parts. As an application, we show that $\Pi_{\mathrm{fin}}^{(k)}$ contains an irrational number for each $k \geq 3$.

Also, we show that $\Pi_\infty^{(k)}$ has cardinality of the continuum. In particular, $\Pi_\infty^{(k)} \neq \Pi_{\mathrm{fin}}^{(k)}$.

1 Introduction

Let $\mathcal{F}$ be a (possibly infinite) family of *k-graphs* (that is, k-uniform set systems). We call elements of $\mathcal{F}$ *forbidden.* A k-graph G is *$\mathcal{F}$-free* if no member $F \in \mathcal{F}$ is a subgraph of G, that is, we cannot obtain F by deleting some vertices and edges from G. The *Turán function* $\mathrm{ex}(n, \mathcal{F})$ is the maximum number of edges that an $\mathcal{F}$-free k-graph on n vertices can have. This is one of the central questions of extremal combinatorics that goes back to the fundamental paper of Turán [45]. We refer the reader to the surveys of the Turán function by Füredi [22], Keevash [27], and Sidorenko [41].

As it was observed by Katona, Nemetz, and Simonovits [26], the ratio $\mathrm{ex}(n, \mathcal{F})/\binom{n}{k}$ is non-increasing in n. In particular, the limit

$$\pi(\mathcal{F}) := \lim_{n\to\infty} \frac{\mathrm{ex}(n, \mathcal{F})}{\binom{n}{k}}$$

exists. It is called the *Turán density* of $\mathcal{F}$. Let $\Pi_\infty^{(k)}$ consist of all possible Turán densities of k-graph families and let $\Pi_{\mathrm{fin}}^{(k)}$ be the set of all possible Turán densities when *finitely many* k-graphs are forbidden. Clearly, $\Pi_{\mathrm{fin}}^{(k)} \subseteq \Pi_\infty^{(k)}$.

*The author was supported by the European Research Council (grant agreement no. 306493) and the National Science Foundation of the USA (grant DMS-1100215). This project was initiated during the workshop *"Hypergraph Turán Problem"* held at the American Institute of Mathematics, Palo Alto, March 21–25, 2011.

For $k=2$, the celebrated Erdős-Stone-Simonovits Theorem [16, 17] determines the Turán density for every family $\mathcal{F}$. In particular, we have

$$\Pi_{\mathrm{fin}}^{(2)}=\Pi_{\infty}^{(2)}=\left\{\frac{m-1}{m}: m=1,2,3,\ldots,\infty\right\}. \tag{1}$$

(It is convenient to allow empty forbidden families, so $1\in\Pi_{\mathrm{fin}}^{(k)}$ for every k.)

Unfortunately, the Turán function for *hypergraphs* (that is, k-graphs with $k\geq 3$) is much more difficult and many problems (even rather basic ones) are wide open.

Arguably, the case when $|\mathcal{F}|=1$ is the most interesting one. However, even very simple forbidden hypergraphs turned out to be notoriously difficult. For example, the famous conjecture of Turán from 1941 that $\pi(\{K_4^3\})=5/9$ is still open, where K_m^k denotes the complete k-graph on m vertices. For no $3\leq k<m$ is the value of $\pi(\{K_m^k\})$ known, despite the \$1000 prize of Erdős. Razborov [36, Page 247] writes that *"these questions became notoriously known ever since as some of the most difficult open problems in discrete mathematics"*.

On the other hand, some Turán-type results stop being true if only one subgraph is to be forbidden. One such example is the Ruzsa-Szemerédi theorem [38] that $\mathrm{ex}(n,\mathcal{F})=o(n^2)$, where $\mathcal{F}$ consists of all 3-graphs with 6 vertices and at least 3 edges (while $\mathrm{ex}(n,\{F\})=\Omega(n^2)$ for every $F\in\mathcal{F}$). Some other problems (such as various intersection questions for uniform set systems, see e.g. [22]) can be restated in terms of the Turán function and require that more than one subgraph is forbidden. Also, new interesting phenomena (such as, for example, non-principality, see [3, 32]) appear when one allows more than one forbidden k-graph. Last but not least, by solving (perhaps more tractable) cases with $|\mathcal{F}|>1$ we may get more insight about the case $|\mathcal{F}|=1$. In fact, some proofs that determine $\pi(\{F\})$ proceed by forbidding some extra hypergraphs whose addition does not affect the Turán density, see e.g. [2, 4, 19, 31, 39].

Little is known about $\Pi_{\mathrm{fin}}^{(k)}$ and $\Pi_{\infty}^{(k)}$ for $k\geq 3$. Brown and Simonovits [8, Theorem 1] noted that for every $\mathcal{F}$ and $\varepsilon>0$ there is a finite $\mathcal{F}'\subseteq\mathcal{F}$ with $\pi(\mathcal{F}')\leq\pi(\mathcal{F})+\varepsilon$ (while, trivially, $\pi(\mathcal{F})\leq\pi(\mathcal{F}')$). It follows that $\Pi_{\infty}^{(k)}$ lies in the closure of $\Pi_{\mathrm{fin}}^{(k)}$. Here we show the following results about $\Pi_{\infty}^{(k)}$ with the first one implying that in fact $\Pi_{\infty}^{(k)}$ is the closure of $\Pi_{\mathrm{fin}}^{(k)}$.

Proposition 1 *For every $k\geq 3$ the set $\Pi_{\infty}^{(k)}\subseteq[0,1]$ is closed.*

Theorem 2 *For every $k\geq 3$ the set $\Pi_{\infty}^{(k)}$ has cardinality of the continuum.*

Since the number of finite families of k-graphs (up to isomorphism) is countable, Theorem 2 implies that $\Pi_{\mathrm{fin}}^{(k)}\neq\Pi_{\infty}^{(k)}$ for $k\geq 3$, answering one part of a question of Baber and Talbot [2, Question 31].

Erdős [14] proved that $\Pi_{\infty}^{(k)}\cap(0,k!/k^k)=\emptyset$, that is, if the Turán density is positive, then it is at least $k!/k^k$. Let us call a real $\alpha\in[0,1]$ a *jump for k-graphs* if there is $\varepsilon>0$ such that $\Pi_{\infty}^{(k)}\cap(\alpha,\alpha+\varepsilon)=\emptyset$. For example, every $\alpha\in[0,1]$ is a jump for graphs by (1) and every $\alpha\in[0,k!/k^k)$

is a jump for k-graphs by [14]. The break-through paper of Frankl and Rödl [21] showed that non-jumps exist for every $k \geq 3$, disproving the \$1000 conjecture of Erdős that $\Pi_\infty^{(k)}$ is well-ordered with respect to the usual order on the reals. Further results on (non-) jumps were obtained in [1, 20, 33] and many other papers. Our Theorem 2 shows that $\Pi_\infty^{(k)}$ is "very far" from being well-ordered for $k \geq 3$. Since each jump is followed by an interval disjoint from $\Pi_\infty^{(k)}$, at most countably many elements of $\Pi_\infty^{(k)}$ can be jumps. Thus, by Theorem 2, the set of non-jumps has cardinality of the continuum.

Very few explicit numbers were proved to belong to $\Pi_{\text{fin}}^{(k)}$. For example, before 2006 the only known members of $\Pi_{\text{fin}}^{(3)}$ were 0, 2/9, 4/9, 3/4, and 1 (see [4, 11, 23]). Then Mubayi [31] showed that $(m-1)(m-2)/m^2 \in \Pi_{\text{fin}}^{(3)}$ for every $m \geq 4$. Very recently, Baber and Talbot [2] and Falgas-Ravry and Vaughan [18] determined a few further elements of $\Pi_{\text{fin}}^{(3)}$; their proofs are computer-generated, being based on the flag algebra approach of Razborov [35]. In all the cases when an explicit element of $\Pi_{\text{fin}}^{(k)}$ is known, this limit density is achieved, informally speaking, by taking a finite pattern and blowing it up optimally. Here we generalise these results (as far as $\Pi_{\text{fin}}^{(k)}$ is concerned) by showing that *every* finite pattern where, moreover, we are allowed to iterate the whole construction recursively inside a specified set of parts, produces an element of $\Pi_{\text{fin}}^{(k)}$.

Let us give some formal definitions. (We refer the reader to Section 2.3 for an illustrative example.) A *pattern* is a triple $P = (m, E, R)$ where m is a positive integer, E is a collection of k-multisets on $[m] := \{1, \ldots, m\}$, and R is a subset of $[m]$. (By a *k-multiset* we mean an unordered collection of k elements with repetitions allowed.) Let $V_1, \ldots, V_m$ be disjoint sets and let $V = V_1 \cup \ldots \cup V_m$. The *profile* of a k-set $X \subseteq V$ (with respect to $V_1, \ldots, V_m$) is the k-multiset on $[m]$ that contains $i \in [m]$ with multiplicity $|X \cap V_i|$. For a k-multiset $Y \subseteq [m]$ let $Y(\!(V_1, \ldots, V_m)\!)$ consist of all k-subsets of V whose profile is Y. We call this k-graph the *blow-up of* Y and the k-graph

$$E(\!(V_1, \ldots, V_m)\!) := \bigcup_{Y \in E} Y(\!(V_1, \ldots, V_m)\!)$$

is called the *blow-up of* E (with respect to $V_1, \ldots, V_m$).

A *P-construction on a set* V is any k-graph G that can be recursively obtained as follows. Either let G be the empty k-graph on V (and stop) or take an arbitrary partition $V = V_1 \cup \ldots \cup V_m$ where we require that if $i \in R$ then $V_i \neq V$. Add all edges of $E(\!(V_1, \ldots, V_m)\!)$ to G. Furthermore, for every $i \in R$ take an arbitrary P-construction on V_i and add all these edges to G. (If $R = \emptyset$, then there is nothing to add and we have $G = E(\!(V_1, \ldots, V_m)\!)$.) Let p_n be the maximum number of edges that can be obtained on n vertices in this way:

$$p_n := \max\{\, |G| : G \text{ is a } P\text{-construction on } [n]\,\}. \tag{2}$$

It is not hard to show (see Lemma 10) that the ratio $p_n/\binom{n}{k}$ is non-increasing and therefore tends to a limit which we denote by Λ_P and call the *Lagrangian* of P:

$$\Lambda_P := \lim_{n\to\infty} \frac{p_n}{\binom{n}{k}}. \tag{3}$$

For $i \in [m]$ let $P - i$ be the pattern obtained from P by *removing index* i, that is, we remove i from R and delete all multisets containing i from E (and relabel the remaining indices to form the set $[m-1]$). In other words, $(P-i)$-constructions are precisely those P-constructions where we always let the i-th part be empty. Let us call P *minimal* if Λ_{P-i} is strictly smaller than Λ_P for every $i \in [m]$. For example, the 2-graph pattern $P := (3, \{\,\{1,2\},\{1,3\}\,\}, \emptyset)$ is not minimal as $\Lambda_P = \Lambda_{P-3} = 1/2$.

Theorem 3 *For every minimal pattern P there is a finite family $\mathcal{F}$ of k-graphs such that for all $n \geq 1$ we have $\mathrm{ex}(n, \mathcal{F}) = p_n$ and, moreover, every maximum $\mathcal{F}$-free k-graph on $[n]$ is a P-construction.*

Corollary 4 *For every pattern P we have that $\Lambda_P \in \Pi^{(k)}_{\mathrm{fin}}$.* ∎

Corollary 4 answers questions posed by Baber and Talbot [2, Question 29] and by Falgas-Ravry and Vaughan [18, Question 4.4]; we refer the reader to Section 7 for details.

Chung and Graham [10, Page 95] conjectured that $\Pi^{(k)}_{\mathrm{fin}}$ consists of rational numbers only. The following theorem disproves this conjecture for every $k \geq 3$. (Note that the conjecture is true for $k = 2$ by (1).) Independently, Chung and Graham's conjecture was disproved by Baber and Talbot [2] who discovered a family of only three forbidden 3-graphs whose Turán density is irrational. We should mention that Theorems 3 and 5 rely on the Strong Removal Lemma of Rödl and Schacht [37] so they give families $\mathcal{F}$ of **huge** size.

Theorem 5 *For every $k \geq 3$ the set $\Pi^{(k)}_{\mathrm{fin}}$ contains an irrational number.*

This paper is organised as follows. Some further notation is given in Section 2. The proof of Theorem 3 is presented in Section 3; it is preceded by a number of auxiliary results. Sections 4, 5, and 6 contain the proofs of respectively Theorem 5, Proposition 1, and Theorem 2. Finally, Section 7 presents some concluding remarks and open questions.

2 Notation

Let us introduce some further notation complementing and expanding that from the Introduction. Some other (infrequently used) definitions are given shortly before they are needed for the first time in this paper.

Recall that a *k-multiset* D is an unordered collection of k elements $x_1, \ldots, x_k$ with repetitions allowed. Let us denote this as $D = \{\!\{\, x_1, \ldots, x_k \,\}\!\}$. The *multiplicity* $D(x)$ of x in D is the number of times that x appears. If the underlying set is understood to be $[m]$, then we can represent D as the ordered m-tuple $(D(1), \ldots, D(m))$ of multiplicities. Thus, for example, the profile of $X \subseteq V_1 \cup \ldots \cup V_m$ is the multiset on $[m]$ whose multiplicities are $(|X \cap V_1|, \ldots, |X \cap V_m|)$. Also, let

$x^{(r)}$ denote the sequence consisting of r copies of x; thus the multiset consisting of r copies of x is denoted by $\{\!\{ x^{(r)} \}\!\}$. If we need to emphasise that a multiset is in fact a set (that is, no element has multiplicity more than 1), we call it a *simple set*.

For $D \subseteq [m]$ and sets $U_1, \ldots, U_m$, denote $U_D := \cup_{i \in D} U_i$. Let $\binom{X}{m} := \{Y \subseteq X : |Y| = m\}$ consist of all m-subsets of a set X. The *standard* $(m-1)$*-dimensional simplex* is

$$\mathbb{S}_m := \{\mathbf{x} \in \mathbb{R}^m : x_1 + \ldots + x_m = 1,\ \forall\, i \in [m]\ x_i \geq 0\}. \tag{4}$$

2.1 Hypergraphs

We usually identify a k-graph G with its edge set. For example, $X \in G$ means that X is an edge of G and $|G|$ denotes the number of edges. When we need to refer to the vertex set, we write $V(G)$ and denote $v(G) := |V(G)|$. The *(edge) density* of G is

$$\rho(G) := \frac{|G|}{\binom{v(G)}{k}}.$$

The *complement of* G is $\overline{G} := \{X \subseteq V(G) : |X| = k,\ X \notin G\}$. For $x \in V(G)$ its *link* is the $(k-1)$-hypergraph

$$G_x := \{X \subseteq V(G) : x \notin X,\ X \cup \{x\} \in G\}.$$

For $U \subseteq V(G)$ its *induced subgraph* is $G[U] := \{X \in G : X \subseteq U\}$. The vertex sets of $\overline{G}$, G_x, and $G[U]$ are by default $V(G)$, $V(G) \setminus \{x\}$, and U respectively. The *degree* of $x \in V(G)$ is $d_G(x) := |G_x|$. Let $\Delta(G)$ and $\delta(G)$ denote respectively the maximum and minimum degrees of the k-graph G.

An *embedding* of a k-graph F into G is an injection $f : V(F) \to V(G)$ such that $f(X) \in G$ for every $X \in F$. An embedding is *induced* if non-edges are mapped to non-edges.

2.2 Pattern Specific Definitions

Let $P = (m, E, R)$ be a pattern and G be a P-construction on $[n]$. The initial partition $V(G) = V_1 \cup \ldots \cup V_m$ is called the *level*-1 *partition* and V_i's are called *level*-1 *parts*. For each $i \in R$ we denote the corresponding partition of V_i as $V_{i,1} \cup \ldots \cup V_{i,m}$ and call these parts *level*-2 *parts*. This notation generalises in the obvious way with $V_{i_1,\ldots,i_s}$ for $(i_1, \ldots, i_s) \in R^{s-1} \times [m]$ consisting of those vertices of G that, for every $j = 1, \ldots, s$, belong to the i_j-th part on level j. Also, we denote $V_\emptyset := V(G)$.

The *length* of a sequence $\mathbf{i} = (i_1, \ldots, i_s)$ is $|\mathbf{i}| = s$. The sequence $\mathbf{i}$ is *legal* if $i_j \in R$ for all $j \in [s-1]$ and $i_s \in [m]$; this includes the empty sequence.

We collect all parts that appear in the P-construction G into a single vector

$$\mathbf{V} := (V_\emptyset, V_1, \ldots, V_m, \ldots)$$

and call $\mathbf{V}$ the *partition structure* of G; its index set is some subset of legal sequences.

For convenience, we view the partition structure as vertical with a level's index (called *height*) increasing as we go up. In particular, the partition $V_1 \cup \ldots \cup V_m$ is called *bottom*. By default, the profile of $X \subseteq V(G)$ is taken with respect to the bottom parts, that is, its multiplicities are $(|X \cap V_1|, \ldots, |X \cap V_m|)$. The *branch* $\mathrm{br}_{\mathbf{V}}(x)$ of a vertex $x \in V(G)$ is the (unique) maximal sequence $\mathbf{i}$ such that $x \in V_{\mathbf{i}}$.

Given P, let $\mathcal{F}_\infty$ consist of those k-graphs F that do not embed into a P-construction:

$$\mathcal{F}_\infty := \{k\text{-graph } F : \text{every } P\text{-construction } G \text{ is } F\text{-free}\}. \tag{5}$$

For an integer n, let $\mathcal{F}_n$ consist of all members of $\mathcal{F}_\infty$ with at most n vertices:

$$\mathcal{F}_n := \{F \in \mathcal{F}_\infty : v(G) \le n\}. \tag{6}$$

Let the *Lagrange polynomial* of E be

$$\lambda_E(x_1, \ldots, x_m) := k! \sum_{D \in E} \prod_{i=1}^{m} \frac{x_i^{D(i)}}{D(i)!}. \tag{7}$$

This definition is motivated by the fact that, for every partition $[n] = V_1 \cup \ldots \cup V_m$ we have that

$$\rho(E(\!(V_1, \ldots, V_m)\!)) = \lambda_E\left(\frac{|V_1|}{n}, \ldots, \frac{|V_m|}{n}\right) + o(1), \qquad \text{as } n \to \infty; \tag{8}$$

see also Lemma 14 that relates λ_E and Λ_P. The special case of (7) when E is a k-graph (i.e. E consists of simple sets) has been successfully applied to Turán-type problems, with the basic idea going back to Motzkin and Straus [30]. Also, our definition of Λ_P is a generalisation of the well-known *hypergraph Lagrangian* $\Lambda_E := \Lambda_{(m,E,\emptyset)}$, see e.g. [2].

For $i \in [m]$ let the *link* E_i consist of all $(k-1)$-multisets A such that if we increase the multiplicity of i in A by one, then the obtained k-multiset belongs to E. We call a pattern P *proper* if it is minimal and $0 < \Lambda_P < 1$. Trivially, every minimal pattern $P = (m, E, R)$ satisfies that

$$E_i \neq \emptyset, \qquad \text{for every } i \in [m]. \tag{9}$$

2.3 An Example

To illustrate the above definitions, let us consider a specific simple example:

$$P := (\,2, \{\, \{\!\{1, 2, 2\}\!\}\, \}, \{1\}\,). \tag{10}$$

Here a P-construction on V is obtained by partitioning $V = V_1 \cup V_2$ with $V_1 \neq V$ and adding all triples that have exactly two vertices in V_2. Next, we apply recursion to V_1: namely, we partition $V_1 = V_{1,1} \cup V_{1,2}$ with $V_{1,1} \neq V_1$ and add all triples that intersect $V_{1,1}$ and $V_{1,2}$ in respectively one and two vertices. Next, we repeat inside $V_{1,1}$, and so on. We can always stop; for example, we may choose to do this after three iterations by letting $V_{1,1,1}$ span the empty 3-graph. In this case, the partition structure is

$$\mathbf{V} = (V_\emptyset, V_1, V_2, V_{1,1}, V_{1,2}, V_{1,1,1}, V_{1,1,2}),$$

where $V_\emptyset := V$. If we take a vertex x in V_2, $V_{1,2}$, $V_{1,1,1}$, and $V_{1,1,2}$ then its branch is respectively (2), $(1,2)$, $(1,1,1)$, and $(1,1,2)$. This defines the branch of every vertex as these four sets partition V; there is no vertex whose branch is, for example, $(1,1)$.

We have $\lambda_E(x_1, x_2) = 6 \cdot x_1 \cdot (x_2^2/2) = 3x_1x_2^2$. It is not hard to show (cf Lemma 14) that $\Lambda_P = 2\sqrt{3} - 3$ and an example of a P-construction attaining this density is to use a ratio close to $1 : \sqrt{3}$ for each partition. Thus, this is an example of a 3-graph pattern whose Lagrangian is irrational.

3 Proof of Theorem 3

The proof of Theorem 3 is rather long and relies on a number of auxiliary results. Very briefly, it proceeds as follows. The starting point is the easy observation (Lemma 7) that by forbidding $\mathcal{F}_\infty$ we restrict ourselves to k-graphs that embed into a P-construction; thus $\mathrm{ex}(n, \mathcal{F}_\infty) = p_n$. The deep and powerful Strong Removal Lemma of Rödl and Schacht [37] (stated as Lemma 21 here) implies that for every $\varepsilon > 0$ there is M such that every $\mathcal{F}_M$-free k-graph with $n \geq M$ vertices can be made $\mathcal{F}_\infty$-free by removing at most $\varepsilon\binom{n}{k}$ edges. It follows that every maximum $\mathcal{F}_M$-free k-graph G on $[n]$ is $2\varepsilon\binom{n}{k}$-close in the edit distance to a P-construction, see Lemma 22. Although the obtained $\varepsilon > 0$ can be made arbitrarily small by choosing M large, the author did not see any simple way of ensuring that $\varepsilon \to 0$ for some fixed M as $n \to \infty$. Nonetheless our key Lemma 20 shows that some small but *constant* $\varepsilon > 0$ suffices to ensure that there is a partition $V(G) = V_1 \cup \ldots \cup V_m$ such that $G \setminus (\cup_{i\in R} G[V_i]) = E(\!(V_1, \ldots, V_m)\!)$, that is, G follows *exactly* the bottom level of some P-construction (but nothing is stipulated about what happens inside the "recursive" parts V_i). The maximality of G implies that each $G[V_i]$ with $i \in R$ is maximum $\mathcal{F}_M$-free (cf Lemma 9), allowing us to apply induction.

3.1 Basic Properties of Patterns

Here, let $P = (m, E, R)$ be an arbitrary pattern and let all definitions of Sections 1 and 2 apply. In particular, p_n, Λ_P, $\mathcal{F}_\infty$ and $\mathcal{F}_n$ are defined by respectively (2), (3), (5), and (6).

Lemma 6 *Any induced subgraph (resp. any blow-up) of a P-construction H is (resp. embeds into) a P-construction.*

Proof. Let $\mathbf{V}$ be the partition structure of H. If $H' := H[X]$ is an induced subgraph, then we can initially let $V'_{\mathbf{i}} := V_{\mathbf{i}} \cap X$ for each index $\mathbf{i}$. This need not be a partition structure as we may have $V'_{i_1,\ldots,i_s} = V'_{i_1,\ldots,i_{s-1}}$ for some $(i_1, \ldots, i_s) \in R^s$, which is not allowed by the definition:

namely, the partition of $V'_{i_1,\ldots,i_{s-1}}$ has the i_s-th part equal to the whole set (and $i_s \in R$). We can fix one such occurrence by removing i_s from all indices that begin with $(i_1,\ldots,i_s)$. Formally, we remove all parts $V'_{i_i,\ldots,i_{s-1},j_s,\ldots}$ with $j_s \neq i_s$ (note that they are all empty) and relabel each part $V'_{i_1,\ldots,i_s,j_{s+1},\ldots,j_t}$ into $V'_{i_1,\ldots,i_{s-1},j_{s+1},\ldots,j_t}$. We keep fixing all such occurences one by one. Since, for example, $\sum_{V'_{\mathbf{i}}\in\mathbf{V}'} |\mathbf{i}|$ strictly decreases each time, this procedure stops. The final vector $\mathbf{V}'$ shows that H' is a P-construction.

If we insert a new vertex into a P-construction by putting it into the same part as some existing vertex x, then we add all those edges (and possibly some further ones) as when we just clone x. Thus every blow-up of H, which can be obtained by a sequence of cloning steps and vertex removals, embeds into a P-construction. ∎

Lemma 7 *The following are equivalent for an arbitrary k-graph G on n vertices: 1) G is $\mathcal{F}_n$-free; 2) G is $\mathcal{F}_\infty$-free; 3) G embeds into a P-construction; 4) G embeds into a P-construction H with $v(H) = n$.*

Proof. The equivalence of 1), 2), and 3) follows from the definitions of $\mathcal{F}_n$ and $\mathcal{F}_\infty$. Statements 3) and 4) are equivalent by Lemma 6. ∎

It follows from Lemma 7 that $\mathrm{ex}(n, \mathcal{F}_n) = \mathrm{ex}(n, \mathcal{F}_\infty) = p_n$.

Lemma 8 *Let $s \in \mathbb{N} \cup \{\infty\}$. If G is $\mathcal{F}_s$-free, then any blow-up of G is $\mathcal{F}_s$-free.*

Proof. Let G' be obtained from G by adding a clone x' of some vertex x of G. Take any $U \subseteq V(G')$ with $|U| \leq s$. If at least one of x and x' is not in U, then $G'[U]$ is isomorphic to a subgraph of G and cannot be in $\mathcal{F}_s$; so suppose otherwise. Since G is $\mathcal{F}_s$-free, there is an embedding f of $G[U \setminus \{x'\}]$ into some P-construction. By Lemma 6, $G[U]$ is also embeddable. It follows that G' is $\mathcal{F}_s$-free. ∎

Lemma 9 *Let $s \in \mathbb{N} \cup \{\infty\}$. Let G be a k-graph on $V = V_1 \cup \ldots \cup V_m$ obtained by taking $E(\!(V_1,\ldots,V_m)\!)$ and putting arbitrary $\mathcal{F}_s$-free k-graphs into parts V_i with $i \in R$. Then G is $\mathcal{F}_s$-free.*

Proof. Take an arbitrary $U \subseteq V(G)$ with $|U| \leq s$. Let $U_i := V_i \cap U$. Note that $G[U_i]$ has no edges for $i \in [m] \setminus R$ and embeds into some P-construction H_i for $i \in R$ (because $|U_i| \leq s$ and $G[U_i] \subseteq G[V_i]$ is $\mathcal{F}_s$-free). By combining the partition structure of each H_i together with the level-1 decomposition $U = U_1 \cup \ldots \cup U_m$, we see that $G[U]$ embeds into a P-construction, giving the required. ∎

Lemma 10 *The ratio $p_n/\binom{n}{k}$ is non-increasing with n. In particular, the limit in (3) exists.*

Proof. Let $\ell < n$ and take a maximum P-construction G on $[n]$. Every ℓ-subset of $[n]$ spans at most p_ℓ edges by Lemma 6. Averaging over all $\binom{n}{\ell}$ ℓ-subsets gives that $p_n \le p_\ell\binom{n}{\ell}/\binom{n-k}{\ell-k} = p_\ell\binom{n}{k}/\binom{\ell}{k}$, as required. ∎

Lemma 11 *For every $\varepsilon > 0$ and $s \in \mathbb{N}\cup\{\infty\}$ there is n_0 such that every maximum $\mathcal{F}_s$-free k-graph G with $n \ge n_0$ vertices has minimum degree at least $(\Lambda_P - \varepsilon)\binom{n-1}{k-1}$.*

Proof. Let n be large and G be as stated. Clearly, $|G| \ge p_n$. The average degree of G is $k|G|/n \ge kp_n/n \ge (\Lambda_P - \varepsilon/2)\binom{n-1}{k-1}$. If some x has degree smaller than $(\Lambda_P - \varepsilon)\binom{n-1}{k-1}$, then by deleting x and adding a clone y' of a vertex y whose degree is at least the average, we increase $|G|$ by at least $|G_y| - |G_x| - \binom{n-2}{k-2} > 0$. This preserves the $\mathcal{F}_s$-freeness by Lemma 8, contradicting the maximality of G. ∎

Lemma 12 *We have $\Lambda_P = 1$ if and only if at least one of the following holds.*

1. *There is $i \in [m]$ such that $\{\!\{\, i^{(k)} \,\}\!\} \in E$;*
2. *There are $i \in R$ and $j \in [m] \setminus \{i\}$ such that $\{\!\{\, i^{(k-1)}, j \,\}\!\} \in E$.*

Proof. The converse implication is obvious: we can get the complete k-graph on $[n]$ by taking $V_i = [n]$ in the first case and by taking $V_i = [n-1]$, $V_j = \{n\}$, and recursing inside V_i in the second case.

Let us show the direct implication. Suppose that the above multisets are not present in E. Let $n \to \infty$ and let G be a maximum P-construction on $[n]$ with the bottom partition $[n] = V_1 \cup \ldots \cup V_m$.

Suppose first that there is a part V_i with $n - o(n)$ vertices for infinitely many n, say $i = 1$. Assume that $1 \in R$ for otherwise the complement $\overline{G}$ has at least $\binom{|V_1|}{k} = \Omega(n^k)$ edges. Since V_1 is not allowed to be the whole vertex set $[n]$, we can assume that e.g. $V_2 \ne \emptyset$. Fix $x \in V_2$. The degree of x in G is at most $(n - |V_1|)\binom{n}{k-2} = o(n^{k-1})$: since $\{\!\{\, 1^{(k-1)}, 2 \,\}\!\} \notin E$, each edge of the link $(k-1)$-graph G_x has to contain at least one vertex outside of V_1. This contradicts Lemma 11.

Thus some two parts, say V_1 and V_2, have $\Omega(n)$ vertices each. Assume that $1 \in R$ for otherwise at least $\Omega(n^k)$ edges (those inside V_1) are missing from G. Since $\{\!\{\, 1^{(k-1)}, 2 \,\}\!\} \notin E$, all edges that intersect V_1 in $k-1$ vertices and V_2 in one vertex are not present. Again, at least $\Omega(n^k)$ edges are missing from G, as required. ∎

The proof of Lemma 12 shows that if $\Lambda_P = 1$, then the complete k-graph is a P-construction. This satisfies Theorem 3 if we take $\mathcal{F} = \emptyset$. Also, if $\Lambda_P = 0$, then only empty k-graphs are realisable as P-constructions and Theorem 3 is also satisfied: let $\mathcal{F} = \{K_k^k\}$ consist of a single edge. Thus it is enough to prove Theorem 3 for proper patterns (that is, minimal patterns with Lagrangian strictly between 0 and 1).

3.2 Properties of Proper Patterns

In this section we let $P = (m, E, R)$ be an arbitrary pattern that is proper. Here we establish some properties of P.

Lemma 13 *For every P-construction G on n vertices with minimum degree $\delta(G) = \Omega(n^{k-1})$, each bottom part V_i has at most $(1 - \Omega(1))n$ vertices as $n \to \infty$.*

Proof. For $i \in [m] \setminus R$ the claim follows from $\delta(G)n/k \le |G| \le \binom{n}{k} - \binom{|V_i|}{k}$. Let $i \in R$. Since $V_i \neq V(G)$, pick any vertex $x \in V_j$ with $j \neq i$. Since $\{\!\{ i^{(k-1)}, j \}\!\} \notin E$ by Lemma 12, every edge through x contains at least one other vertex outside of V_i. Thus $d_G(x) \le (n - |V_i|)\binom{n-2}{k-2}$, implying the required. ∎

Let

$$\mathbb{S}_m^* := \{\mathbf{x} \in \mathbb{R}^m : x_1 + \ldots + x_m = 1,\ \forall\, i \in [m]\ 0 \le x_i < 1\}$$

be obtained from $\mathbb{S}_m$ by excluding the standard basis vectors, where $\mathbb{S}_m$ is defined by (4). Let us call a vector $\mathbf{x} \in \mathbb{R}^m$ *optimal* if $\mathbf{x} \in \mathbb{S}_m^*$ and

$$\Lambda_P = \lambda_E(\mathbf{x}) + \Lambda_P \sum_{i \in R} x_i^k. \tag{11}$$

Let $\mathcal{X}$ be the set of all optimal $\mathbf{x}$. Note that when we define $\mathcal{X}$ we restrict ourselves to $\mathbb{S}_m^*$ (i.e. we do not allow any standard basis vector to be included into $\mathcal{X}$).

In a sense (with the formal statements appearing in Lemmas 14 and 15 below), $\mathcal{X}$ is precisely the set of optimal limiting ratios that lead to asymptotically maximum P-constructions. Let us illustrate this on the case when P is as in (10). Suppose that we want to determine Λ_P. Let $[n] = V_1 \cup V_2$ be the bottom partition in a maximum P-construction G. Let $x_i := |V_i|/n$ for $i = 1, 2$. By Lemma 10, $\rho(G)$ and $\rho(G[V_1])$ are close to Λ_P. (Note that we cannot have $x_1 = o(1)$ by Lemmas 11 and 13.) Thus, we conclude that $\Lambda_P = 3x_1x_2^2 + \Lambda_P x_1^3 + o(1)$, which is exactly (11) if we ignore the error term. Solving for Λ_P and excluding x_2, we have to maximise $g(x) := 3x(1-x)^2/(1-x^3)$ for $x \in (0, 1)$. In this particular case, the maximum is $2\sqrt{3} - 3$ and it is attained inside $(0, 1)$ at the unique point $\alpha := (\sqrt{3}-1)/2$. It follows that $\Lambda_P = 2\sqrt{3}-3$, (11) has a unique solution in $\mathbb{S}_2^*$, and $\mathcal{X} = \{(\alpha, 1-\alpha)\}$. Note that although $(x_1, x_2) = (1, 0)$ satisfies (11), we have that $\lim_{x \to 1} g(x) = 0 < \Lambda_P$. This justifies why we exclude the standard basis vectors from $\mathcal{X}$.

Lemma 14 *Let $f(\mathbf{x}) := \lambda_E(\mathbf{x}) + \Lambda_P \sum_{i \in R} x_i^k$ be the right-hand side of (11). Then the following claims hold.*

1. *$\mathcal{X} \neq \emptyset$.*

2. *$f(\mathbf{x}) \le \Lambda_P$ for all $\mathbf{x} \in \mathbb{S}_m$. (Thus, by Part 1, $\mathcal{X}$ is precisely the set of elements in $\mathbb{S}_m^*$ that maximise f.)*

3. $\mathcal{X}$ does not intersect the boundary of $\mathbb{S}_m$.

4. For every $\mathbf{x} \in \mathcal{X}$ and $j \in [m]$ we have $\frac{\partial f}{\partial_j}(\mathbf{x}) = k\Lambda_P$.

5. $\mathcal{X}$ is a closed subset of $\mathbb{S}_m$.

6. For every $\varepsilon > 0$ there is $\alpha > 0$ such that for every $\mathbf{y} \in \mathbb{S}_m$ with $\max(y_1, \ldots, y_m) \le 1 - \varepsilon$ and $f(\mathbf{y}) \ge \Lambda_P - \alpha$ there is $\mathbf{x} \in \mathcal{X}$ with $\|\mathbf{x} - \mathbf{y}\|_\infty \le \varepsilon$.

7. There is $\beta > 0$ such that for every $\mathbf{x} \in \mathcal{X}$ and every $i \in [m]$ we have $x_i \ge \beta$.

Proof. Let G be a maximum P-construction on $[n]$ with the bottom partition $V_1 \cup \ldots \cup V_m$. By passing to a subsequence of n, we can assume that, for every $i \in [m]$, the ratio $|V_i|/n$ tends to some limit x_i. By Lemmas 11 and 13, $\mathbf{x} = (x_1, \ldots, x_m)$ belongs to $\mathbb{S}_m^*$. By Lemma 9, for each $i \in R$ the induced subgraph $G[V_i]$ is a maximum P-construction. By Lemma 10, we have that $|G[V_i]| = (\Lambda_P x_i^k + o(1))\binom{n}{k}$. Now, (8) shows that $\mathbf{x}$ satisfies (11). Thus $\mathbf{x} \in \mathcal{X}$, so this set is non-empty.

Let $\mathbf{x} \in \mathbb{S}_m$. If we use the approximate ratios $x_1 : \ldots : x_m$ for the bottom partition $V_1 \cup \ldots \cup V_m$ and put a maximum P-construction on each V_i with $i \in R$, then the obtained P-construction has edge density $f(\mathbf{x}) + o(1)$. Thus $f(\mathbf{x}) \le \Lambda_P$ for all $\mathbf{x} \in \mathbb{S}_m$, proving Part 2.

Suppose that $\mathcal{X}$ intersects the boundary of $\mathbb{S}_m$, that is, $\mathcal{X}$ contains some $\mathbf{x}$ with zero entries. Without loss of generality, assume that $x_1, \ldots, x_{m'}$ are positive while all other entries are 0. Since $\mathbf{x} \in \mathbb{S}_m^*$, we have $m' \ge 2$. Let $P' = (m', E', R')$ be obtained from P by removing the indices $m' + 1, \ldots, m$. Consider a P'-construction H where the bottom partition $U_1 \cup \ldots \cup U_{m'}$ has approximate ratios $x_1 : \ldots : x_{m'}$ while each part U_i with $i \in R'$ spans a maximum P'-construction. By the definition of $\Lambda_{P'}$, we have

$$\begin{aligned} \Lambda_{P'} \;\ge\; \rho(H) + o(1) \;&=\; \lambda_{E'}(x_1, \ldots, x_{m'}) + \Lambda_{P'} \sum_{i \in R'} x_i^k + o(1) \\ &=\; \lambda_E(\mathbf{x}) + \Lambda_{P'} \sum_{i \in R} x_i^k + o(1). \end{aligned}$$

Since $\mathbf{x} \in \mathbb{S}_m^*$, we have $\sum_{i \in R} x_i^k < 1$. Thus $\Lambda_{P'} + o(1) \ge \lambda_E(\mathbf{x})/(1 - \sum_{i \in R} x_i^k) = \Lambda_P$, where we used identity (11) that $\mathbf{x} \in \mathcal{X}$ has to satisfy. This contradicts the minimality of P and proves Part 3.

Let $\mathbf{x} \in \mathcal{X}$. By Part 3, $\mathbf{x}$ lies in the interior of $\mathbb{S}_m$. Since $\mathbf{x}$ maximises f subject to $x_1 + \ldots + x_m = 1$, we conclude that all partial derivatives of f coincide at $\mathbf{x}$. Furthermore, this common value is $k\Lambda_P$, which follows from the easy identity $\sum_{i=1}^m x_i \frac{\partial f}{\partial_i}(\mathbf{x}) = kf(\mathbf{x})$, establishing Part 4

From Part 2 we know that $\mathcal{X}$ is precisely the set of elements of $\mathbb{S}_m^*$ that maximise $f(\mathbf{x})$. Clearly, $f : \mathbb{S}_m \to \mathbb{R}$ is a continuous function. Thus, in order to prove Part 5 it is enough to show that $\mathcal{X}$ cannot accumulate to any element of the set $\mathbb{S}_m \setminus \mathbb{S}_m^*$ that consists of the standard basis vectors. The proof will essentially be a translation of the argument of Lemma 13 into a more analytic language. Let $\mathbf{x} \in \mathcal{X}$. Take any index $i \in [m]$. If $i \in [m] \setminus R$, then $\Lambda_P = f(\mathbf{x}) \le 1 - x_i^k + \Lambda_P(1 - x_i)^k$, so x_i

cannot be arbitrarily close to 1. Let $i \in R$. Since P is proper, each monomial of $\lambda_E(\mathbf{x})$ contains at least two factors different from x_i by Lemma 12. Thus when we take the j-th derivative of f for $j \in [m]$, each monomial will have some factor x_s with $s \neq i$; of course, $x_s \leq 1 - x_i$. As the sum of the coefficients of the degree-k polynomial f is, rather roughly, at most m^k, we conclude that $\frac{\partial f}{\partial_j}(\mathbf{x}) \leq km^k(1 - x_i)$. By Part 4 we conclude that $1 - x_i \geq \Lambda_P/m^k$, that is, x_i is separated from 1. This establishes Part 5.

Suppose that Part 6 is false. Then there is $\varepsilon > 0$ such that for every $i \in \mathbb{N}$ there is $\mathbf{y}_i \in \mathbb{S}_m$ violating it with $\alpha = 1/i$. By the compactness of $\mathbb{S}_m$ the sequence $(\mathbf{y}_1, \mathbf{y}_2, \ldots)$ accumulates to some $\mathbf{y}$. The vector $\mathbf{y}$ belongs to $\mathbb{S}_m^*$ by the assumption on each $\mathbf{y}_i$. By the continuity of f we have $f(\mathbf{y}) \geq \Lambda_P$, that is, $\mathbf{y} \in \mathcal{X}$, a contradiction to $\mathbf{y}$ being ε-far from $\mathcal{X}$.

Part 7 is proved in a similar way as Part 6. (Alternatively, it directly follows from Parts 3 and 5.) ∎

Informally speaking, the following lemma implies, among other things, that all part ratios of bounded height in a P-construction of large minimum degree approximately follow some optimal vectors. For example, if P is defined by (10), then Part 1 of Lemma 15 gives that, for any fixed ℓ, every P-construction G on $[n]$ with $\delta(G) \geq (\Lambda_P - o(1))\binom{n-1}{k-1}$ satisfies $|V_{1^{(s)},2}| : |V_{1^{(s)},1}| = \sqrt{3} + o(1)$ for each $0 \leq s \leq \ell$.

Lemma 15 *For every $\varepsilon > 0$ and $\ell \in \mathbb{N}$ there are constants $\alpha_0, \varepsilon_0, \ldots, \alpha_\ell, \varepsilon_\ell, \alpha_{\ell+1} \in (0, \varepsilon)$ and $n_0 \in \mathbb{N}$ such that the following holds. Let G be an arbitrary P-construction G on $n \geq n_0$ vertices with the partition structure $\mathbf{V}$ such that the minimal degree $\delta(G) \geq (\Lambda_P - \alpha_0)\binom{n-1}{k-1}$. Take arbitrary $\mathbf{i} \in R^s$ with $0 \leq s \leq \ell$ and denote $\mathbf{v}_\mathbf{i} := (|V_{\mathbf{i},1}|/|V_\mathbf{i}|, \ldots, |V_{\mathbf{i},m}|/|V_\mathbf{i}|)$. Then:*

1. *$\|\mathbf{v}_\mathbf{i} - \mathbf{x}\|_\infty \leq \varepsilon_s$ for some $\mathbf{x} \in \mathcal{X}$;*
2. *$|V_{\mathbf{i},j}| \geq (\beta/2)^{s+1} n$ for all $j \in [m]$, where β is returned by Part 7 of Lemma 14;*
3. *$\delta(G[V_{\mathbf{i},j}]) \geq (\Lambda_P - \alpha_{s+1})\binom{|V_{\mathbf{i},j}|-1}{k-1}$ for all $j \in R$.*

Proof. We choose positive constants in this order

$$\alpha_{\ell+1} \gg \varepsilon_\ell \gg \alpha_\ell \gg \ldots \gg \varepsilon_0 \gg \alpha_0 \gg 1/n_0,$$

each being sufficiently small depending on P, ε, β, and the previous constants. Let G and $\mathbf{i}$ be as in the lemma. We use induction on $s = 0, 1, \ldots, \ell$. Let $U_j := V_{\mathbf{i},j}$ for $j \in [m]$, $U := V_\mathbf{i} = U_1 \cup \ldots \cup U_m$, and $\mathbf{u} := \mathbf{v}_\mathbf{i} = (|U_1|/|U|, \ldots, |U_m|/|U|)$. By the inductive assumption on $\delta(G[U])$ (or by the assumption of the lemma if $s = 0$ when $U = V_\emptyset = V(G)$), we have that

$$|G[U]| \geq \delta(G[U])|U|/k \geq (\Lambda_P - \alpha_s)\binom{|U|}{k}.$$

Since $\alpha_s \ll \varepsilon_s$, Lemma 13 and Part 6 of Lemma 14 (when applied to $\mathbf{y} = \mathbf{u}$) give the desired $\mathbf{x}$, proving Part 1.

For all $j \in [m]$, we have $|U_j| \geq (x_j - \varepsilon_s)|U| \geq (\beta/2)|U|$, which is at least $(\beta/2)^{s+1}n$ by the inductive assumption, proving Part 2.

Finally, take arbitrary $j \in R$ and $y \in U_j$. The degree of y in $E(\!(U_1, \ldots, U_m)\!)$ is

$$d_{E(\!(U_1,\ldots,U_m)\!)}(y) = \left(\frac{1}{k} \times \frac{\partial \lambda_E}{\partial_j}(\mathbf{u}) + o(1)\right)\binom{|U|-1}{k-1}.$$

Since $\|\mathbf{u} - \mathbf{x}\|_\infty \leq \varepsilon_s \ll \alpha_{s+1}$, we have by Part 4 of Lemma 14 that, for example,

$$\frac{\partial}{\partial_j}\lambda_E(\mathbf{u}) - \alpha_{s+1}^2 \leq \frac{\partial}{\partial_j}\lambda_E(\mathbf{x}) = \frac{\partial}{\partial_j}f(\mathbf{x}) - k\Lambda_P x_j^{k-1} = k\Lambda_P - k\Lambda_P x_j^{k-1}.$$

Thus, by the (inductive) assumption on the minimal degree of $G[U]$, we have

$$\begin{aligned}
d_{G[U_j]}(y) &= d_{G[U]}(y) - d_{E(\!(U_1,\ldots,U_m)\!)}(y) \\
&\geq \left((\Lambda_P - \alpha_s) - (\Lambda_P - \Lambda_P x_j^{k-1} + 2\alpha_{s+1}^2/k)\right)\binom{|U|-1}{k-1} \\
&\geq \left(\Lambda_P(u_j - \varepsilon_s)^{k-1} - 3\alpha_{s+1}^2/k\right)\binom{|U|-1}{k-1} \;\geq\; (\Lambda_P - \alpha_{s+1})\binom{|U_j|-1}{k-1},
\end{aligned}$$

where we used $|u_j - x_j| \leq \varepsilon_s$ and $x_j \geq \beta \gg \alpha_{s+1} \gg \varepsilon_s \gg \alpha_s$. This finishes the proof of Lemma 15. ∎

Recall that the link E_i of $i \in [m]$ consists of all $(k-1)$-multisets on $[m]$ such that if we increase the multplicity of i by one, then the obtained k-multiset belongs to E.

Lemma 16 *If distinct $i, j \in [m]$ satisfy $E_i \subseteq E_j$, then $i \in R$, $j \notin R$, and $E_i \neq E_j$. In particular, no two vertices of the pattern $P = (m, E, R)$ have the same links in E.*

Proof. Take some optimal $\mathbf{x} \in \mathcal{X}$. By Part 3 of Lemma 14, all coordinates of $\mathbf{x}$ are non-zero. Define $\mathbf{x}' \in \mathbb{S}_m$ by $x'_i = 0$, $x'_j = x_i + x_j$, and $x'_h = x_h$ for all other indices h. We claim that

$$\lambda_E(\mathbf{x}') \geq \lambda_E(\mathbf{x}). \tag{12}$$

One way to show (12) is to use (8). Consider some $F := E(\!(V_1, \ldots, V_m)\!)$. The assumption $E_i \subseteq E_j$ implies that if decrease the multiplicity of i in some $A \in E$ but increase the multiplicity of j by the same amount, then the new multiset necessarily belongs to E. Thus if we remove a vertex y from V_i and add a vertex y' to V_j, then the obtained k-graph F' has at least as many edges as F. (In fact, we have that $F_y \subseteq F'_{y'}$.) Since $\mathbf{x}'$ is obtained from $\mathbf{x}$ by shifting weight from x_i to x_j, (12) follows.

Also, $m \geq 3$ for otherwise $\{\!\{ j^{(k)} \}\!\} \in E$, contradicting Lemma 12. Thus $\mathbf{x}' \in \mathbb{S}_m^*$.

We conclude that $j \notin R$ for otherwise we get a contradiction to Part 2 of Lemma 14 by using (12) and the trivial inequality $x_i^k + x_j^k < (x_i + x_j)^k$. Likewise, $i \in R$ for otherwise the vector $\mathbf{x}'$, that has a zero coordinate, would belong to $\mathcal{X}$. Finally, we see that $E_i \neq E_j$ by swapping the roles of i and j in the above argument. ∎

A map $h : [m] \to [m]$ is an *automorphism* of the pattern P if h is bijective, $h(R) = R$, and h is an automorphism of E (that is, $h(E) = E$). Let us call a P-construction G with the bottom partition $V_1 \cup \ldots \cup V_m$ *rigid* if for every embedding f of G into a P-construction H with the bottom partition $U_1 \cup \ldots \cup U_m$ such that $f(V(G))$ intersects at least two different parts U_i, there is an automorphism h of P such that $f(V_i) \subseteq U_{h(i)}$ for every $i \in [m]$.

For example, the pattern P in (10) has no non-trivial automorphism and a rigid P-construction can be obtained by taking any $E(\!(V_1, V_2)\!)$ with $|V_1| \geq 1$ and $|V_2| \geq 3$. Thus the following Lemma 17 is trivially true for this particular P.

Lemma 17 *For all large n, every maximum P-construction G on $[n]$ is rigid.*

Since the proof of Lemma 17 in general is long and complicated, some informal discussion may be helpful here. It is not surprising that the proof is far simpler if $R = \emptyset$. In fact, an example of a rigid P-construction in this case can be obtained by letting each V_i have more than $(k-2)m$ vertices. Indeed, take any embedding f of $G = E(\!(V_1, \ldots, V_m)\!)$ into $E(\!(U_1, \ldots, U_m)\!)$. For every $i \in [m]$ at least $k-1$ vertices of V_i go into $U_{h(i)}$ for some $h(i)$. It is not hard to see that if we map each part V_i entirely into $U_{h(i)}$, then the new map is also an embedding. Since P is minimal, h has to be surjective and some extra work shows that necessarily $f(V_i) \subseteq U_{h(i)}$ for all $i \in [m]$. (In fact, if furthermore E consists of simple k-sets only, then $|V_i| \geq 1$ is enough for rigidity.)

The case $R \neq \emptyset$ is more complicated, although the main ideas (such as using the function h that specifies where a large part of V_i is mapped to) are roughly the same. One complication is that for a non-minimal pattern P there can be embeddings that map the bottom edges into different levels. For example, let

$$P = (5, \{\, \{\!\{1, 2, 3\}\!\}, \{\!\{1, 2, 4\}\!\}, \{\!\{1, 2, 5\}\!\}, \{\!\{3, 4, 5\}\!\}\,\}, \{5\})$$

and let f map the bottom parts $V_1, \ldots, V_5$ into respectively $U_1, U_2, U_{3,1}, U_{3,2}, U_{3,3}$. Here, P is obtained from the pattern $(3, \{\, \{\!\{1, 2, 3\}\!\}\,\}, \{3\})$ by "expanding" the third part up one level. Thus our proof of Lemma 17 should in particular catch all such redundancies.

Proof of Lemma 17. Let $n \to \infty$ and G be a maximum P-construction on $[n]$ with the partition structure $\mathbf{V}$. Take any embedding f of G into some P-construction H with the bottom partition $V(H) = U_1 \cup \ldots \cup U_m$ such that $f(V(G))$ intersects at least two different parts U_i.

Claim 17.1 *The map f is an induced embedding (that is, $f(X)$ is an edge if and only if X is).*

Proof of Claim. If some non-edge $D \in \overline{G}$ is mapped by f into an edge of H, then the k-graph $G \cup \{D\}$ embeds into a P-construction (the very same map f embeds it into H). However, this contradicts the maximality of G. ❙

By Lemma 11 and Part 2 of Lemma 15, the size of each V_i tends to infinity. By the pigeonhole principle, there is a function $h : [m] \to [m]$ such that

$$|f(V_i) \cap U_{h(i)}| \geq k, \quad \text{for all } i \in [m]. \tag{13}$$

Claim 17.2 *We can choose h in (13) so that, additionally, $h(R) \subseteq R$ and h assumes at least two different values.*

Proof of Claim. Suppose that $R \neq \emptyset$ and we cannot satisfy the first part of the claim for some $i \in R$, that is, for each $s \in R$ we have $|f(V_i) \cap U_s| < k$. Thus $G[V_i]$ with exception of at most $(k-1)|R|$ vertices is embeddable into $H[U_{[m]\setminus R}]$. By the maximality of G, Lemmas 11 and 15 give that $|V_i| \to \infty$ and $\rho(G[V_i]) = \Lambda_P + o(1)$. This means that $(P - R)$-constructions can contain arbitrarily large subgraphs of edge density $\Lambda_P + o(1)$, that is, $\Lambda_{P-R} \geq \Lambda_P$. However, this contradicts the minimality of P.

Let us restrict ourselves to those h with $h(R) \subseteq R$. Suppose that we cannot fulfil the second part of the claim. Then there is $j \in [m]$ such that $|f(V_i) \cap U_j| \geq k$ for every $i \in [m]$. Since $E \neq \emptyset$, the induced subgraph $G[f^{-1}(U_j)]$ is non-empty (it has at least k vertices from each V_i) and is mapped entirely into U_j. Thus $j \in R$. Since $f(V(G))$ intersects at least two different parts U_i, we can pick some $x \in V_i$ with $f(x) \in U_s$ and $s \neq j$. Fix some $(k-1)$-multiset $D \in E_i$. (Note that $E_i \neq \emptyset$ by (9).) Take an edge $D' \ni x$ of G so that $D' \setminus \{x\}$ is a subset of $f^{-1}(U_j)$ and has profile D; it exists because each part V_g contains at least k vertices of $f^{-1}(U_j)$. The k-set $f(D')$ is an edge of H as f is an embedding. However, it has $k-1$ vertices in U_j and one vertex in U_s. Thus the k-multiset $\{\{ j^{(k-1)}, s \}\}$ belongs to E. Since $j \in R$, this contradicts Lemma 12. The claim is proved. ▎

Claim 17.3 *Each h satisfying Claim 17.2 is a bijection.*

Proof of Claim. For $j \in [m]$ let $U_j' := \cup_{i \in h^{-1}(j)} f(V_i) \subseteq V(H)$. (Thus $U_j' = \emptyset$ for j not in the image of h.) Let H' be the P-construction on $f(V(G))$ such that $U_1' \cup \ldots \cup U_m'$ is the bottom partition of H' and, for $i \in R$, $H'[U_i']$ is the image of the P-construction $G[f^{-1}(U_i')]$ under the bijection f.

We have just defined a new P-construction H' so that each part V_i of G is entirely mapped by f into the $h(i)$-th part of H', that is, *all* vertices of G follow h now. This H' will be used only for proving that h is a bijection. The reader should be able to derive from the proof of Claim 17.3 that in fact $U_j' \subseteq U_j$ and $H' = H[f(V(G))]$ (but we will not use these properties).

Let us show first that the same map f is an embedding of G into H'. First, take any bottom edge $D \in G$ such that $f(D)$ intersects two different parts U_i'. Let $D' \in G$ have the same profile as D and satisfy

$$D' \subseteq \cup_{i \in [m]} \left(V_i \cap f^{-1}(U_{h(i)}) \right), \tag{14}$$

which is possible because there are at least k vertices available in each part V_i. Since $f(D') \cap U_i = f(D') \cap U_i'$ for all $i \in [m]$, the f-image of D' has the same profile X with respect to the partitions $U_1 \cup \ldots \cup U_m$ and $U_1' \cup \ldots \cup U_m'$. Thus $X \in E$. Next, as each $f(V_i)$ lies entirely inside $U_{h(i)}'$, the sets $f(D)$ and $f(D')$ have the same profiles with respect to parts U_i'. Thus $f(D)$ is an edge of $E(\!(U_1', \ldots, U_n')\!)$, as required. Next, take any $i \in [m]$ and let $G' := G[f^{-1}(U_i')]$. Assume that $i \in [m] \setminus R$ for otherwise $f(G') \subseteq H'$ by the definition of H'. We claim that G' has no edges in this case. Since $h(R) \subseteq R$, we have $h^{-1}(i) \cap R = \emptyset$. Thus it remains to derive a contradiction by assuming that a bottom edge D of G belongs to G'. As before, we can find an edge $D' \in G$ that satisfies (14) and has the same profile as D' with respect to $V_1, \ldots, V_m$. However, f maps this D' inside a non-recursive part U_i of H, a contradiction. Thus f is an embedding of G into H'.

Thus, by considering H' instead of H (and without changing h) we have that $f(V_i) \subseteq U'_{h(i)}$ for all $i \in [m]$.

Suppose on the contrary to the claim that $|h^{-1}(s)| \geq 2$ for some $s \in [m]$. Let $A := h^{-1}(s)$ and $B := [m] \setminus A$. Since h assumes at least two different values, the set B is non-empty.

Note that U'_s is *externally H'-homogeneous*, meaning that any permutation σ of $V(H')$ that fixes every vertex outside of U'_s is a symmetry of the set of H'-edges that intersect the complement of U'_s, that is, $\sigma(H' \setminus \binom{U'_s}{k}) = H' \setminus \binom{U'_s}{k}$. It follows from Claim 17.1 that $f^{-1}(U'_s) = V_A$ is externally G-homogeneous. (Recall that we denote $V_A := \cup_{i \in A} V_i$.) Since each V_i has at least k elements, we conclude that A is externally E-homogeneous.

Suppose first that $A \cap R \neq \emptyset$. By the above homogeneity, if we replace $G[V_A]$ by any P-construction, then the new k-graph on V is still a P-construction. Also, recall that each V_i has size $\Omega(n)$ by Lemmas 11 and 15. Thus, by the maximality of G, the edge density of $G[V_A]$ is $\Lambda_P + o(1)$. Also, $\rho(G[V_i]) = \Lambda_P + o(1)$ for $i \in A \cap R$. Consider the pattern $Q := P - B$ obtained by removing B from P. Without loss of generality assume that $A = [a]$. For $i \in A$ let $x_i := |V_i|/|V_A|$. The obtained vector $\mathbf{x} \in \mathbb{S}_a$ satisfies $\Lambda_P = \lambda_Q(\mathbf{x}) + \sum_{i \in A \cap R} \Lambda_P x_i^k + o(1)$. On the other hand, if we use the same vector $\mathbf{x}$ for the bottom ratios and put a maximum Q-construction on each recursive part, then we get overall density at most $\Lambda_Q + o(1)$. Thus $\Lambda_Q \geq \lambda_Q(\mathbf{x}) + \sum_{i \in A \cap R} \Lambda_Q x_i^k + o(1)$. Since G is a maximum P-construction, we have that each part V_i has $\Omega(n)$ vertices; thus no x_i can be equal $1 - o(1)$ and we have that $1 - \sum_{i \in A \cap R} x_i^k = \Omega(1)$. These inequalities imply that $\Lambda_Q \geq \lambda_Q(\mathbf{x})/(1 - \sum_{i \in A \cap R} x_i^k) + o(1) = \Lambda_P + o(1)$, contradicting the minimality of P.

Finally, suppose that $A \cap R = \emptyset$. Since V_A is externally G-homogeneous and A consists of at least two indices $i \neq j$, we have that E contains at least one multiset entirely inside A (for otherwise $E_i = E_j$, contradicting Lemma 16). Since $f(V_A) = U'_s$, we have that $s \in R$. By the maximality of G and Claim 17.1 it follows that the edge density of $H'[U'_s]$ (and thus of $G[V_A]$) is $\Lambda_P + o(1)$. Thus $\Lambda_{P-B} \geq \Lambda_P$, a contradiction proving the claim. ▌

It follows from Claim 17.3 that each h satisfying Claim 17.2 is an automorphism of P. By relabelling the parts of H, we can assume for notational convenience that h is the identity mapping. Now we are ready to prove the lemma, namely that $f(V_i) \subseteq U_i$ for every $i \in [m]$.

Suppose on the contrary that $f(x) \in U_j$ for some $x \in V_i$ and $j \in [m] \setminus \{i\}$. It follows that $E_i \subseteq E_j$. By Lemma 16 this inclusion is strict and $i \in R$. Pick some X from $E_j \setminus E_i \neq \emptyset$. We can find $D \in H$ such that $D \setminus \{f(x)\}$ has the profile X with respect to both $U_1 \cup \ldots \cup U_m$ and $f(V_1) \cup \ldots \cup f(V_m)$. But then $f^{-1}(D)$ is not an edge of G because its profile $X \cup \{i\}$ is not in E. (Note that it is impossible that $X = \{\!\{ i^{(k-1)} \}\!\}$ as this would give that E contains $\{\!\{ i^{(k-1)}, j \}\!\}$, the profile of $D \in H$, contradicting Lemma 12.) Thus f is not induced, contradicting Claim 17.1. This shows that G is rigid. ∎

Lemma 18 *Every rigid P-construction G with the partition structure $\mathbf{V}$ such that $|V_{\mathbf{i}}| \geq k$ for every legal $\mathbf{i}$ with $|\mathbf{i}| \leq 2$ remains rigid after the addition of any new vertex.*

Proof. It is enough to show that if we take any embedding f of G into a P-construction with the partition structure $\mathbf{U}$ such that $f(V_i) \subseteq U_i$ for $i \in [m]$ and add one vertex x to some part V_i, then any extension of f to x maps it into U_i. If $i \in [m] \setminus R$, then x and some $y \in V_i$ have the same links in $V' := V(G) \setminus \{x, y\}$; since the part containing $f(y)$ is determined by the values of f on V', the same applies to $f(x)$, as required. So let $i \in R$. Since $|V_{i,j}| \geq k$ for each $j \in [m]$, the link of x in G necessarily contains at least one $(k-1)$-set entirely inside V_i by (9). This forces by Lemma 12 that $f(x) \in U_i$, finishing the proof. ∎

Later (in the proof of Lemma 20) we will need the existence of a rigid P-construction such that the recursion goes for exactly ℓ levels for some ℓ and every part at height at most ℓ has many vertices. This can be achieved as follows. Take large n and let G be a maximum P-construction on $[n]$. It is rigid by Lemma 17. Also, by Lemma 11 and Part 2 of Lemma 15, G satisfies the assumptions of Lemma 18. Thus we can add some extra vertices to the P-construction G, without increasing its maximum height ℓ while achieving the following property:

Lemma 19 *There are $\ell \in \mathbb{N}$ and a rigid P-construction with the partition structure $\mathbf{V} = (V_{\mathbf{i}} : |\mathbf{i}| \leq \ell)$ such that for every legal sequence $\mathbf{i}$ of length at most ℓ we have $|V_{\mathbf{i}}| \geq (k-1)\max(m, k)$.* ∎

3.3 Key Lemmas

In this section, $P = (m, E, R)$ is still a proper pattern. Let us call two k-graphs with the same number of vertices *s-close* if one can be made isomorphic to the other by changing at most s edges.

Lemma 20 *There are $c_0 > 0$ and $M_0 \in \mathbb{N}$ such that the following holds. Let G be a maximum $\mathcal{F}_{M_0}$-free k-graph on $n \geq 2$ vertices that is $c_0\binom{n}{k}$-close to some P-construction. Then there is a partition $V(G) = V_1 \cup \ldots \cup V_m$ such that no V_i is equal to V and*

$$G \setminus (\cup_{i \in R} G[V_i]) = E(\!(V_1, \ldots, V_m)\!). \tag{15}$$

Proof. Clearly, it is enough to establish the existence of M_0 such that the conclusion of the lemma holds for every sufficiently large n. (Indeed, it clearly holds for $n \leq M_0$ by Lemma 7, so we can simply increase M_0 at the end to take care of finitely many exceptions; alternatively, one can decrease c_0.)

Let F be the rigid construction returned by Lemma 19. Let ℓ be the maximum height of F and let $\mathbf{W}$ be its partition structure. (Our proof also works if $R = \emptyset$, when $\ell = 1$; in fact, some parts can be simplified in this case.) Let $M_0 := v(F) + k$.

We choose some constants c_i in this order $c_4 \gg c_3 \gg c_2 \gg c_1 \gg c_0 > 0$, each being sufficiently small depending on the previous ones. Let n tend to infinity.

Let G be a maximum $\mathcal{F}_{M_0}$-free k-graph on $[n]$ that is $c_0\binom{n}{k}$-close to some P-construction H. We can assume that $V(H) = [n]$ and the vertices of H are already re-labelled so that $|G \triangle H| \le c_0\binom{n}{k}$. Let $\mathbf{V}$ be the partition structure of H. In particular, the bottom partition of H is $V_1 \cup \ldots \cup V_m$.

One of the technical difficulties that we are going to face is that some part V_i with $i \in R$ may in principle contain almost every vertex of $V(G)$ (so every other part V_j has $o(n)$ vertices). This means that the "real" approximation to G starts only at some higher level inside V_i. On the other hand, Lemma 15 gives us a way to rule out such cases: we have to ensure that the minimal degree of H is close to $\Lambda_P\binom{n-1}{k-1}$. So, as our first step, we are going to modify the P-construction H (perhaps at the expense of increasing $|G \triangle H|$ slightly) so that its minimal degree is large.

Namely, let $Z := \{x \in [n] : d_H(x) < (\Lambda_P - 2c_1)\binom{n-1}{k-1}\}$. By Lemma 11 we can assume that

$$\delta(G) \ge (\Lambda_P - c_1)\binom{n-1}{k-1}. \tag{16}$$

Thus every vertex of Z contributes at least $c_1\binom{n-1}{k-1}/k$ to $|G \triangle H|$. We conclude that $|Z| \le c_0 n/c_1$. Fix an arbitrary $y \in [n] \setminus Z$. Let us change H by making all vertices in Z into clones of y (and updating $\mathbf{V}$ accordingly). Clearly, we have now

$$\delta(H) \ge (\Lambda_P - 2c_1)\binom{n-1}{k-1} - |Z|\binom{n-2}{k-2} \ge (\Lambda_P - 3c_1)\binom{n-1}{k-1} \tag{17}$$

while $|G \triangle H| \le c_0\binom{n}{k} + |Z|\binom{n-1}{k-1} \le c_1\binom{n}{k}$.

If we end up with an improper partition structure (e.g. $V_i = V(H)$ for some $i \in R$), then we correct this as in the proof of Lemma 6 without changing the k-graph H.

By Lemma 15 we can conclude that (in the new k-graph H) all part ratios up to height ℓ are close to optimal ones and $|V_{\mathbf{i}}| \ge 2c_4 n$ for each legal sequence $\mathbf{i}$ of length at most ℓ.

Let $A := E(\!(V_1, \ldots, V_m)\!) \setminus G$ consist of what we shall call *absent* edges. Let us call a k-multiset D on $[m]$ *bad* if $D \notin E$ and $D \neq \{\!\{\, i^{(k)} \,\}\!\}$ for some $i \in R$. Let

$$B := (G \setminus E(\!(V_1, \ldots, V_m)\!)) \setminus \left(\cup_{i \in R}\binom{V_i}{k}\right)$$

and call edges in B *bad*. Equivalently, an edge of G is bad if its profile is bad. Define $a := |A|$ and $b := |B|$. Our aim is to achieve that $a = b = 0$.

Our next modification is needed to ensure later that (23) holds. Roughly speaking, we want a property that the number of bad edges cannot be decreased much if we move one vertex between parts. Unfortunately, we cannot just take a partition structure $\mathbf{V}$ that minimises b because then we do not know how to guarantee the high minimum degree of H (another property important in our proof). Nonetheless, we can simultaneously satisfy both properties with some extra work.

Namely, we modify H as follows (updating A, B, $\mathbf{V}$, etc, as we proceed). If there is a vertex $x \in [n]$ such that by moving it to another part V_i we decrease b by at least $c_2\binom{n-1}{k-1}$, then we pick $y \in V_i$ of maximum H-degree and make x a clone of y. (Note that the new value of b depends only on the index i of the part V_i but not on the choice of $y \in V_i$.) Clearly, we perform this operation at most $c_1\binom{n}{k}/c_2\binom{n-1}{k-1} = c_1 n/(c_2 k)$ times because we initially had $b \le |G \bigtriangleup H| \le c_1\binom{n}{k}$. Thus, we have at all steps of this process (which affects at most $c_1 n/(c_2 k)$ vertices of H) that, trivially,

$$\begin{aligned} |V_{\mathbf{i}}| &\ \ge\ 2c_4 n - \frac{c_1 n}{c_2 k} \ \ge\ c_4 n, \qquad \text{for all legal } \mathbf{i} \text{ with } |\mathbf{i}| \le \ell, && (18)\\ |G \bigtriangleup H| &\ \le\ c_1\binom{n}{k} + \frac{c_1 n}{c_2 k}\binom{n-1}{k-1} \ \le\ c_2\binom{n}{k}. && (19)\end{aligned}$$

It follows that at every step each part V_i had a vertex of degree at least $(\Lambda_P - c_2/2)\binom{n-1}{k-1}$ for otherwise the edit distance between H and G at that moment would be at least $\frac{c_2}{3}\binom{n-1}{k-1} \times c_4 n$ by (16) and (18), contradicting the first inequality in (19). This implies that every time we clone a vertex it has a high degree. Thus we have by (17) that, additionally to (18) and (19), the following holds at the end of this process:

$$\delta(H) \ge (\Lambda_P - \max(3c_1, c_2/2))\binom{n-1}{k-1} - \frac{c_1 n}{c_2 k}\binom{n-2}{k-2} \ \ge\ (\Lambda_P - c_2)\binom{n-1}{k-1}. \qquad (20)$$

If we take the union of $E(\!(V_1, \ldots, V_m)\!)$ with $\cup_{i \in R} G[V_i]$, then the obtained k-graph is still $\mathcal{F}_{M_0}$-free by Lemma 9 and has exactly $a - b + |G|$ edges. The maximality of G implies that

$$b \ge a. \qquad (21)$$

Suppose that $b > 0$ for otherwise $a = b = 0$ and the lemma is proved. Let

$$H' := H \setminus (\cup_{\mathbf{i} \in R^\ell} H[V_{\mathbf{i}}])$$

be obtained from H by "truncating" it down to the first ℓ levels.

Let us show that the maximal degree of B is small, namely that

$$\Delta(B) < c_3\binom{n-1}{k-1}. \qquad (22)$$

Suppose on the contrary that $d_B(x) \ge c_3\binom{n-1}{k-1}$ for some $x \in [n]$.

It may be helpful to informally illustrate our argument leading to a contradiction on the special case when P is as in (10). Assume that Lemma 19 returns $\ell = 1$ and $F = E(\!(W_1, W_2)\!)$ with $|W_i| = 6$ (although some smaller W_i's will also suffice). Here we have $H' = E(\!(V_1, V_2)\!)$. Suppose, for example, that the vertex x contradicting (22) is in V_2. Let $B_{x,2} := B_x = G_x \cap (\binom{V_2}{2} \cup \binom{V_1}{2})$ be the link of x in the bad 3-graph B. Next, let $B_{x,1} := G_x \cap K_2^2(\!(V_1, V_2)\!)$ be the set of pairs that would form a bad edge with x if x is moved to V_1. By our assumption, we have $|B_{x,2}| \ge c_3\binom{n-1}{2}$. Also, $|B_{x,1}| \ge (c_3 - c_2)\binom{n-1}{2}$ for otherwise we would have moved x to V_1, thus decreasing b substantially. Take any $\mathbf{D} = (D_1, D_2)$, where $D_i \in B_{x,i}$. Consider arbitrary 6-vertex sets $Z_1 \subseteq V_1$ and $Z_2 \subseteq V_2 \setminus \{x\}$

such that $Z_1 \cup Z_2$ contains the set $D_1 \cup D_2$. It is impossible that $E(\!(Z_1, Z_2)\!) \subseteq G$ because the 3-graph $F_{\mathbf{D}}$ obtained from $E(\!(Z_1, Z_2)\!)$ by adding the edges $D_1 \cup \{x\}$ and $D_2 \cup \{x\}$ belongs to $\mathcal{F}_{13}$. (Indeed, if we could embed $F_{\mathbf{D}}$ into a P-construction, then all vertices of Z_1 and Z_2 would have to go into "correct" parts by the rigidity of $F \cong E(\!(Z_1, Z_2)\!)$, leaving no way to fit x.) Thus $Z_1 \cup Z_2$ contains at least one absent edge. Finally, our lower estimates on $|B_{x,i}|$ translate with some work into a lower bound on a that contradicts (19).

Let us give the general argument. For $i \in [m]$ let the $(k-1)$-graph $B_{x,i}$ consist of those $D \in G_x$ such that if we add i to the profile of D then the obtained k-multiset is bad. In other words, if we move x to V_i, then $B_{x,i}$ will be the link of x with respect to the (updated) bad k-graph B. By the definition of H, we have

$$|B_{x,i}| \geq (c_3 - c_2)\binom{n-1}{k-1}, \qquad \text{for every } i \in [m]. \tag{23}$$

For $\mathbf{D} = (D_1, \ldots, D_m) \in \prod_{i=1}^m B_{x,i}$ let $F_{\mathbf{D}}$ be the k-graph that is constructed as follows. Recall that F is the rigid P-construction given by Lemma 19 and $\mathbf{W}$ is its partition structure. By relabelling vertices of F, we can assume that $x \notin V(F)$ while $D := \cup_{i=1}^m D_i$ is a subset of $V(F)$ so that for every $y \in D$ we have $\mathrm{br}_F(y) = \mathrm{br}_{H'}(y)$, that is, y has the same branches in both F and H'. (Note that both k-graphs have the same maximum height ℓ.) This is possible because each part of F of height at most ℓ has at least $m(k-1) \geq |D|$ vertices. Finally, add x as a new vertex and the sets $D_i \cup \{x\}$ for $i \in [m]$ as edges, obtaining the k-graph $F_{\mathbf{D}}$.

Claim 20.1 *For every $\mathbf{D} \in \prod_{i=1}^m B_{x,i}$ we have $F_{\mathbf{D}} \in \mathcal{F}_\infty$.*

Proof of Claim. Suppose on the contrary that we have an embedding f of $F_{\mathbf{D}}$ into some P-construction with the partition structure $\mathbf{U}$. By the rigidity of F, we can assume that $f(W_i) \subseteq U_i$ for every $i \in [m]$. Let $i \in [m]$ satisfy $f(x) \in U_i$. But then the edge $D_i \cup \{x\} \in F_{\mathbf{D}}$ is mapped into a non-edge because $f(D_i \cup \{x\})$ has bad profile with respect $U_1, \ldots, U_m$ by the choice of $D_i \in B_{x,i}$, a contradiction. ▮

For every vector $\mathbf{D} \in \prod_{i=1}^m B_{x,i}$ and every map $f : V(F_{\mathbf{D}}) \to V(G)$ such that f is the identity on $D \cup \{x\}$ and f preserves branches of height up to ℓ on all other vertices, the image $f(F_{\mathbf{D}})$ has to contain some $X \in \overline{G}$ by Claim 20.1. (Note that G is $F_{\mathbf{D}}$-free since $v(F_{\mathbf{D}}) \leq M_0$.) Also,

$$f(F_{\mathbf{D}} \setminus \{D_1 \cup \{x\}, \ldots, D_m \cup \{x\}\}) \subseteq H',$$

that is, the "base" copy of F on which $F^{\mathbf{D}}$ was built is embedded by f into H'. On the other hand, each of the edges $D_1 \cup \{x\}, \ldots, D_m \cup \{x\}$ of $F_{\mathbf{D}}$ that contain x is mapped to an edge of G (to itself). Thus $X \in H' \setminus G$ and $X \not\ni x$. Any such X can appear, very roughly, for at most $\binom{w}{k-1}^m (w+1)!\, n^{w-k}$ choices of $(\mathbf{D}, f)$, where $w := v(F) = v(F_{\mathbf{D}}) - 1$. On the other hand, the total number of choices of $(\mathbf{D}, f)$ is at least $\prod_{i=1}^m |B_{x,i}| \geq ((c_3 - c_2)\binom{n-1}{k-1})^m$ times $(c_4 n/2)^{w-(k-1)m}$ (since every part of H' has at least c_4n vertices by (18)). We conclude that

$$|H \setminus G| \geq |H' \setminus G| \geq \frac{\left((c_3 - c_2)\binom{n-1}{k-1}\right)^m \times (c_4n/2)^{w-(k-1)m}}{\binom{w}{k-1}^m (w+1)!\, n^{w-k}} > c_2\binom{n}{k}.$$

However, this contradicts (19). Thus (22) is proved.

Next, we show (in Claim 20.3 below) that *every* bad edge D intersects $\Omega(c_3n^{k-1})$ absent edges. Again, let us first illustrate the proof on the case when P is as in (10). Suppose that $D = \{y_1, y_2, z\}$ with, say, $y_1, y_2 \in V_1$ and $z \in V_2$. For $i = 1, 2$, let D_{y_i} be an edge of $G[V_1]$ such that $D_{y_i} \cap D = \{y_i\}$; there are many such edges because we have by Part 3 of Lemma 15 that, for example,

$$d_{G[V_i]}(y) \geq c_4\binom{n-1}{k-1}, \quad \text{for all } i \in R \text{ and } y \in V_i. \tag{24}$$

Let $\mathbf{D} = (D_{y_1}, D_{y_2})$. Take arbitrary 6-sets $Z_i \subseteq V_i \setminus D$ such that $Z_1 \supseteq (D_{y_1} \cup D_{y_2}) \setminus D$. Let $F^{\mathbf{D}}$ be the 3-graph obtained from $E(\!(Z_1, Z_2 \cup \{z\})\!)$ by adding vertices y_1, y_2 and edges D, D_{y_1}, D_{y_2}. It is not hard to show that $F^{\mathbf{D}}$ belongs to $\mathcal{F}_{15}$. Thus $E(\!(Z_1, Z_2 \cup \{z\})\!) \not\subseteq G$ and we arrive at some $Y \in A$. It is impossible that the obtained absent edge Y is disjoint from D for at least half of choices of $(D_{y_1}, D_{y_2}, Z_1, Z_2)$, for otherwise $|A|$ is too large. If $Y \cap D$ is not empty, then it consists of the unique vertex z. Some counting gives the desired lower bound on the number of absent edges intersecting D. Note that this counting would not work if the obtained $Y \in A$ could share more than one vertex with D. This is the reason why we do not allow $Z_1 \cup Z_2$ to share more than one vertex with D; we make these sets disjoint in the general proof for the notational convenience.

Let us present the general argument. Let $D \in B$ be an arbitrary bad edge. For each $i \in R$ and $y \in D \cap V_i$ pick some $D_y \in G[V_i]$ such that $D_y \cap D = \{y\}$; it exists by (24). Let $\mathbf{D} := (D, \{D_y : y \in D \cap V_R\})$. (Recall that $V_R = \cup_{i \in R} V_i$.) We define the k-graph $F^{\mathbf{D}}$ using the rigid k-graph F as follows. By re-labelling $V(F)$, we can assume that $X \subseteq V(F)$, where

$$X := \cup_{y \in D \cap V_R} D_y \setminus \{y\}, \tag{25}$$

so that for every $x \in X$ its branches in F and H' coincide. Again, there is enough space inside F to accommodate all $|X| \leq k(k-1)$ vertices of X. Assume also that D is disjoint from $V(F)$. The vertex set of $F^{\mathbf{D}}$ is $V(F) \cup D$. Starting with the edge-set of F, add D and each D_y with $y \in D \cap V_R$. Finally, for every $y \in D \cap V_i$ with $i \in [m] \setminus R$ pick some $z \in W_i$ and add $\{Z \cup \{y\} : Z \in F_z\}$ to the edge set, obtaining the k-graph $F^{\mathbf{D}}$. The last step can be viewed as enlarging the part W_i by $D \cap V_i$ and adding those edges that are stipulated by the pattern P and intersect D in at most one vertex.

Claim 20.2 *For every* $\mathbf{D}$ *as above, we have* $F^{\mathbf{D}} \in \mathcal{F}_\infty$.

Proof of Claim. Suppose on the contrary that we have an embedding f of $F^{\mathbf{D}}$ into some P-construction with the partition structure $\mathbf{U}$. We can assume by the rigidity of F, that $f(W_i) \subseteq U_i$ for each i.

Take $y \in D \cap V_i$ with $i \in R$. The $(k-1)$-set $f(D_y \setminus \{y\})$ lies entirely inside U_i. We cannot have $f(y) \in U_j$ with $j \neq i$ because otherwise the profile of the edge $f(D_y)$ is $\{\!\{ i^{(k-1)}, j \}\!\}$, contradicting Lemma 12. Thus $f(y) \in U_i$.

Next, take any $y \in D \cap V_i$ with $i \in [m] \setminus R$. Pick some $z \in W_i$. By the rigidity of F, if we fix the restriction of f to $V(F) \setminus \{z\}$, then U_i is the only part where z can be mapped to. By definition, y and z have the same link $(k-1)$-graphs in $F^{\mathbf{D}}$ when restricted to $V(F) \setminus \{y, z\}$. Hence, $f(y)$ necessarily belongs to W_i.

Thus the edge $f(D)$ has the same profile as $D \in B$, a contradiction. ▮

Claim 20.3 *For every $D \in B$ there are at least $kc_3\binom{n-1}{k-1}$ absent edges $Y \in A$ with $|D \cap Y| = 1$.*

Proof of Claim. Given D choose the sets D_y, for $y \in D \cap V_R$, as before Claim 20.2. The condition $D_y \cap D = \{y\}$ rules out at most $k\binom{n-2}{k-2}$ edges for this y. Thus by (24) there are, for example, at least $(c_4/2)\binom{n-1}{k-1}$ choices of each D_y. Form the k-graph $F^{\mathbf{D}}$ as above and consider potential injective embeddings f of $F^{\mathbf{D}}$ into G that are the identity on $D \cup X$ and map every other vertex of F into a vertex of H' with the same branch, where X is defined by (25). For every vertex $x \notin D \cup X$ we have at least $c_4 n/2$ choices for $f(x)$ by (18). By Claim 20.2, G does not contain $F^{\mathbf{D}}$ as a subgraph so its image under f contains some $Y \in \overline{G}$. Since f maps D and each D_y to an edge of G (to itself) and

$$f\left(F^{\mathbf{D}} \setminus (\{D\} \cup \{D_y : y \in D \cap V_R\})\right) \subseteq H',$$

we have that $Y \in H'$. The number of choices of $(\mathbf{D}, f)$ is at least

$$\left((c_4/2)\binom{n-1}{k-1}\right)^{|D \cap V_R|} \times (c_4 n/2)^{w-(k-1)|D \cap V_R|} \geq (c_4 n/4k)^w,$$

where $w := v(F)$. Assume that for at least half of the time the obtained set Y intersects D for otherwise we get a contradiction to (19):

$$|H' \setminus G| \geq \frac{1}{2} \times \frac{(c_4 n/4k)^w}{\binom{w}{k-1}^k (w+k)!\, n^{w-k}} > c_2 \binom{n}{k}.$$

By the definitions of $F^{\mathbf{D}}$ and f, we have that $|Y \cap D| = 1$ and $Y \in A$. Each such $Y \in A$ is counted for at most $\binom{w}{k-1}^k (w+k)!\, n^{w-k+1}$ choices of f. Thus the number of such Y is at least $\frac{1}{2}(c_4 n/4k)^w / (\binom{w}{k-1}^k (w+k)!\, n^{w-k+1})$, implying the claim. ▮

Let us count the number of pairs (Y, D) where $Y \in A$, $D \in B$, and $|Y \cap D| = 1$. On one hand, each bad edge $D \in B$ creates at least $kc_3\binom{n-1}{k-1}$ such pairs by Claim 20.3. On the other hand, we trivially have at most $ak\Delta(B)$ such pairs. By (21), we have $bkc_3\binom{n-1}{k-1} \leq ak\Delta(B) \leq bk\Delta(B)$. Since $b \neq 0$, we obtain a contradiction to (22). This finishes the proof of Lemma 20. ▮

Let us state a special case of a result of Rödl and Schacht [37, Theorem 6] that we will need.

Lemma 21 (Strong Removal Lemma [37]) *For every k-graph family $\mathcal{F}$ and $\varepsilon > 0$ there are $\delta > 0$, m, and n_0 such that the following holds. Let G be a k-graph on $n \geq n_0$ vertices such that for every $F \in \mathcal{F}$ with $v(F) \leq m$ the number of F-subgraphs in G is at most $\delta n^{v(F)}$. Then G can be made $\mathcal{F}$-free by removing at most $\varepsilon\binom{n}{k}$ edges.* ∎

Lemma 22 *For every $c_0 > 0$ there is M_1 such that every maximum $\mathcal{F}_{M_1}$-free G with $n \geq M_1$ vertices is $c_0\binom{n}{k}$-close to a P-construction.*

Proof. Lemma 21 gives M_1 such that any $\mathcal{F}_{M_1}$-free k-graph G on $n \geq M_1$ vertices can be made into an $\mathcal{F}_\infty$-free k-graph G' by removing at most $c_0\binom{n}{k}/2$ edges. By Lemma 7, G' embeds into some P-construction H with $v(H) = v(G')$. Assume that $V(H) = V(G')$ and the identity map is an embedding of G' into H.

Since H is $\mathcal{F}_{M_1}$-free, the maximality of G implies that $|G| \geq |H|$. Thus $|H \setminus G'| \leq c_0\binom{n}{k}/2$ and we can transform G' into H by changing at most $c_0\binom{n}{k}/2$ further edges. ∎

3.4 Proof of Theorem 3: Putting All Together

We are ready to prove Theorem 3. It is trivially true if $\Lambda_P = 0$ or 1 by the discussion after Lemma 12, so we can assume that P is proper. Apply Lemma 20 which returns c_0 and M_0. Next, Lemma 22 on input c_0 returns some M_1.

Let us show that $M = \max(M_0, M_1)$ works in Theorem 3. We use induction on n. Let G be any maximum $\mathcal{F}_M$-free k-graph on $[n]$. Suppose that $n > M$ for otherwise we are done by Lemma 7. Thus Lemma 22 applies and shows that G is $c_0\binom{n}{k}$-close to some P-construction. Lemma 20 returns a partition $[n] = V_1 \cup \ldots \cup V_m$ such that (15) holds.

Let $i \in R$ be arbitrary. By Lemma 9 if we replace $G[V_i]$ by a maximum $\mathcal{F}_M$-free k-graph, then the new k-graph on V is still $\mathcal{F}_M$-free. By the maximality of G, we conclude that $G[V_i]$ is a maximum $\mathcal{F}_M$-free k-graph. By the induction hypothesis (note that $|V_i| \leq n-1$), $G[V_i]$ is a P-construction.

It follows that G is a P-construction itself, which implies all claims of Theorem 3. ∎

4 Proof of Theorem 5

By Corollary 4 it is enough to exhibit, for every $k \geq 3$, a pattern P such that Λ_P is irrational.

Given $k \geq 3$, let ℓ be any prime number that does not divide k such that $2 \leq \ell < k$. If k is odd, we can take $\ell = 2$. For even k we can take ℓ to be any prime with $k/2 < \ell < k$; it exists by Bertrand's postulate. Take $P = (2, E, \{1\})$, where E consists of the single multiset $\{\!\{ 1^{(k-\ell)}, 2^{(\ell)} \}\!\}$. In other words, a P-construction on V is obtained by partitioning $V = V_1 \cup V_2$ with $V_1 \neq V$, adding all k-sets that intersect V_1 in exactly $k-\ell$ vertices, and doing recursion inside V_1. If $k=3$, we get the familiar pattern from (10).

Let $\mathbf{x}=(x_1,x_2)\in\mathcal{X}$ be an optimal vector. We have by (11) that $\Lambda_P/\binom{k}{\ell}=r(x_1)$, where

$$r(x):=\frac{(1-x)^\ell x^{k-\ell}}{1-x^k}.$$

By Part 2 of Lemma 14 the real x_1 maximises r over $(0,1)$. Thus x_1 is a root of the derivative of r. We have

$$r'(x)=-\frac{(1-x)^\ell x^{k-\ell-1}}{(1-x^k)^2}\,g(x),$$

where $g(x):=\ell(x^{k-1}+\ldots+1)-k$. Since $x_1\neq 0,1$, it is a root of g.

The polynomial $g(x)$ is irreducible (by Eisenstein's criterion). Indeed, if we can factorise $g(x)=(a_mx^m+\ldots+a_0)(b_{k-1-m}x^{k-1-m}+\ldots+b_0)$ in $\mathbb{Z}[x]$, then exactly one of a_m and b_{k-1-m} is divisible by ℓ, say b_{k-1-m} is. Since $b_0a_0=\ell-k$, we have that ℓ does not divide b_0. Thus there is i such that ℓ does not divide b_i but ℓ divides each b_j with $j>i$. But then the coefficient at x^{m+i} is congruent to a_mb_i modulo ℓ, which implies that $m=0$ and the first factor is equal to ±1.

Thus x_1 is irrational but we still have to show that Λ_P is irrational. Suppose on the contrary that $\Lambda_P=s/t$ with $s,t\in\mathbb{Z}$. Note that $(x-1)g(x)=\ell x^k-kx+k-\ell$. Thus $x_1^k=(kx_1-k+\ell)/\ell$. Substituting this in $\Lambda_P=\binom{k}{\ell}r(x_1)$, we infer that

$$\frac{s}{t}=\binom{k}{\ell}\frac{(1-x_1)^\ell x_1^{k-\ell}}{1-(kx_1-k+\ell)/\ell}=\binom{k-1}{\ell-1}(1-x_1)^{\ell-1}x_1^{k-\ell}.$$

Thus x_1 is a root of the polynomial

$$h(x):=t\,(k-1)!\,(1-x)^{\ell-1}x^{k-\ell}-s\,(\ell-1)!\,(k-\ell)!\ \in\ \mathbb{Z}[x]$$

which has to be divisible by the irreducible polynomial g. Since these polynomials have the same degree $k-1$, h is a constant multiple of g. But the highest two coefficients of g are the same while those of h have different signs, a contradiction. Thus Λ_P is irrational, proving Theorem 5. ∎

5 Proof of Proposition 1

Here we prove Proposition 1. The proof is motivated by the emerging theory of the limits of discrete structures, see e.g. [5, 13, 28, 29]. Also, an intermediate result that we obtain (Theorem 24) may be of independent interest. We make the presentation essentially self-contained by restricting ourselves to only one aspect of hypergraph limits. In particular, we do not rely on the machinery developed by Elek and Szegedy [13].

Let F and G be k-graphs. A *homomorphism* from F to G is a map $f:V(F)\to V(G)$, not necessarily injective, such that $f(A)\in G$ for every $A\in F$. (Thus, embeddings are precisely injective homomorphisms.) Let the *homomorphism density* $t(F,G)$ be the probability that a random map

$V(F) \to V(G)$, with all $v(G)^{v(F)}$ choices being equally likely, is a homomorphism. For example, we have $t(K_k^k, G) = k!\,|G|/v(G)^k$.

Let $\mathcal{G}^{(k)}$ consist of all k-graphs up to isomorphism. A sequence $(G_i)_{i=1}^{\infty}$ of k-graphs *converges to a function* $\phi : \mathcal{G}^{(k)} \to [0,1]$ if the sequence is *increasing* (i.e. $v(G_1) < v(G_2) < \ldots$) and for every k-graph F we have $\lim_{i\to\infty} t(F, G_i) = \phi(F)$. Clearly, the convergence is not affected if we modify $o(v(G_i)^k)$ edges in each G_i. Let $\mathrm{LIM}^{(k)}$ consist of all possible functions ϕ that can be obtained in the above manner.

Given a family $\mathcal{F}$ of forbidden k-graphs, let $\mathcal{T}(\mathcal{F}) \subseteq \mathrm{LIM}^{(k)}$ consist of all possible limits of increasing sequences $(G_i)_{i=1}^{\infty}$ such that $\lim_{i\to\infty} \rho(G_i) = \pi(\mathcal{F})$ and each G_i is $\mathcal{F}$-free. In other words, $\mathcal{T}(\mathcal{F})$ is the set of the limits of almost maximum $\mathcal{F}$-free k-graphs. The standard diagonalisation argument shows that every increasing sequence has a convergent subsequence; in particular, $\mathcal{T}(\mathcal{F}) \neq \emptyset$. Let $\mathcal{T}^{(k)}$ be the union of $\mathcal{T}(\mathcal{F})$ over all k-graph families $\mathcal{F}$. We have

$$\Pi_{\infty}^{(k)} = \{\phi(K_k^k) : \phi \in \mathcal{T}^{(k)}\}. \tag{26}$$

Let the *blow-up closure* $\overline{\mathcal{F}}$ of $\mathcal{F} \subseteq \mathcal{G}^{(k)}$ consist of all k-graphs F such that some blow-up of F is not $\mathcal{F}$-free. Clearly, $\mathcal{F} \subseteq \overline{\mathcal{F}}$. Also, it is easy to see that by applying the blow-up closure twice we get the same family $\overline{\mathcal{F}}$.

Lemma 23 *For every $\mathcal{F} \subseteq \mathcal{G}^{(k)}$ and $\varepsilon > 0$ there is n_0 such that any $\mathcal{F}$-free k-graph G with $n \geq n_0$ vertices can be made $\overline{\mathcal{F}}$-free by removing at most $\varepsilon\binom{n}{k}$ edges.*

In particular, it follows that $\pi(\mathcal{F}) = \pi(\overline{\mathcal{F}})$ and $\mathcal{T}(\mathcal{F}) = \mathcal{T}(\overline{\mathcal{F}})$.

Proof. Let Lemma 21 on input $(\mathcal{F}, \varepsilon)$ return m and $\delta > 0$. Let n be large and G be an arbitrary $\mathcal{F}$-free k-graph on $[n]$. For each $F \in \overline{\mathcal{F}}$ there is s such that G is $F(s)$-free. As it is well known (see e.g. [8, Theorem 3]), G contains at most $\delta n^{v(F)}$ copies of F for all large n. Since there are only finitely many non-isomorphic k-graphs on at most m vertices, we can satisfy the above bound for all such k-graphs by taking n large. Now Lemma 21 applies, giving the first part.

Thus any increasing sequence $(G_i)_{i=1}^{\infty}$ of asymptotically maximum $\mathcal{F}$-free k-graphs can be converted into that for $\overline{\mathcal{F}}$ by modifying $o(v(G_i)^k)$ edges in each G_i. This modification does not affect, for any fixed F, the limit of $t(F, G_i)$ as $i \to \infty$, implying the second part. ∎

Recall that the Lagrangian of a k-graph G on $[n]$ is $\Lambda_G = \max\{\lambda_G(\mathbf{x}) : \mathbf{x} \in \mathbb{S}_n\}$; equivalently, Λ_G is the Lagrangian Λ_P of the pattern $P := (n, G, \emptyset)$ as defined by (3). We have the following characterisation of the set $\mathcal{T}^{(k)}$.

Theorem 24 *For $\phi \in \mathrm{LIM}^{(k)}$, the following are equivalent:*

1. *$\phi \in \mathcal{T}^{(k)}$;*

2. *ϕ is a limit of an increasing k-graph sequence $(G_i)_{i=1}^{\infty}$ such that $\rho(G_i) - \Lambda_{G_i} \to 0$;*

3. $\phi(F) = 0$ *for every* k*-graph* F *with* $t(K_k^k, F) > \phi(K_k^k)$.

Proof. **1) $\Rightarrow$ 2):** Let $\phi \in \mathcal{T}^{(k)}$, say $\phi \in \mathcal{T}(\mathcal{F})$. By Lemma 23 we can assume that $\overline{\mathcal{F}} = \mathcal{F}$. Let $(G_i)_{i=1}^{\infty}$ be a sequence of almost maximum $\mathcal{F}$-free k-graphs that converges to ϕ. Take any i and let $n := v(G_i)$. Since $\overline{\mathcal{F}} = \mathcal{F}$, any blow-up of G_i is still $\mathcal{F}$-free. Also, the limit superior of the edge densities attained by increasing blow-ups of G_i is exactly Λ_{G_i}. Thus $\Lambda_{G_i} \le \pi(\mathcal{F}) = \rho(G_i) + o(1)$. On the other hand, we have $\Lambda_{G_i} \ge \lambda_{G_i}(1/n, \ldots, 1/n) = k!\,|G_i|/n^k$, giving the converse inequality.

2) $\Rightarrow$ 3): Let ϕ satisfy 2). Suppose on the contrary that some F on $[n]$ violates 3). Pick a sequence $(G_i)_{i=1}^{\infty}$ given by 2). By the definition of convergence, G_i contains F as a subgraph for all large i. But then

$$\Lambda_{G_i} \ge \Lambda_F \ge \frac{k!\,|F|}{n^k} = t(K_k^k, F) \ge \phi(K_k^k) + \Omega(1),$$

contradicting $\phi(K_k^k) = \lim_{i\to\infty} \rho(G_i) = \lim_{i\to\infty} \Lambda_{G_i}$.

3) $\Rightarrow$ 1): Given ϕ as in 3), let

$$\mathcal{F} := \{F \in \mathcal{G}^{(k)} : \phi(F) = 0\}.$$

Let H_n be a maximum $\mathcal{F}$-free k-graph on $[n]$. Since $H_n \notin \mathcal{F}$, we have $\phi(H_n) > 0$. Thus, by 3),

$$\phi(K_k^k) \ge t(K_k^k, H_n) = \frac{k!\,|H_n|}{n^k} \ge \rho(H_n) + O(1/n).$$

By letting $n \to \infty$, we obtain that

$$\phi(K_k^k) \ge \pi(\mathcal{F}). \tag{27}$$

Let us show that $\phi \in \mathcal{T}(\mathcal{F})$. Take any increasing sequence $(G_i)_{i=1}^{\infty}$ convergent to ϕ. Let $F \in \mathcal{F}$. Since $\phi(F) = 0$, the number of F-subgraphs in G_i is at most $o(v(G_i)^{v(F)})$. By Lemma 21, we can remove $o(v(G_i)^k)$ edges in each G_i, obtaining an $\mathcal{F}$-free k-graph G_i'. Thus

$$\pi(\mathcal{F}) \ge \lim_{i\to\infty} \rho(G_i') = \phi(K_k^k)$$

and, by (27), this is equality. The obtained sequence $(G_i')_{i=1}^{\infty}$ of almost maximum $\mathcal{F}$-free k-graphs still converges to ϕ. This shows that $\phi \in \mathcal{T}(\mathcal{F}) \subseteq \mathcal{T}^{(k)}$. ∎

Let us view $\mathrm{LIM}^{(k)}$ as a subset of $[0,1]^{\mathcal{G}^{(k)}}$, where the latter set is endowed with the product (or pointwise convergence) topology. If we identify each k-graph G with the sequence $(t(F,G))_{F\in\mathcal{G}^{(k)}}$, then this topology gives exactly the above convergence. Moreover, the set $\mathrm{LIM}^{(k)}$, as the topological closure of $\mathcal{G}^{(k)}$, is a closed subset of $[0,1]^{\mathcal{G}^{(k)}}$.

Corollary 25 *For every* $k \ge 2$ *the set* $\mathcal{T}^{(k)}$ *is a closed subset of* $[0,1]^{\mathcal{G}^{(k)}}$.

Proof. The third characterisation of Theorem 24 shows that

$$\mathcal{T}^{(k)} = \cap_{F\in\mathcal{G}^{(k)}} \left(\{\phi \in \mathrm{LIM}^{(k)} : \phi(F) = 0\} \cup \{\phi \in \mathrm{LIM}^{(k)} : \phi(F) \ge t(K_k^k, F)\}\right). \tag{28}$$

For every $F \in \mathcal{G}^{(k)}$, the map $\phi \mapsto \phi(F)$ is continuous (it is just the projection of $[0,1]^{\mathcal{G}^{(k)}}$ onto the F-th coordinate). We see by (28) that $\mathcal{T}^{(k)}$, as the intersection of closed sets, is itself closed. ∎

Proof of Proposition 1. Let $a_i \in \Pi_\infty^{(k)}$ with $a_i \to a$ as $i \to \infty$. By Part 2 of Theorem 24 we can find, for each $i \in \mathbb{N}$, a k-graph H_i such that $v(H_i) > i$ and both Λ_{H_i} and edge density of H_i are within $1/i$ from a_i. By passing to a subsequence we can additionally assume that the k-graphs H_i converge to some $\phi \in \mathrm{LIM}^{(k)}$. This ϕ satisfies Part 2 of Theorem 24 and thus belongs to $\mathcal{T}^{(k)}$. Thus $a = \phi(K_k^k)$ belongs to $\Pi_\infty^{(k)}$, as required.

Alternatively, by Tychonoff's theorem, $[0,1]^{\mathcal{G}^{(k)}}$ is compact. By Corollary 25, $\mathcal{T}^{(k)}$ is compact. By (26), $\Pi_\infty^{(k)}$ is a continuous image of $\mathcal{T}^{(k)}$, so it is compact too. Hence $\Pi_\infty^{(k)} \subseteq [0,1]$ is closed. ∎

6 Proof of Theorem 2

Let $k \geq 3$. Let $\alpha < 1$ be a non-jump for k-graphs (that is, $(\alpha, \alpha+\varepsilon) \cap \Pi_\infty^{(k)} \neq \emptyset$ for every $\varepsilon > 0$). It exists by the result of Frankl and Rödl [21]. Pick m so that $\gamma := \Lambda_{K_m^k} > \alpha$. (Such m exists as the assignment $x_i = 1/m$ shows that $\Lambda_{K_m^k} \geq k!\binom{m}{k}/m^k$, which tends to 1 as $m \to \infty$.) Let $\tau := \gamma/(k(m-1))$. By Part 2 of Theorem 24, we can pick, inductively for $i = 1, 2, \ldots$, a k-graph H_i such that $\beta_i := \Lambda_{H_i}$ belongs to (α, γ) and

$$0 < \beta_i - \alpha < (\beta_{i-1} - \beta_i)\,\tau^{2k}, \qquad \text{for all } i \geq 2. \tag{29}$$

Informally speaking, we require that $\beta_1 > \beta_2 > \ldots$ tend to α rather fast.

Next, we introduce a new concept that is similar to that of a P-construction. Namely, for an infinite set $A = \{a_1 < a_2 < \ldots\} \subseteq \mathbb{N}$, an *$A$-configuration* is a k-graph G that can be recursively obtained as follows. Take a partition $V = V_1 \cup \ldots \cup V_m$ of the vertex set and add $K_m^k(\!(V_1, \ldots, V_m)\!)$ to the edge set (that is, add all k-sets that intersect every part in at most one vertex). Inside V_2 put some blow-up of H_{a_1}. Inside V_1 put any A'-configuration, where $A' := A \setminus \{a_1\}$.

Note that we allow a part to be everything, e.g. we allow $V_1 = V$. Let $p_{A,n}$ be the maximum size of an A-configuration on n vertices. Let $\mathcal{F}_A$ consist of all k-graphs that do not embed into an A-configuration. It is routine to see that Lemmas 6–11, with the obvious modifications, apply to A-configurations as well. In particular, we have $\mathrm{ex}(n, \mathcal{F}_A) = p_{A,n}$ for all n. Let Λ_A be the limit of $p_{A,n}/\binom{n}{k}$ as $n \to \infty$; averaging shows that this ratio is non-increasing (cf Lemma 10). Thus $\Lambda_A = \pi(\mathcal{F}_A)$.

In order to show that $|\Pi_\infty^{(k)}| \geq 2^{\aleph_0}$ it is enough to show that $\Lambda_A \neq \Lambda_B$ for every pair of infinite distinct sets $A, B \subseteq \mathbb{N}$. We prove the stronger claim that $\Lambda_A > \Lambda_B$ provided

$$\min A \setminus B < \min B \setminus A, \tag{30}$$

where we agree that $\min X = \infty$ if X is empty.

Let $A = \{a_1 < a_2 < \ldots\}$, $B = \{b_1 < b_2 < \ldots\}$, and $\min A \setminus B = a_i$. Fix large ℓ and let $n \to \infty$. Take a maximum B-configuration G. Let $\mathbf{V}$ be its partition structure, defined in the obvious way. (For example, every index in $\mathbf{V}$ is of the form $(1^{(j)}, s)$ for some $j \geq 0$ and $s \in [m]$.)

If, for some $j \leq i$ and infinitely many n, the part $V_{1^{(j-1)},2}$ (that is, the second part of the j-th level of G) is empty, we remove this b_j from B. Clearly, the k-graph G remains a maximum B-configuration. Also, this does not violate (30). Thus, by passing to a subsequence of n, we can assume that $V_{1^{(j-1)},2} \neq \emptyset$ for all $j \leq i$. Furthermore, by relabelling parts (if needed), we can assume that for every $j \geq 0$

$$\min\left(|V_{1^{(j)},1}|, |V_{1^{(j)},2}|\right) \geq \max\left\{|V_{1^{(j)},h}| : h = 3, \ldots, m\right\}. \tag{31}$$

Let us show by induction on $j = 1, \ldots, \ell$ that

$$\min\left(|U_1|, |U_2|\right) \geq \tau\,|U|, \tag{32}$$

where $U := V_{1^{(j-1)}}$ and $U_h := V_{1^{(j-1)},h}$ for $h \in [m]$. Since $G' := G[U]$ is a maximum $\{b_j, b_{j+1}, \ldots\}$-configuration on at least $\tau^{j-1}n$ vertices, its edge density is, for example, at least $\gamma + o(1)$. The argument of Lemma 11 shows that $\delta(G') \geq (\gamma + o(1))\binom{|U|-1}{k-1}$. It is impossible that $U_2 = U$ for otherwise $\Lambda_{G'} \leq \Lambda_{H_{b_j}} < \gamma$ and G' cannot be maximum for large n. This, our assumption that $U_2 \neq \emptyset$ and (31) imply that both U_1 and U_2 are non-empty. Let $h = 1$ or 2. It is impossible that $|U_h| > (1-\gamma/k)|U|$ for otherwise a vertex x of $U_{3-h} \neq \emptyset$ has too small G'-degree as every edge of G'_x has at least one other vertex in $U \setminus U_h$. By (31) we get that $|U_{3-h}| \geq (|U| - |U_h|)/(m-1) \geq \tau\,|U|$. This proves (32).

Recall that $i \in \mathbb{N}$ is defined by $a_i = \min A \setminus B$. Let $U_1 \cup \ldots \cup U_m$ be the partition of $U := V_{1^{(i-1)}}$ in the B-configuration G. Let G' be obtained from G by replacing $G[U_2]$ with a maximum blow-up of H_{a_i} (instead of H_{b_i}) and replacing $G[U_1]$ with the $\{a_j : j > i\}$-configuration that has the same partition structure as the $\{b_j : j > i\}$-configuration $G[U_1]$. Clearly, G' is an A-configuration. Since $|U_2| \geq \tau|U|$ by (32), the change inside U_2 increases the number of edges by at least $(\beta_{a_i} - \beta_{a_i+1})\tau^k\binom{u}{k} + o(n^k)$, where $u := |U|$. On the other hand, when we modify $G[U_1]$, we replace, for $j > i$, a blow-up of H_{b_j} by another blow-up whose density is at least $\alpha + o(1)$. Let n_j be the number of vertices in this part. By (32), $n_j \leq (1-\tau)^{j-i}|U|$ for all $j \leq \ell$. Thus

$$\begin{aligned}
|G[U_1]| - |G'[U_1]| &\leq \sum_{j=i+1}^{\ell-1} (\beta_{b_j} - \alpha)\binom{n_j}{k} + \sum_{j \geq \ell}\binom{n_j}{k} + o(n^k)\\
&\leq (\beta_{a_i+1} - \alpha)\binom{u}{k}\sum_{j=i+1}^{\ell-1}(1-\tau)^{(j-i)k} + \binom{n_\ell + n_{\ell+1} + \ldots}{k} + o(n^k)\\
&\leq \left((\beta_{a_i+1} - \alpha)\tau^{-k} + (1-\tau)^{(\ell-i)k}\right)\binom{u}{k}.
\end{aligned}$$

This is strictly less than $(\beta_{a_i} - \beta_{a_i+1})\tau^k\binom{u}{k}$ by (29) (and since $\ell = \ell(A,B)$ is large). Thus $|G'| \geq |G| + \Omega(n^k)$ and indeed $\Lambda_A > \Lambda_B$, finishing the proof of Theorem 2. ∎

7 Concluding Remarks

If we consider graphs (the case $k = 2$), then the Stability Theorem of Erdős [15] and Simonovits [42] answers the question about the possible asymptotic structure of maximum $\mathcal{F}$-free graphs. However, if we need a more precise answer, then the picture is much more complicated and many questions remain open, including the general *inverse problem* of describing graphs that are maximum $\mathcal{F}$-free for some family $\mathcal{F}$ (see e.g. [42, 43, 44]). The situation with extremal problems for digraphs and multigraphs is similar (see e.g. [6, 7, 9, 40]).

Although very few instances of the hypergraph Turán problem have been solved, there is a variety of constructions giving best known lower bounds. So it is likely that $\Pi^{(k)}_{\text{fin}}$ contains many further elements in addition to the values given by Corollary 4. For example, we do not know if there is a pattern P that gives the same (or better) lower bound $\pi(\{K_5^4\}) \geq \frac{11}{16}$ as the construction of Giraud [24] (see also [12] for generalisations). Roughly speaking, Giraud's construction takes an arbitrary 2-colouring of vertices and pairs of vertices (with an optimal colouring of pairs being quasi-random) and decides if a quadruple X is an edge depending on the colouring induced by X. It would be interesting to decide if Corollary 4 can be extended to cover constructions of this type.

In the special case when E consists of simple k-sets and $R = \emptyset$, Λ_P is equal by Lemma 14 to the well-studied Lagrangian of the k-graph E, see e.g. [2]. Thus Corollary 4 implies that every value of the Lagrangian belongs to $\Pi^{(k)}_{\text{fin}}$, answering a question of Baber and Talbot [2, Question 29].

One can show that every proper pattern $P = (m, E, R)$ with $R \neq \emptyset$ is *complex*, meaning that the number of non-isomorphic s-vertex subgraphs in a large maximum P-construction grows faster than any polynomial of s. Indeed, by Lemma 17 for every ℓ there is a P-construction F with the partition structure $\mathbf{V}$ which is *ℓ-rigid*, meaning that for every $\mathbf{i} \in R^s$ with $s \leq \ell$ the induced P-construction $F[V_{\mathbf{i}}]$ is rigid. Additionally, we can assume that $|V_{\mathbf{i}}| \geq k$ for each legal $\mathbf{i}$ of length at most $\ell + 1$. Thus if we add any $n - v(F)$ vertices, the new k-graph F' is still ℓ-rigid by Lemma 18. There are at least ℓ different parts at the bottom ℓ levels for placing these extra vertices. The rigidity implies that the number of pairwise non-isomorphic k-graphs F' with n vertices that we can obtain this way is at least $\binom{n-v(F)+\ell-1}{\ell-1}$ (the number of solutions to $n - v(F) = x_1 + \ldots + x_\ell$ in non-negative integers) divided by $\ell!$. Moreover, each such F' will appear an an induced subgraph in every large maximum P-construction by Lemmas 11 and 15. Since ℓ can be chosen arbitrarily large, P is indeed complex. Thus Theorem 3 answers the question of Falgas-Ravry and Vaughan [18, Question 4.4] to solve an explicit Turán problem with a complex extremal configuration (if one agrees that the family $\mathcal{F}$ in Theorem 3 is "explicit").

Let $\Pi^{(k)}_m$ consist of all possible Turán densities $\pi(\mathcal{F})$ where $\mathcal{F}$ is a family consisting of at most m forbidden k-graphs.

Question 26 (Baber and Talbot [2]) *Let $k \geq 3$. Which of the following trivial inclusions $\Pi^{(k)}_1 \subseteq \Pi^{(k)}_2 \subseteq \ldots \subseteq \Pi^{(k)}_i \subseteq \ldots \subseteq \Pi^{(k)}_{\text{fin}}$ are strict?*

It is open even whether $\Pi_1^{(k)} = \Pi_{\text{fin}}^{(k)}$ for $k \geq 3$.

Question 27 (Jacob Fox (personal communication)) *Does $\Pi_{\text{fin}}^{(k)}$ contain a transcendental number?*

Since there are only countably many algebraic numbers, Theorem 2 implies that $\Pi_\infty^{(k)}$ has a transcendental number for every $k \geq 3$.

Question 28 (Frank, Peng, Rödl, and Talbot [20]) *Let $k \geq 3$. Is there $\alpha_k < 1$ such that no value in $(\alpha_k, 1)$ is a jump for k-graphs?*

Note that by Proposition 1 the last condition is equivalent to $\Pi_\infty^{(k)} \supseteq [\alpha, 1]$. It is still open if $\Pi_\infty^{(k)}$ contains some interval of positive length for $k \geq 3$. On the other hand, the arsenal of tools for proving that some real does *not* belong to $\Pi_\infty^{(k)}$ is very limited for $k \geq 3$. In addition to the old result of Erdős [14] that $\Pi_\infty^{(k)} \cap (0, k!/k^k) = \emptyset$, the only other such result is by Baber and Talbot [1] that $(0.2299, 0.2315) \cap \Pi_\infty^{(3)} = \emptyset$. The proof in [1] uses flag algebras and is computer-generated.

Hatami and Norine [25] showed that the question whether a given linear inequality in subgraph densities is always valid is undecidable.

Question 29 *Is the validity of $\pi(\mathcal{F}) \leq \alpha$ decidable, where the input is a finite family $\mathcal{F}$ of k-graphs and a rational number α?*

A related open question is whether every true inequality $\pi(\mathcal{F}) \leq \alpha$ admits a finite proof in Razborov's Cauchy-Schwarz calculus [34, 35] (see also [25, Appendix A]).

If $k = 2$, then the answer to Question 29 is in the affirmative by the Erdős-Stone-Simonovits Theorem [17, 16]. Brown, Erdős, and Simonovits [6] obtained a positive solution to the version of Question 29 for the class of directed multigraphs.

As we have already mentioned, our proof of Theorem 3 relies on the Strong Removal Lemma. So the size of the obtained family $\mathcal{F}$ is huge (even for small concrete P). This is in contrast to many previous results and conjectures that forbid very few hypergraphs. The main place in our proof that makes $|\mathcal{F}|$ huge is the application of the Removal Lemma in the proof of Lemma 22. If, for some concrete P, Lemma 22 can be deduced in an alternative way, then one might be able to obtain an explicit and reasonably sized $\mathcal{F}$ for which Theorem 3 holds for all large n. (Note that we did not try to optimise our other lemmas for the sake of brevity and generality.) So, some of our results and techniques might be useful for small forbidden families as well. Also, the new ideas introduced for proving Theorem 3 (in particular the method of Lemma 20) might be applicable to other instances of the Turán problem.

Acknowledgements

The author thanks Zoltán Füredi for helpful discussions and the anonymous referee for the comments that greatly improved the presentation of this paper.

References

[1] R. Baber and J. Talbot. Hypergraphs do jump. *Combin. Probab. Computing*, 20:161–171, 2011.

[2] R. Baber and J. Talbot. New Turán densities for 3-graphs. *Electronic J. Combin.*, 19:21pp., 2012.

[3] J. Balogh. The Turán density of triple systems is not principal. *J. Combin. Theory* (A), 100:176–180, 2002.

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

[5] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219:1801–1851, 2008.

[6] W. G. Brown, P. Erdős, and M. Simonovits. Algorithmic solution of extremal digraph problems. *Trans. Amer. Math. Soc.*, 292:421–449, 1985.

[7] W. G. Brown, P. Erdős, and M. Simonovits. Inverse extremal digraph problems. In *Finite and infinite sets, Vol. I, II (Eger, 1981)*, volume 37 of *Colloq. Math. Soc. János Bolyai*, pages 119–156. North-Holland, Amsterdam, 1984.

[8] W. G. Brown and M. Simonovits. Digraph extremal problems, hypergraph extremal problems and the densities of graph structures. *Discrete Math.*, 48:147–162, 1984.

[9] W. G. Brown and M. Simonovits. Extremal multigraph and digraph problems. In *Paul Erdős and his mathematics, II (Budapest, 1999)*, volume 11 of *Bolyai Soc. Math. Stud.*, pages 157–203. János Bolyai Math. Soc., Budapest, 2002.

[10] F. Chung and R. L. Graham. *Erdős on Graphs: His Legacy of Unsolved Problems.* A.K.Peters, Wellesley, 1998.

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

[12] D. de Caen, D. L. Kreher, and J. Wiseman. On constructive upper bounds for the Turán numbers $T(n, 2r+1, 2r)$. *Congr. Numer.*, 65:277–280, 1988.

[13] G. Elek and B. Szegedy. A measure-theoretic approach to the theory of dense hypergraphs. *Adv. Math.*, 231:1731–1772, 2012.

[14] P. Erdős. On extremal problems of graphs and generalized graphs. *Israel J. Math.*, 2:183–190, 1964.

[15] P. Erdős. Some recent results on extremal problems in graph theory. Results. In *Theory of Graphs (Internat. Sympos., Rome, 1966)*, pages 117–123 (English); pp. 124–130 (French). Gordon and Breach, New York, 1967.

[16] P. Erdős and M. Simonovits. A limit theorem in graph theory. *Stud. Sci. Math. Hungar.*, pages 51–57, 1966.

[17] P. Erdős and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.

[18] V. Falgas-Ravry and E. R. Vaughan. Applications of the semi-definite method to the Turán density problem for 3-graphs. *Combin. Probab. Computing*, 22:21–54, 2013.

[19] P. Frankl and Z. Füredi. Extremal problems whose solutions are the blowups of the small Witt-designs. *J. Combin. Theory* (A), 52:129–147, 1989.

[20] P. Frankl, Y. Peng, V. Rödl, and J. Talbot. A note on the jumping constant conjecture of Erdős. *J. Combin. Theory* (B), 97:204–216, 2007.

[21] P. Frankl and V. Rödl. Hypergraphs do not jump. *Combinatorica*, 4:149–159, 1984.

[22] Z. Füredi. Turán type problems. In *Surveys in Combinatorics*, volume 166 of *London Math. Soc. Lecture Notes Ser.*, pages 253–300. Cambridge Univ. Press, 1991.

[23] Z. Füredi, O. Pikhurko, and M. Simonovits. The Turán density of the hypergraph $\{abc, ade, bde, cde\}$. *Electronic J. Combin.*, 10:7pp., 2003.

[24] G. R. Giraud. Remarques sur deux problèmes extrémaux. *Discrete Math.*, 84:319–321, 1990.

[25] H. Hatami and S. Norine. Undecidability of linear inequalities in graph homomorphism densities. *J. Amer. Math. Soc.*, 24:547–565, 2011.

[26] G. O. H. Katona, T. Nemetz, and M. Simonovits. On a graph problem of Turán (In Hungarian). *Mat. Fiz. Lapok*, 15:228–238, 1964.

[27] P. Keevash. Hypergraph Turán problem. In R. Chapman, editor, *Surveys in Combinatorics*, pages 83–140. Cambridge Univ. Press, 2011.

[28] L. Lovász. *Large Networks and Graph Limits.* Colloquium Publications. Amer. Math. Soc, 2012.

[29] L. Lovász and B. Szegedy. Limits of dense graph sequences. *J. Combin. Theory* (B), 96:933–957, 2006.

[30] T. S. Motzkin and E. G. Straus. Maxima for graphs and a new proof of a theorem of Turán. *Can. J. Math.*, 17:533–540, 1965.

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

[32] D. Mubayi and O. Pikhurko. Constructions of non-principal families in extremal hypergraph theory. *Discrete Math.*, 308:4430–4434, 2008.

[33] Y. Peng and C. Zhao. Generating non-jumping numbers recursively. *Discrete Applied Math.*, 156:1856–1864, 2008.

[34] A. Razborov. Flag algebras. *J. Symb. Logic*, 72:1239–1282, 2007.

[35] A. Razborov. On 3-hypergraphs with forbidden 4-vertex configurations. *SIAM J. Discr. Math.*, 24:946–963, 2010.

[36] A. A. Razborov. On the Fon-der-Flaass interpretation of extremal examples for Turán's $(3, 4)$-problem. *Proc. Steklov Inst. Math.*, 274:247–266, 2011. Translated from *Trudy Mat. Inst. Steklova.*

[37] V. Rödl and M. Schacht. Generalizations of the Removal Lemma. *Combinatorica*, 29:467–501, 2009.

[38] I. Z. Ruzsa and E. Szemerédi. Triple systems with no six points carrying three triangles. In A. Hajnal and V. Sós, editors, *Combinatorics II*, pages 939–945. North Holland, Amsterdam, 1978.

[39] A. Sidorenko. The maximal number of edges in a homogeneous hypergraph containing no prohibited subgraphs. *Math Notes*, 41:247–259, 1987. Translated from *Mat. Zametki.*

[40] A. Sidorenko. Boundedness of optimal matrices in extremal multigraph and digraph problems. *Combinatorica*, 13:109–120, 1993.

[41] A. Sidorenko. What we know and what we do not know about Turán numbers. *Graphs Combin.*, 11:179–199, 1995.

[42] M. Simonovits. A method for solving extremal problems in graph theory, stability problems. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 279–319. Academic Press, 1968.

[43] M. Simonovits. Extermal graph problems with symmetrical extremal graphs. Additional chromatic conditions. *Discrete Math.*, 7:349–376, 1974.

[44] M. Simonovits. Extremal graph problems and graph products. In *Studies in pure mathematics*, pages 669–680. Birkhäuser, Basel, 1983.

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