Polynomial to exponential transition in Ramsey theory

Dhruv Mubayi* Alexander Razborov†

February 21, 2020

Abstract

Given $s \geq k \geq 3$, let $h^{(k)}(s)$ be the minimum t such that there exist arbitrarily large k-uniform hypergraphs H whose independence number is at most polylogarithmic in the number of vertices and in which every s vertices span at most t edges. Erdős and Hajnal conjectured (1972) that $h^{(k)}(s)$ can be calculated precisely using a recursive formula and Erdős offered \$500 for a proof of this. For $k = 3$ this has been settled for many values of s including powers of three but it was not known for any $k \geq 4$ and $s \geq k + 2$.

Here we settle the conjecture for all $s \geq k \geq 4$. We also answer a question of Bhat and Rödl by constructing, for each $k \geq 4$, a quasirandom sequence of k-uniform hypergraphs with positive density and upper density at most $k!/(k^k - k)$. This result is sharp.

MSC classification codes: 05D10 (primary), 05C35, 05C55, 05C65 (secondary)

1 Introduction

Write $K_N^{(k)}$ for the complete k-uniform hypergraph (henceforth k-graph) on N vertices. The *Ramsey number* $r_k(s, n)$ is the minimum N such that every red/blue coloring of the edges of $K_N^{(k)}$ contains a monochromatic red copy of $K_s^{(k)}$ or a monochromatic blue copy of $K_n^{(k)}$. In order to shed more light on the growth rate of these classical Ramsey numbers, Erdős and Hajnal [12] in 1972 considered the following more general parameter.

Definition 1.1. For integers $2 \leq k < s < n$ and $2 \leq t \leq \binom{s}{k}$, let $r_k(s, t; n)$ be the minimum N such that every red/blue coloring of the edges of $K_N^{(k)}$ contains a monochromatic blue copy of $K_n^{(k)}$ or has a set of s vertices which contains at least t red edges.

Note that $r_k(s, \binom{s}{k}; n) = r_k(s, n)$ so $r_k(s, t; n)$ includes classical Ramsey numbers. In addition, the case $(k, s, t, n) = (k, k + 1, k + 1, k + 1)$ was investigated in relation to the Erdős-Szekeres theorem

*Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL, 60607 USA. Research partially supported by NSF grants DMS-1300138 and 1763317. Email: mubayi@uic.edu

†Departments of Mathematics and Computer Science, University of Chicago, IL, 60637, USA, razborov@math.uchicago.edu and Steklov Mathematical Institute, Moscow, 119991, Russia, razborov@mi.ras.ru.

and Ramsey numbers of ordered tight paths as well as to high dimensional tournaments by several researchers [9, 10, 16, 27, 29, 30]; the very special case $(3,4,3,n)$ has connections to quasirandom hypergraph constructions [3, 22, 25, 26].

The main conjecture of Erdős and Hajnal [12] for $r_k(s,t;n)$ is that, as t grows from 1 to $\binom{s}{k}$, there is a well-defined value $t_1 = h_1^{(k)}(s)$ at which $r_k(s,t_1-1;n)$ is polynomial in n while $r_k(s,t_1;n)$ is exponential in a power of n, another well-defined value $t_2 = h_2^{(k)}(s)$ at which it changes from exponential to double exponential in a power of n and so on, and finally a well-defined value $t_{k-2} = h_{k-2}^{(k)}(s) < \binom{s}{k}$ at which it changes from twr_{k-2} to twr_{k-1} in a power of n. They were not able to offer a conjecture as to what $h_i^{(k)}(s)$ is in general, except when $i=1$ or when $s=k+1$.

The problem of determining $r_k(k+1,t;n)$ for $t=2$ and $t=3$ has essentially been solved. For general t, the methods of Erdős and Rado [13] show that there exists $c=c(k,t)>0$ such that $r_k(k+1,t;n) \le \mathrm{twr}_{t-1}(n^c)$ for $3 \le t \le k$. Erdős and Hajnal conjectured that this gives the correct tower growth rate for $r_k(k+1,t;n)$. When $k \ge 6$, the first author and Suk [31] settled their conjecture in almost all cases in a strong form.

Perhaps the main open problem about $r_k(s,t;n)$ posed by Erdős and Hajnal [12] was to determine the value of $t_1 = h_1^{(k)}(s)$; namely the value of t at which $r_k(s,t;n)$ transitions from polynomial to super polynomial growth. This is the problem we address in this paper. The following function plays an important role.

Definition 1.2. Given positive integers s,k, call a partition $s_1+\cdots+s_k=s$ *nontrivial* if $0 \le s_i < s$ for each i. For $0 \le s < k$, let $g_k(s)=0$ and for $s \ge k \ge 3$, let $g_k(s)$ be the maximum of

$$\sum_{i=1}^{k} g_k(s_i) + \prod_{i=1}^{k} s_i$$

where the maximum is taken over all nontrivial partitions $s_1+\cdots+s_k=s$.

We will interpret $g_k(s)$ as the maximum number of edges in the s-vertex k-graph obtained by first partitioning s vertices into k parts, taking all edges that intersect all parts, and then recursing this construction within each part. Erdős and Hajnal commented without proof that it is easy to see that $g_k(s)$ is achieved by taking a partition that is as equitable as possible. We will prove this in the Appendix, and also prove an asymptotic version of this fact later (see (17)). As an easy exercise, this implies that

$$g_k(s) = (1+o(1))\frac{k!}{k^k-k}\binom{s}{k} \qquad k \text{ is fixed, } s \to \infty. \tag{1}$$

Erdős and Hajnal proved that $r_k(s,g_k(s);n)$ is polynomial in n for all fixed $s > k \ge 3$. In other words, they showed that every N-vertex k-graph ($k \ge 3$ fixed) in which every s-set spans at most $g_k(s)-1$ edges has independence number at least N^ϵ where $\epsilon > 0$ depends only on s,k. Therefore

$$h_1^{(k)}(s) \ge g_k(s)+1.$$

They conjectured the following for which Erdős later offered \$500 (see [28] page 21 and [5] Problem (85)).

Conjecture 1.3 (Erdős-Hajnal)**.** *Fix $s \geq k \geq 3$. Then $h_1^{(k)}(s) = g_k(s) + 1$, or equivalently, $r_k(s, g_k(s) + 1; n)$ is at least exponential in a power of n.*

For $k = 3$, Erdős and Hajnal [12] proved that Conjecture 1.3 follows from the following conjecture.

Conjecture 1.4 (Erdős-Hajnal)**.** *For every edge-coloring of the complete graph with vertex set $[n]$ by three colors I, II, III, the number of triangles $\{a, b, c\}$ with $a < b < c$ for which $\{a, b\}$ has color I, $\{b, c\}$ has color II, and $\{a, c\}$ has color III is at most $g_3(n)$.*

Conlon-Fox-Sudakov [6] connected Conjecture 1.4 to the maximum number $T(s)$ of directed triangles in an s-vertex tournament (It is worth noting that the hypergraphs in Conjecture 1.4 were also considered in [33] due to their connection to hypergraph Turán theory.) They determined $T(s)$ exactly and observed that this also settles Conjecture 1.4 for many values of s including powers of 3. Consequently, their approach gave a solution to Conjecture 1.3 when $k = 3$ and these s-values; they also proved that $h_1^{(3)}(s) = (1/4)\binom{s}{3} + O(s \log s)$. However, their method using $T(s)$ does not apply for any $k > 3$ as it does not capture the recursive structure from Definition 1.2 needed to prove Conjecture 1.3. Indeed, the set of extremal configurations for $T(s)$ consists of *all* (out-)regular tournaments; the recursive construction is just one (and unnecessarily complicated) example in this class. Thus Conjecture 1.3 was known only when $s = k + 1$ and when $k = 3$ and s is as described above. In fact, Erdős and Hajnal stated in [12] that they were much less certain about Conjecture 1.3 when $k \geq 4$ than when $k = 3$.

In this paper we prove Conjecture 1.3 for all $k \geq 4$.

Theorem 1.5. *$h_1^{(k)}(s) = g_k(s) + 1$ for all $s \geq k \geq 4$.*

Our method also answers a question posed by Bhat and Rödl [3] about quasirandom sequences. The density of a k-graph $H = (V, E)$ is $d(H) = |E|/\binom{|V|}{k}$. Let $\mathcal{H} = \{H_n\}_{n=1}^{\infty}$ be a sequence of k-graphs with $H_n = (V_n, E_n)$ such that $|V_n| \to \infty$ as $n \to \infty$. Define the density $d(\mathcal{H})$ of $\mathcal{H}$ as $d(\mathcal{H}) = \lim_{n\to\infty} d(H_n)$ (we only consider sequences where the limit exists) and the upper density

$$\overline{d}(\mathcal{H}) \stackrel{\text{def}}{=} \lim_{s\to\infty} \max_{n} \max_{S\in\binom{V_n}{s}} d(H_n[S])$$

(note that for any fixed s, $H_n[S]$ can take only finitely many values, up to isomorphism). One can show by a simple averaging argument that $\overline{d}(\mathcal{H})$ exists.

Definition 1.6. A k-graph sequence $\{H_n\}_{n=1}^{\infty}$ is ρ-quasirandom if for every $\epsilon > 0$ there exists n_0 such that for $n > n_0$, every $W \subset V = V(H_n)$ with $|W| \geq \epsilon|V|$ satisfies $d(H_n[W]) \in [\rho(1-\epsilon),\ \rho(1+\epsilon)]$.

An important result of Erdős [11] states that every k-graph sequence with positive density contains arbitrarily large complete k-partite subgraphs and hence has upper density at least $k!/k^k$ (the case $k = 2$ was done earlier by Kövári-Sós-Turán [23] and by Erdős (see [11])); the value $k!/k^k$ cannot be increased as shown by complete k-partite k-graphs. This is a fundamental tool for hypergraph problems, and shows that every $\rho \in (0, k!/k^k)$ is a "jump" for k-graphs (see [17] for background on jumps).

Bhat and Rödl [3] improved this result of Erdős in the quasirandom setting: they showed that for each $k \geq 3$ and $\rho > 0$, every ρ-quasirandom k-graph sequence has upper density at least $k!/(k^k - k)$, thus showing that every $\rho \in (0, k!/(k^k - k))$ is a jump in this setting. It is well-known that $k!/(k^k - k)$ cannot be increased for $k = 3$ (the simplest example is to take the 3-graph of all cyclic triples in a random tournament) and Bhat and Rödl asked whether the same is true for $k \geq 4$. We answer this positively, showing that the result in [3] is sharp for all $k \geq 3$.

Theorem 1.7. *For each $k \geq 4$, there exists $\rho > 0$ and a ρ-quasirandom k-graph sequence with upper density $k!/(k^k - k)$.*

We note that our proof of Theorem 1.7 yields $\rho = k^{-\Omega(k^2)}$ which is much smaller than $k!/(k^k - k)$ and it remains open to prove the theorem with $\rho = k!/(k^k - k)$ (for $k = 3$ this is true).

2 Reduction to inducibility

As mentioned in the introduction, Erdős and Hajnal showed that the $k = 3$ case of Conjecture 1.3 follows from Conjecture 1.4, which asks for the maximum number of rainbow colored triangles (with some additional properties) in an edge-colored ordered graph. This is an example of a question about the inducibility of colored, directed structures. In fact, Erdős and Hajnal observed that Conjecture 1.4 could be replaced by another slightly different question about inducibility (where we use only two colors and count certain 2-colored triangles) and, as mentioned earlier, Conlon, Fox and Sudakov [6] considered yet another inducibility problem, namely the determination of $T(s)$.

Our approach to Conjecture 1.3 is to formulate a novel question about the inducibility of colored directed structures and solve it exactly. It is perhaps interesting that the "universal" character of the structure we consider below allows us to get around many technical difficulties plaguing previous research on inducibility.

Theorem 2.1. **(Main Result)** *Let $s \geq k \geq 4$ and R be an arbitrary k-vertex tournament whose edges are colored with the $\binom{k}{2}$ distinct colors from $\binom{[k]}{2}$. Then the number of copies of R in any s-vertex tournament whose edges are colored from $\binom{[k]}{2}$ is at most $g_k(s)$.*

We immediately get Theorem 1.5 as a consequence.

Proof of Theorem 1.5. Fix $s \geq k \geq 4$. We are to show that $h_1^{(k)}(s) \leq g_k(s) + 1$. In other words: there exists $C = C(k) > 0$ and, for all $N > k$, an N-vertex k-graph H with $\alpha(H) \leq C \log N$ such that every s vertices of H span at most $g_k(s)$ edges. Fix a k-vertex tournament R whose edges are colored with $\binom{k}{2}$ distinct colors. Next consider the random N-vertex tournament $T = T_N$ whose edges are randomly colored with the same $\binom{k}{2}$ colors; thus, each pair gets a particular orientation and color with probability $p = 1/((k-1)k)$. Now form the k-graph $H = H(T) = (V, E)$ with $V = V(T)$ and $E = \{K \subset V : H[K] \cong R\}$. In other words, the edges of H correspond to copies of R. By Theorem 2.1, every s vertices of H span at most $g_k(s)$ edges. On the other hand, the probability that a given k-set of vertices in H induces a copy of R is $k!p^{\binom{k}{2}} > 0$. Hence the expected number of t-sets in H that are independent is at most $\binom{N}{t}2^{-\epsilon t^2}$ for appropriate $\epsilon = \epsilon(k) > 0$. Indeed, given any t-set A, pick up in it $\ell \geq \Omega(t^2/k^2)$ k-subsets $B_1, \ldots, B_\ell$ such that $|B_i \cap B_j| \leq 1$ whenever $i \neq j$, and notice that the events "B_i spans a copy of R" are mutually independent. This expectation is less than one as long as $t > C \log N$ and $C = C(k)$ is sufficiently large. □

Remark 2.2. For the remaining case $k = 3$, we believe that the Erdős-Hajnal conjecture still holds but it may require new techniques and ideas: many crucial calculations in this paper completely fall apart.

Remark 2.3. As mentioned above, there is nothing specific about the kind of combinatorial structures we are considering here, and Theorem 1.5 is implied by results analogous to Theorem 2.1 for arbitrary structures. For example, [1] gives Theorem 1.5 for $k = 5,\ s = 5^t$, [37] gives it for all k sufficiently large and $s \leq 2^{\sqrt{k}}$, the well-known Pippenger-Golumbic conjecture [32] about the inducibility of C_k would imply it for $k \geq 5,\ s = k^t$, and the conjecture from [36] about $\vec{C}_4$ would imply it for $k = 4,\ s = 4^t$. See [2, 4, 14, 15, 18, 19, 21, 24] for results about inducibility for other structures.

Next we show how our approach also answers the question of Bhat and Rödl about ρ-quasirandom hypergraph sequences. It is convenient to use the following theorem from [7] which is a hypergraph generalization of the Chung-Graham-Wilson characterization of graph quasirandomness. In what follows M_k is a specific linear k-graph with $v = k2^{k-1}$ vertices and $e = 2^k$ edges (see [7] or [25] for the precise definition); in particular $M_2 = C_4$. We write the result from [7] in the language of hypergraph sequences.

Theorem 2.4 (Conlon-Han-Person-Schacht [7])**.** *Fix $k \geq 2$, $0 < \rho < 1$, and a sequence of k-graphs $\mathcal{H} = \{H_n\}_{n=1}^{\infty}$ of density ρ each with $H_n = (V_n, E_n)$ and $|V_n| \to \infty$ as $n \to \infty$. Then $\mathcal{H}$ is ρ-quasirandom iff the number of (labeled) copies of M_k in H_n is $|V_n|^{k2^{k-1}}\rho^{2^k}(1 + o(1))$ as $n \to \infty$.*

Proof of Theorem 1.7. We use the proof of Theorem 1.5 above to construct the desired sequence. Using the notation there, for each $k \geq 4$, let $\rho = k!((k-1)k)^{-\binom{k}{2}}$ be the probability that a k-set induces a copy of R. For $n \geq 1$, let $\epsilon_n = 1/n$. Standard probabilistic arguments together with the construction of H in Theorem 1.5 imply that there exists a k-graph $H_n = (V_n, E_n)$ whose edge set comprises copies of R such that $|V_n| \to \infty$ and the number of copies of M_k in H_n is $|V_n|^{k2^{k-1}}\rho^{2^k}(1 \pm \epsilon_n)$. Indeed, since M_k is linear (meaning that every two edges of M_k share at most one vertex) the expected number of (labeled) copies of M_k in H_n is $|V_n|^{k2^{k-1}}\rho^{2^k}$ and Chebyshev's inequality implies that there is an H_n where the number of copies of M_k in H_n is $|V_n|^{k2^{k-1}}\rho^{2^k}(1 \pm \epsilon_n)$. Now let $\mathcal{H} = \{H_n\}_{n=1}^{\infty}$. We have just shown that the number of copies of M_k in H_n is $|V_n|^{k2^{k-1}}\rho^{2^k}(1+o(1))$ so Theorem 2.4 implies that $\mathcal{H}$ is ρ-quasirandom. On the other hand, for each $s, n > 0$, $k \geq 4$ and $S \in \binom{V_n}{s}$ we have $d(H_n[S]) \leq g_k(s)/\binom{s}{k}$ by Theorem 2.1. Consequently, $\overline{d}(\mathcal{H}) \leq \lim_{s\to\infty} g_k(s)/\binom{s}{k} \leq k!/(k^k - k)$. □

3 Proof of asymptotic result

Recall that the *inducibility* $i(R)$ is defined as

$$i(R) \stackrel{\text{def}}{=} \lim_{s\to\infty} \max_{|V(H)|=s} \frac{i(R; H)}{\binom{s}{k}},$$

where $i(R;H)$ is the number of copies of R in an s-vertex $\binom{[k]}{2}$-colored tournament H. In this section we prove the following.

Theorem 3.1. *Let $k \geq 4$ and R be an arbitrary k-vertex tournament whose edges are colored with the $\binom{k}{2}$ distinct colors from $\binom{[k]}{2}$. Then*

$$i(R) = \frac{k!}{k^k - k}$$

(which, by (1) *is equal to* $\lim_{s\to\infty} \frac{g_s(k)}{\binom{s}{k}}$*).*

The proof of Theorem 3.1 is much cleaner than the proof of our main result $\max_{|V(H)|=s} i(R;H) = g_k(s)$ presented in Section 4 since it avoids dealing with unnecessary details about the number of vertices. It also gives the reader the overall structure of our argument. Moreover, as we will show in Corollary 3.2, the asymptotic result in Theorem 3.1 immediately implies an exact result whenever s is a power of k.

To make our argument both clean and rigorous, we use the language of Flag Algebras. But since in order to prove Theorem 2.1 we will have to "discretize" it anyway (so Theorem 3.1 is sort of a warm-up), we skip the traditional crash course in Flag Algebras and assume a certain degree of familiarity with the method. The reader interested only in the end result can safely proceed to Section 4 (or, if willing to believe that all this can be made completely rigorous, follow the proof on the intuitive level).

Proof of Theorem 3.1. Let T_k be the theory [34, §2] of $\binom{[k]}{2}$-colorings of edges of a complete graph, and let T be the disjoint union of T_k and the theory $T_{\text{Tournament}}$ of tournaments. Let $R \in \mathcal{M}_k[T]$ be any model with $V(R) = [k]$ such that its restriction to T_k is the canonical (that is, the edge (i,j) is colored with the color $\{i,j\}$) model, and let Γ be the underlying tournament. As always, we denote by 1 the (only) type [34, §2.1] of size 1.

For a color $c \in \binom{[k]}{2}$, there are two 1-flags in $\mathcal{F}_2^1$ colored by c: α_c (in which the distinguished vertex is the tail) and β_c (distinguished = head). Let

$$S_i \stackrel{\text{def}}{=} \sum_{j \in N_\Gamma^+(i)} \alpha_{\{i,j\}} + \sum_{j \in N_\Gamma^-(i)} \beta_{\{i,j\}}$$

(this is an element of $\mathcal{A}^1$) and note that $\sum_i S_i = 1$. Define also

$$P_i \stackrel{\text{def}}{=} \prod_{j \in N_\Gamma^+(i)} \alpha_{\{i,j\}} \cdot \prod_{j \in N_\Gamma^-(i)} \beta_{\{i,j\}}. \tag{2}$$

Let us now fix $\phi \in \text{Hom}^+(\mathcal{A}^0[T], \mathbb{R})$ (see [34, Definition 5]) maximizing $\phi(R)$ [34, §4.1] and let

$$a_k \stackrel{\text{def}}{=} \frac{k!}{k^k - k} \left(= \frac{(k-1)!}{k^{k-1} - 1} \right).$$

Our goal is to prove that

$$\phi(R) \leq a_k, \tag{3}$$

and we can assume w.l.o.g. that $\phi(R) \geq a_k$.

Let $\boldsymbol{\phi^1}$ be the distribution over $\mathrm{Hom}^+(\mathcal{A}^1[T], \mathbb{R})$ rooted at ϕ [34, Definition 10], and let $\mathcal{S}^1(\phi)$ be the support of this distribution. Combinatorially, $\boldsymbol{\phi^1}$ should be thought of as a uniform distribution over vertices (except that we do have any such thing as a "vertex" here). Let us study an individual element $\phi^1 \in \mathcal{S}^1(\phi)$.

Assume for simplicity that $\phi^1(S_1) \geq \phi^1(S_2) \geq \cdots \geq \phi^1(S_k)$; our goal is to bound $\phi^1(S_2)$ from above (the trivial bound is 1/2). More specifically, note first (recall that $k \geq 4$) that

$$\frac{(k-1)^{k-1}}{k^{k-1}-1} \geq \left(1 - \frac{1}{k}\right)^{k-1} \geq e^{-1} \geq 2^{2-k}.$$

Hence the equation

$$z^{k-1} + (1-z)^{k-1} = \frac{(k-1)^{k-1}}{k^{k-1}-1} \tag{4}$$

has two roots in the interval $(0, 1)$; let $z_k \in (0, 1/2)$ be the smallest one. We claim that $\phi^1(S_2) \leq z_k$.

Let $\mu_k^1(R) \in \mathcal{A}^1$ be the sum of all k possible 1-flags that can be obtained from R [34, §4.3]. Then one consequence of the extremality of ϕ (and the fact that $\phi^1 \in \mathcal{S}^1(\phi)$) is that $\phi^1(\mu_k^1(R)) = \phi(R) \geq a_k$ [34, Theorem 4.3].

On the other hand, by the AMGM inequality we have

$$\mu_k^1(R) \leq (k-1)! \sum_i P_i \leq \frac{(k-1)!}{(k-1)^{k-1}} \sum_i S_i^{k-1}, \tag{5}$$

where the partial pre-order $\leq$ on $\mathcal{A}^1$ simply means [34, Definition 6] that the inequality holds upon being evaluated by an arbitrary element of $\mathrm{Hom}^+(\mathcal{A}^1[T], \mathbb{R})$. Comparing the two,

$$\sum_i \phi^1(S_i)^{k-1} \geq \frac{(k-1)^{k-1}}{k^{k-1}-1}.$$

But under the condition $S_2 = z$ ($\leq 1/2$), the left-hand side is clearly maximized when $S_1 = 1 - z$ and $S_3 = \ldots = S_k = 0$. This gives us the claim.

We now have a measurable partition $S^1(\phi) = V_1 \dot\cup \ldots \dot\cup V_k$ according to $\arg\max_i \phi^1(S_i)$ (we resolve conflicts arbitrarily), and we want to incorporate it into our language explicitly. Let T^+ be the extension of T with vertex coloring χ in k colors. We let $p_i \in \mathcal{M}_1[T^+]$ be the one-element model in which the only vertex is colored by i, and let (i) be the corresponding type. Let $I : T \rightsquigarrow T^+$ be the interpretation [34, §2.3.3] erasing vertex coloring. We want to extend ϕ to an element $\phi^+ \in \mathrm{Hom}^+(\mathcal{A}^0[T^+], \mathbb{R})$ that respects the partition $V_1 \dot\cup \ldots \dot\cup V_k$ (we will actually need only its property $\phi^1(S_i) \leq z_k$ for $\phi^1 \notin V_i$). Formally, we claim the existence of ϕ^+ with the following two properties:

1. $\phi = \phi^+ \circ \pi^I$ (for the definition of algebra homomorphisms $\pi^I, \pi^{I,\sigma}, \pi^{\sigma,\eta}$ etc. see [34, §2.3]);
2. For any $\psi \in \mathcal{S}^{(i)}(\phi^+)$ and any $i' \neq i$, $\psi\left(\pi^{I,(i)}(S_{i'})\right) \leq z_k$.

Combinatorially, the existence of such an extension is entirely obvious, and the simplest way to give a rigorous definition in the language of Flag Algebras is by an explicit formula. Namely, for a type σ of the theory T^+ that has size k, we first define the "labelled density" $\phi^+(\langle\sigma\rangle)$ as

$$\phi^+(\langle\sigma\rangle) \stackrel{\text{def}}{=} \phi(\langle I(\sigma)\rangle) \cdot \mathbf{P}\left[\bigwedge_{i=1}^{k} \left(\boldsymbol{\phi^{I(\sigma)}} \circ \pi^{I(\sigma),i} \in V_i\right)\right]. \tag{6}$$

Then we let

$$\phi^+(\sigma) \stackrel{\text{def}}{=} (S_k : \mathrm{Aut}(\sigma))\phi^+(\langle\sigma\rangle).$$

It is straightforward to check that so defined ϕ^+ is an element of $\mathrm{Hom}^+(\mathcal{A}^0[T^+], \mathbb{R})$ that satisfies properties 1) and 2) above.

From now on we will often omit from the notation operators π^I and $\pi^{I,\sigma}$ (as well as ϕ^+); thus, the algebra $\mathcal{A}^{I(\sigma)}[T]$ is identified with its image under $\pi^{I,\sigma}$ in $\mathcal{A}^\sigma[T^+]$. When σ has to be specified, we write f^σ for the image of $f \in \mathcal{A}^{I(\sigma)}[T]$ in $\mathcal{A}^\sigma[T^+]$; we will be primarily interested in the case when σ has size 1, i.e. $\sigma = (i)$ for some $i \in [k]$. Thus, property 2) above simplifies to $\psi\left(S_{i'}^{(i)}\right) \leq z_k$ for any $i' \neq i$ and $\psi \in \mathcal{S}^{(i)}(\phi^+)$ etc.

For $j \neq i$, let $P_{ij} \in \mathcal{A}^1[T]$ be the product P_i with either $\alpha_{i,j}$ or $\beta_{i,j}$ removed. Then the AMGM inequality implies the bound

$$\psi\left(P_{i'j}^{(i)}\right) \leq \left(\frac{z_k}{k-2}\right)^{k-2} \left(i' \neq i,\ \psi \in \mathcal{S}^{(i)}(\phi^+)\right). \tag{7}$$

Now, R splits in T^+ as follows:

$$\pi^I(R) = R_m + R_g + R_b$$

("m, g, b" stand for "monochromatic", "good" and "bad", respectively), where R_m is the sum of m models in $I^{-1}(R)$ in which all vertices are colored in the same color, R_g is the model with $\chi = \mathrm{id}$ and R_b is the sum of all remaining models. We will estimate these three terms (evaluated by ϕ^+) separately.

The bound on R_m (that, combinatorially, is the density of monochromatic copies of R) is obtained by exploiting the extremality of ϕ one more time:

$$\phi^+(R_m) \leq \phi(R) \cdot \sum_i \phi^+(p_i)^k. \tag{8}$$

To make this rigorous, whenever $\phi^+(p_i) > 0$, we form the restriction $\pi^{p_i} : \mathcal{A}^0[T^+] \longrightarrow \mathcal{A}^0_{p_i}[T^+]$, where $\mathcal{A}^0_{p_i}$ is the localization of $\mathcal{A}^0$ by the element p_i (see [34, §2.3.2] with $\sigma = 0$) that combinatorially corresponds to the restriction on V_i. Then $\phi^+ \circ \pi^{p_i} \circ \pi^I \in \mathrm{Hom}^+(\mathcal{A}^0[T], \mathbb{R})$, hence the extremality of ϕ implies that $(\phi^+ \circ \pi^{p_i} \circ \pi^I)(R) \leq \phi(R)$. On the other hand, by unrolling definitions we see that $(\phi^+ \circ \pi^{p_i} \circ \pi^I)(R) = \frac{\phi^+(R_{m,i})}{\phi^+(p_i)^k}$, where $R_{m,i}$ is the model with $\chi \equiv i$. Multiplying by p_i^k and summing over all i gives us (8).

For bounding R_g, we let Δ be the sum of all rainbow (i.e., with bijective χ) models in $\mathcal{M}_k[T^+]$ *different* from R_g. Then we clearly have

$$R_g = k!\prod_{i=1}^{k} p_i - \Delta, \tag{9}$$

and we need to bound Δ from below. We will do it in terms of the element $\delta \in \mathcal{A}_2^0[T^+]$ which is the sum of all models M with the set of vertices $\{u, v\}$ that are *transversal* ($\chi(u) \neq \chi(v)$) and are either *miscolored* (the edge color of (u, v) is different from $\{\chi(u), \chi(v)\}$) or *disoriented* (the orientation of (u, v) is different from the orientation of $(\chi(u), \chi(v))$ in Γ), or both. Let δ_{ij} be the contribution to δ made by those models in $\mathcal{M}_2[T^+]$ for which $\{\chi(u), \chi(v)\} = \{i, j\}$. Then

$$\delta = \sum_{1\leq i<j\leq k} \delta_{ij}.$$

Now, if we extend any model in δ_{ij} to an arbitrary rainbow model in $\mathcal{M}_k[T^+]$, we will actually get a model in Δ. This implies

$$\Delta \geq \frac{k!}{2}\left(\prod_{\nu\neq i,j} p_\nu\right)\delta_{ij}$$

for any $i \neq j$ (the factor 2 in the enumerator accounts for the symmetry interchanging i and j). Multiplying this by $p_i p_j$ and summing up over all such pairs, we arrive at

$$\left(\sum_{1\leq i<j\leq k} p_i p_j\right)\Delta \geq \frac{k!}{2}\left(\prod_i p_i\right)\delta. \tag{10}$$

At this point we have to take care of the case when all but one of the p_is are equal to 0. Assuming that, say, $p_1 = 1$, we know that $\phi^1(S_2) \leq z_k$ for *any* $\phi^1 \in S^1(\phi)$. Let $R^{1,i} \in \mathcal{F}^1[T]$ be obtained from R by placing the distinguished vertex into i (so that $\mu_k^1(R) = \sum_i R^{1,i}$). Then the local version of (5) gives us

$$R^{1,2} \leq \frac{(k-1)!}{(k-1)^{k-1}} S_2^{k-1} \leq \frac{(k-1)!}{(k-1)^{k-1}} z_k^{k-1}. \tag{11}$$

On the other hand, $R = k[\![R^{1,2}]\!]_1$ (see [34, §2.2] for the averaging operator $[\![\cdot]\!]_\sigma$). This gives us

$$\phi(R) \leq \frac{k!}{(k-1)^{k-1}} z_k^{k-1} \leq a_k k\left(1 + \frac{1}{k-1}\right)^{k-1} z_k^{k-1} \leq a_k, \tag{12}$$

where for $k = 4$ the last inequality follows from the bound $z_4 \leq 0.3$, and when $k \geq 5$ it suffices to apply the trivial bound $z_k \leq 1/2$. This completes the proof of (3) when $p_1 = 1$.

Thus we can and will assume that $p_i < 1$ for all i and hence we can divide (10) by $\sum_{1\leq i<j\leq k} p_i p_j$. Comparing the result with (9), we arrive at our second estimate

$$R_g \leq k! \cdot \prod_i p_i\left(1 - \frac{\delta}{2\sum_{1\leq i<j\leq k} p_i p_j}\right) = k! \cdot \prod_i p_i\left(1 - \frac{\delta}{1 - \sum_i p_i^2}\right). \tag{13}$$

Let's now turn to upper bounding R_b. We need a few simple remarks first.

Every model $M \in \mathcal{M}_2[T]$ has a unique embedding $\alpha_M : M \longrightarrow R$; intuitively, this embedding corresponds to the "intended" vertex-coloring of M. The mapping α can be extended to $\mathcal{M}_2[T^+]$ simply by letting $\alpha_M \stackrel{\text{def}}{=} \alpha_{I(M)}$ i.e. by ignoring the vertex-coloring χ. Then δ can be described as the sum of all transversal models in $\mathcal{M}_2[T^+]$ for which $\chi \neq \alpha_M$. Let now $b^{(i)}$ be the sum of all (i)-flags that have the form (M, v), where M appears in δ and $v \in V(M)$ is miscolored, that is $\chi(v)(= i) \neq \alpha_M(v)$. This element further splits as

$$b^{(i)} = \sum_{\substack{i' \neq i \\ j}} b^{(i)}_{i',j},$$

where $b^{(i)}_{i',j}$ consists of those (M, v) for which $\alpha_M(v) = i'$ and $\alpha_M(u) = j$ $(V(M) = \{u, v\})$. Going one step further,

$$b^{(i)}_{i',j} = b^{(i)'}_{i',j} + b^{(i)''}_{i',j},$$

where $b^{(i)'}_{i',j}$ consists of those (M, v) in which the second vertex u is also miscolored, that is $\chi(u) \neq j$. Note for the record that

$$\delta = \sum_i [\![\sum_{\substack{i' \neq i \\ j}} b^{(i)'}_{i'j} + 2b^{(i)''}_{i'j}]\!]_{(i)} \tag{14}$$

(the extra coefficient 2 balances off the coefficient $\frac{1}{2}$ that will appear in $\sum [\![b^{(i)''}_{i'j}]\!]_{(i)}$ for those models $M \in \delta$ in which only one vertex is miscolored).

The upper bound on R_b will be actually given in terms of the expression $\sum_i [\![\sum_{\substack{i' \neq i \\ j}} (b^{(i)'}_{i'j} + 2b^{(i)''}_{i'j}) P^{(i)}_{i'j}]\!]_{(i)}$ differing from the right-hand side of (14) only in the extra term $P^{(i)}_{i'j}$. For that we have to bound from below $p^M \left(\sum_i [\![\sum_{\substack{i' \neq i \\ j}} (b^{(i)'}_{i'j} + 2b^{(i)''}_{i'j}) P^{(i)}_{i'j}]\!]_{(i)} \right)$ (see [34, Definition 7] for p^M), where M is a model of size k appearing in R_b. This quantity, however, has a very clean combinatorial meaning. Namely, let $c(M)$ be the number of ordered pairs $\langle i', j \rangle$ such that $\chi(i') \notin \{i', \chi(j)\}$, where those pairs for which $\chi(j) = j$ are counted twice. Then we have

$$p^M \left(\sum_i [\![\sum_{\substack{i' \neq i \\ j}} (b^{(i)'}_{i'j} + 2b^{(i)''}_{i'j}) P^{(i)}_{i'j}]\!]_{(i)} \right) = \frac{c(M)}{k!}. \tag{15}$$

The reason is simply that any pair $\langle i', j \rangle$ as described above determines an embedding of either $b^{(i)'}_{i'j}$ or $b^{(i)''}_{i'j}$ into M, with an appropriate coefficient. But once it is determined, there is precisely one way of assigning the remaining $(k-2)$ vertices to terms in the product $P_{i'j}$ (which is simply (2) with $i := i'$ and the term corresponding to $\{i', j\}$ missing).

We claim that $c(M) \geq 2(k-2)$. Indeed, another way to interpret $c(M)$ is as twice the number of *unordered* pairs $\{i, j\}$ that are transversal $(\chi(i) \neq \chi(j))$ and in which *at least one* of the two vertices is miscolored. Now, if χ is a (non-identical) permutation then the transversality restriction becomes void. Picking arbitrarily any miscolored i and any $j \neq i$ will already give us $(k-1)$ pairs of the desired form. If, on the other hand, χ is not a permutation, let C be any non-trivial χ-colored class: $2 \leq |C| \leq k-1$ (the latter condition holds since $\chi \neq \text{const}$). At least $|C|-1$ vertices in this class are miscolored which gives us $\geq (|C|-1)(k-|C|) \geq k-2$ desired pairs.

Thus, by (15), (7) and (14) we conclude that

$$\begin{aligned} R_b &\leq \frac{k!}{2(k-2)} \sum_i [\![\sum_{\substack{i' \neq i \\ j}} (b_{i'j}^{(i)\prime} + 2b_{i'j}^{(i)\prime\prime}) P_{i'j}^{(i)}]\!]_{(i)} \leq \frac{k!}{2(k-2)} \left(\frac{z_k}{k-2} \right)^{k-2} \sum_i [\![\sum_{\substack{i' \neq i \\ j}} b_{i'j}^{(i)\prime} + 2b_{i'j}^{(i)\prime\prime}]\!]_{(i)} \\ &= \delta \frac{k!}{2(k-2)} \left(\frac{z_k}{k-2} \right)^{k-2}. \end{aligned} \tag{16}$$

Along with (8) and (13), this gives us

$$\phi(R) \leq \phi(R) \sum_i p_i^k + k! \prod_i p_i \left(1 - \frac{\delta}{1 - \sum_i p_i^2} \right) + \frac{\delta k!}{2(k-2)} \left(\frac{z_k}{k-2} \right)^{k-2}.$$

But since the case $\sum_i p_i^k = 1$ was already treated above, in order to finish the proof of (3), it remains to show that

$$a_k \left(1 - \sum_i p_i^k \right) \geq k! \prod_i p_i \left(1 - \frac{\delta}{1 - \sum_i p_i^2} \right) + \frac{\delta k!}{2(k-2)} \left(\frac{z_k}{k-2} \right)^{k-2}$$

or, cancelling the factorials,

$$\frac{1 - \sum_i p_i^k}{k^k - k} \geq \prod_i p_i \left(1 - \frac{\delta}{1 - \sum_i p_i^2} \right) + \frac{\delta}{2(k-2)} \left(\frac{z_k}{k-2} \right)^{k-2}.$$

We now make use of the fact that the right-hand side here is linear in δ, hence it suffices to check our inequality at the end-points of the interval $\delta \in \left[0, 1 - \sum_i p_i^2\right]$.

The left end $\delta = 0$ amounts to

$$(k^k - k) \prod_i p_i + \sum_i p_i^k \leq 1. \tag{17}$$

Let $S \stackrel{\text{def}}{=} [k]^k \setminus \{(i, \ldots, i) : i \in [k]\}$. Since $|S| = k^k - k$, the AMGM inequality implies that

$$\sum_{(i_1, \ldots, i_k) \in S} p_{i_1} \cdots p_{i_k} \geq (k^k - k) \left(\prod_{(i_1, \ldots, i_k) \in S} p_{i_1} \cdots p_{i_k} \right)^{\frac{1}{k^k - k}} = (k^k - k) \prod_i p_i.$$

Adding $\sum_i p_i^k$ to both sides gives us the desired inequality since

$$\sum_{(i_1, \ldots, i_k) \in S} p_{i_1} \cdots p_{i_k} + \sum_i p_i^k = \left(\sum_i p_i \right)^k = 1.$$

The right end $\delta = 1 - \sum_i p_i^2$ leads to

$$\frac{1 - \sum_i p_i^k}{1 - \sum_i p_i^2} \geq \frac{(k^k - k)}{2(k-2)} \left(\frac{z_k}{k-2} \right)^{k-2}.$$

For $k \geq 5$ we simply use the trivial bound $1 - \sum_i p_i^k \geq 1 - \sum_i p_i^2$ so that we have to prove $\frac{k^k - k}{2(k-2)} \left(\frac{z_k}{k-2}\right)^{k-2} \leq 1$. For $k = 5, 6$ we can use numerical bounds[1] $z_5, z_6 \leq 0.3$ and for $k \geq 7$ the trivial bound $z_k \leq 1/2$ suffices.

When $k = 4$, we have to do a bit of extra work. Assume w.l.o.g. that p_1 is the largest. Then $\sum_i p_i^4 \leq p_1^2 \cdot \sum_i p_i^2$ and

$$\frac{1 - \sum_i p_i^k}{1 - \sum_i p_i^2} \geq \frac{1 - p_1^2 \sum_i p_i^2}{1 - \sum_i p_i^2} \geq \frac{4}{3}(1 - p_1^2/4)$$

(the latter inequality follows from $\sum_i p_i^2 \geq 1/4$). Since $z_4 \leq 0.257$, all that remains to prove is $p_1 \leq 0.91$. For that we simply re-use our previous calculation showing that w.l.o.g. we can assume $p_1 < 1$. Indeed, (11) is still true for the flag[2] $(R^{1,2})^{(1)}$. That is, under the additional assumption that the distinguished vertex is in V_1 we have

$$(R^{1,2})^{(1)} \leq \frac{2}{9} z_4^3,$$

and for all other i we still have the same bound but with the trivial estimate $S_2 \leq 1$:

$$(R^{1,2})^{(1)} \leq \frac{2}{9} \ (2 \leq i \leq k).$$

Now the bound (12) reads as

$$\phi(R) \leq \frac{8}{9}(z_4^3 p_1 + 1 - p_1).$$

Since $z_4 \leq 0.257$, this is $\leq \frac{2}{21}$ $(= a_4)$ whenever $p_1 \geq 0.91$. Hence we can assume w.l.o.g. that $p_1 \leq 0.91$ and, as we already observed, this implies (2.1) for $k = 4$.

This completes the proof of Theorem 3.1. □

Corollary 3.2. $h_1^{(k)}(s) = g_k(s) + 1$ *for all* $k \geq 4$ *whenever* s *is a power of* k.

Proof. Let H be a model of T with $|V(H)| = s$. We have to prove that $\binom{n}{k} i(R; H) \leq g_k(s)$. The plan is clear (and well-known): turn H into an element of $\mathrm{Hom}^+(\mathcal{A}^0[T], \mathbb{R})$ by replacing every $v \in V(H)$ with the infinite recursive construction and then apply Theorem 3.1 to it. There are several ways to make this intuition rigorous: we might consider convergent sequences or simply come up with an explicit formula as in [19, Section 2.3]. Let us do it geometrically (cf. [35, Section 2]) as this is the most elegant one.

Consider the infinite lexicographic product $\Omega \stackrel{\text{def}}{=} H \times R^\infty$. Thus, the vertices are infinite sequences $\mathsf{x} = (x_0, x_1, \ldots, x_n, \ldots)$, where $x_0 \in V(H)$ and $x_i \in V(R)$ $(i \geq 1)$. The edge coloring and the orientation between $\mathsf{x} \neq \mathsf{y}$ are read from the first coordinate i in which $x_i \neq y_i$. Further, Ω is equipped with the measure that is the product of uniform measures on $V(H), V(R)$ and all this structure turns Ω into a T-on ([8, Definition 3.2]). Hence we also have ([8, Theorem 6.3]) the

[1] A simple Maple worksheet verifying this fact, as well as several other facts of similar nature below, can be found at `http://homepages.math.uic.edu/~mubayi/papers/ErdosHajnalmw.pdf` and `http://people.cs.uchicago.edu/~razborov/files/ErdosHajnal.mw`

[2] A slightly better bound will be obtained in Lemma 4.7; the improvement is achieved by selecting the *minimal* flag among $(R^{1,2})^{(1)}, \ldots, (R^{1,k})^{(1)}$ rather than arbitrary. But we need not be that precise here.

corresponding algebra homomorphism $\phi \in \mathrm{Hom}^+(\mathcal{A}^0[T], \mathbb{R})$; its values are computed as obvious integrals over Ω. In particular, $\phi(R)$ is given by the "expected" formula

$$\phi(R) = \frac{(s)_k}{s^k} i(R; H) + \frac{a_k}{s^{k-1}}.$$

Along with Theorem 3.1, this leads us, after a bit of manipulations, to the bound

$$\binom{s}{k} i(R; H) \leq \frac{s^k - s}{k^k - k}.$$

When s is a power of k, the right-hand side here is exactly $g_k(s)$ (by an obvious induction). $\square$

4 Proof of Theorem 2.1

Before commencing with the formal proof of Theorem 2.1 we state some facts about partitions. Recall that a partition $n_1 + \cdots + n_k = n$ is *equitable* if $|n_i - n_j| \leq 1$ for all $i \neq j$.

Definition 4.1. Let $p(0,0) = 1$ and for $q > t > 0$, let $p(q,t)$ be the maximum of $\prod_i q_i$ where $q_1 + \cdots + q_t = q$ is a partition of q and each $q_i < q$.

It is easy to see that this maximum is achieved only by an equitable partition.

The following Lemma was stated by Erdős and Hajnal [12]. Since we could not find a proof of this, we will give a proof in the Appendix.

Lemma 4.2. *If $n \geq k \geq 3$, then $g_k(n)$ is achieved by an equitable partition.*

An immediate consequence of Lemma 4.2 is that

$$g_k(n) = p(n,k) \qquad \text{for } n \leq k(k-1). \tag{18}$$

Indeed, for $n \leq k(k-1)$ the equitable partition for n has each part of size less than k.

The next simple lemma collects some useful facts about $p(n,k)$. Its easy proof is left to the reader.

Lemma 4.3. *Let $k \geq 1$.*

a) $p(n+1,k) - p(n,k) = p(n - \lfloor n/k \rfloor, k-1)$.

b) $p(n,k)$ *is* **strictly** *increasing whenever $n \geq k-1$.*

c) for $n \geq n' \geq 1$,

$$p(n+1,k) + p(n'-1,k) \geq p(n,k) + p(n',k).$$

Proof of Theorem 2.1. In our proof we try to keep the notation reasonably consistent with Section 3 although some differences are unavoidable.

Fix $k \geq 4$ and a k-vertex tournament R with vertex set $[k]$ and pair $\{i,j\}$ is colored by $\{i,j\}$ from $C = \binom{[k]}{2}$. Let H be an n-vertex tournament with edges colored from C and let $i(R;H)$ be the number of copies of R in H. We are to prove that $i(R;H) \leq g_k(n)$ and we will proceed by induction on n.

For a vertex x in $V(H)$ and $i \in [k]$, write $d_i(x)$ for the number of copies of R containing x where x plays the role of vertex i in R. More formally, $d_i(x)$ is the number of isomorphic embeddings $\phi : R \to H$ such that $\phi(i) = x$. Let $d(x) = \sum_i d_i(x)$ be the number of copies of R containing x. For $i \in [k]$, let $N_i(x)$ be the set of those $y \in V(H) \setminus \{x\}$ for which there is a copy of R in H containing both x and y in which x plays the role of vertex i in R. Due to uniqueness of the colors of R we have $N_j(x) \cap N_{j'}(x) = \emptyset$ for $j \neq j'$. Moreover, $N_i(x)$ also has a (unique) partition $\cup_{j \neq i} N_i^j(x)$ where $N_i^j(x)$ comprises those y such that x, y lie in a copy of R with x playing the role of i and y playing the role of j [3]. We have

$$d(x) = \sum_{i=1}^{k} d_i(x) \leq \sum_{i=1}^{k} \prod_{j \neq i} |N_i^j(x)| \leq \sum_{i=1}^{k} p(|N_i(x)|, k-1). \tag{19}$$

We now partition $V(H)$ into $V_1 \cup \cdots \cup V_k$, $n_i = |V_i|$, where

$$V_i = \{x \in V(H) : d_i(x) \geq d_j(x) \text{ for all } j \neq i\}^4$$

and subject to this property, minimize $\sum_{i,j} |n_i - n_j|$. Note that $n_i < n$ for all i, for if, say, $n_1 = n$, then $d_1(x) \geq d_2(x)$ for all x, and $\sum d_1(x) = i(R;H) = \sum d_2(x)$ implies that $d_1(x) = d_2(x)$ for all x so we could move a vertex to V_2, contradicting the choice of partition.

Claim 1. $i(R;H) \leq g_k(n)$ for all $n \leq k(k-1)$.

Proof. We proceed by induction on n; the case $n \leq k$ is trivial. Pick a vertex x in H and suppose that there are $i \neq j$ with both $d_i(x)$ and $d_j(x)$ positive. Let $N_i = N_i(x)$ and $m_i = |N_i|$. Then (19) gives $d(x) \leq \sum_i p(m_i, k-1)$. By Lemma 4.3c), this is maximized when there exist $1 \leq i < j \leq k$ with $m_i + m_j = n - 1$ and, since $d_i(x), d_i(x) > 0$, one of m_i, m_j is equal to $k-1$. Consequently,

$$d(x) \leq p(m_i, k-1) + p(m_j, k-1) \leq 1 + p(n-k, k-1).$$

Deleting x we have, by induction, at most $i(R; H-x) \leq g_k(n-1)$ copies of R and hence $i(R;H) \leq i(R; H-x) + d(x) \leq g_k(n-1) + 1 + p(n-k, k-1)$. We claim that

$$g_k(n-1) + 1 + p(n-k, k-1) \leq g_k(n) \tag{20}$$

for $n \leq k(k-1)$. Indeed, applying (18) and Lemma 4.3a), this is equivalent to

$$p(n-1-\lfloor (n-1)/k \rfloor, k-1) \geq 1 + p(n-k, k-1)$$

[3] Thus, in the language of Section 3, the flag S_i provides a simple upper bound on the density of $N_i(x)$ while $\alpha_{\{i,j\}}/\beta_{\{i,j\}}$ upper bound $N_i^j(x)$.

[4] The corresponding definition of V_i in Section 3 considers $|N_i(x)|$ instead of $d_i(x)$.

which in turn follows from $n-1-\lfloor (n-1)/k \rfloor \geq \max(k-1, n-k+1)$ by Lemma 4.3b).

We may now assume that for each vertex x there is a unique i for which $d_i(x) > 0$ (otherwise apply (20)) and this gives a natural k-partition of the vertex set. Moreover, we now also have $i(R; H) \leq p(n,k) = g_k(n)$ by (18) since $n \leq k(k-1)$. □

Claim 1 concludes the base case and we now proceed to the induction step where we may assume that $n > k(k-1)$. We may also assume that $d(x) \geq d_{min} = g_k(n) - g_k(n-1)$ for each vertex x as otherwise we may delete x and apply induction.

The next part of the argument (up to the inequality (28)) closely parallels the one given in Section 3 but we give it here anyway for the sake of completeness.

Partition the copies of R in H as $H_m \cup H_g \cup H_b$ where H_m comprises those copies that lie entirely inside some V_i, H_g comprises those copies that intersect every V_i whose edge coloring coincides with the natural one given by the vertex partition (meaning the map from R to H takes vertex i to a vertex in V_i), and H_b comprises all other copies of R (these include transversal copies but some vertex in any such copy will be in an inappropriate V_i). Let $h_m = |H_m|, h_g = |H_g|$ and $h_b = |H_b|$ so that

$$i(R; H) = h_m + h_g + h_b.$$

We will bound each of these three terms separately. First, note that since $n_i < n$, by induction

$$h_m \leq \sum_j i(R; H[V_j]) \leq \sum_j g_k(n_j). \tag{21}$$

Next we turn to h_g. Let Δ denote the number of k-sets that intersect each V_i but are not counted by h_g. So a k-set counted by Δ either does not form a copy of R, or forms a copy of R but its edge coloring does not coincide with the natural one given by the vertex partition $V_1 \cup \ldots \cup V_k$. Then

$$h_g = \prod_i n_i - \Delta \tag{22}$$

and we need to bound Δ from below. Note that the color or orientation of some pair in every member of Δ does not align with the implicit one given by our partition. With this in mind, let D_{ij} be the set of pairs of vertices $\{v, w\}$ where $v \in V_i, w \in V_j$ such that either the color or orientation of vw does not match that of ij in R. Let $\delta_{ij} = |D_{ij}|/\binom{n}{2}$, $D = \cup_{ij} D_{ij}$ and $\delta = |D|/\binom{n}{2}$. Let us lower bound Δ by counting the misaligned pairs from D and then choosing the remaining $k-2$ vertices, one from each of the remaining parts V_ℓ. This gives, for each $i < j$,

$$\Delta \geq |D_{ij}| \prod_{\ell \neq i,j} n_\ell = \delta_{ij} \binom{n}{2} \prod_{\ell \neq ij} n_\ell = \delta_{ij} \binom{n}{2} \frac{\prod_\ell n_\ell}{n_i n_j}.$$

Since $\sum_{ij} \delta_{ij} \binom{n}{2} = \sum_{ij} |D_{ij}| = |D| = \delta \binom{n}{2}$, we obtain by summing over i, j,

$$\Delta \left(\sum_{1 \leq i < j \leq k} n_i n_j \right) \geq \delta \binom{n}{2} \prod_\ell n_\ell$$

which gives

$$h_g \leq \prod_\ell n_\ell \left(1 - \frac{\delta \binom{n}{2}}{\sum_{1 \leq i < j \leq k} n_i n_j} \right) = \prod_\ell n_\ell \left(1 - \frac{\delta \binom{n}{2}}{\binom{n}{2} - \sum_i \binom{n_i}{2}} \right). \tag{23}$$

Our next task is to upper bound h_b. For a vertex x and $j \in [k]$, recall that $N_j(x) \subset V(H)$ is the set of y such that x, y lie in a copy of R with x playing the role of vertex j in R. For $x \in V_i$, let

$$Z(x) \stackrel{\text{def}}{=} \max_{j \neq i} |N_j(x)| \qquad \text{and} \qquad z_{k,n} \stackrel{\text{def}}{=} \max_{x \in V(H)} \frac{Z(x)}{(n-1)}.$$

Later we will give upper bounds for $z_{k,n}$. For now, let us enumerate the set J of tuples (v, w, f) where $e = \{v, w\} \in D, f \in H_b$, $e \subset f$, and $v \in V_i$, but v plays the role of vertex $i' \neq i$ in the copy f of R. In particular, all $k-1$ pairs (v, x) with $x \in f$ contain color i'. For $m = (v, w, f) \in J$, say that m is 2-sided if $(w, v, f) \in J$ as well; otherwise say that m is 1-sided. Let J_i be the set of i-sided tuples $(i = 1, 2)$. We consider the weighted sum

$$S = 2|J_1| + |J_2|.$$

Observe that each $f \in H_b$ contains at least $k-2$ pairs from D. Indeed, if f is transversal, then it must contain a miscolored vertex which yields at least $k-1$ pairs from D in f. If f is not transversal, then take a largest color class C of f, observe that at least $|C|-1$ of the vertices in C are miscolored, and this yields at least $(|C|-1)(k-|C|) \geq k-2$ pairs from D in f. We conclude that each $f \in H_b$ contributes at least $2(k-2)$ to S since f contains at least $k-2$ pairs $e = \{u, v\} \in D$ and if (v, w, f) is 1-sided it contributes 2 to S while if it is 2-sided then it contributes 2 again since both (v, w, f) and (w, v, f) are counted with coefficient 1. This yields

$$S \geq 2(k-2)h_b. \tag{24}$$

On the other hand, we can bound S from above by first choosing $e \in D$ and then $f \in H_b$ as follows. Call $v \in e = \{v, w\} \in D$ *correct in* e if $v \in V_i$, vw has color $\{i, j\}$ for some j and $v \to w$ in H iff $i \to j$ in R; if v is not correct in e then say that v is *wrong in* e. The definition of D implies that every $e \in D$ has at least one wrong vertex in e (and possibly two wrong vertices). Let

$$D_i = \{\{v, w\} \in D : \{v, w\} \text{ contains exactly } i \text{ wrong vertices}\} \qquad (i = 1, 2).$$

The crucial observation is that

$$(v, w, f) \in J_i \qquad \Longrightarrow \qquad \{v, w\} \in D_i \qquad\qquad (i = 1, 2). \tag{25}$$

The reason this holds is that the color and orientation of $e \in D$ completely determine the role that its endpoints play in every copy of R containing e.

Now, to bound S from above, we use (25) and start by choosing $\{v, w\} \in D_i$ with wrong vertex v and then the remaining $k-2$ vertices of $f \setminus e$. If $v \in V_i$, then, since v is wrong in e and $e \subset f \in H_b$, the remaining $k-2$ vertices of $f \setminus e$ must all lie in $N_j(v) \setminus \{w\}$ for some $j \neq i$. So the number of choices for $f \setminus e$ is at most

$$p(|N_j(v)| - 1, k-2) \leq p(Z(v), k-2) \leq p((n-1)z_{k,n}, k-2)$$

and for each choice of $f \setminus e$, we obtain $m = (v, w, f) \in J_i$. This gives

$$S \leq 2 \sum_{\{v,w\} \in D_1} p((n-1)z_{k,n}, k-2) + 2 \sum_{\{v,w\} \in D_2} p((n-1)z_{k,n}, k-2) = 2\,|D|\, p((n-1)z_{k,n}, k-2).$$

Continuing, we obtain

$$S \le 2\,|D|\,p((n-1)z_{k,n}, k-2) \le 2\,\delta \binom{n}{2} \left(\frac{z_{k,n}}{k-2}\right)^{k-2} (n-1)^{k-2}. \tag{26}$$

Finally, (24) and (26) give

$$h_b \le \frac{S}{2(k-2)} < \frac{\delta\binom{n}{2}}{k-2} \left(\frac{z_{k,n}}{k-2}\right)^{k-2} (n-1)^{k-2}, \tag{27}$$

which is a refined version of (16). Using (21), (23) and (27) we now have

$$i(R;H) < \sum_i g_k(n_i) + \prod_\ell n_\ell \left(1 - \frac{\delta\binom{n}{2}}{\binom{n}{2} - \sum_i \binom{n_i}{2}}\right) + \frac{\delta\binom{n}{2}}{k-2} \left(\frac{z_{k,n}}{k-2}\right)^{k-2} (n-1)^{k-2}.$$

Our final task now is to upper bound the RHS by $g_k(n)$.

Since $\delta\binom{n}{2} \le \sum_{ij} n_i n_j = \binom{n}{2} - \sum_i \binom{n_i}{2}$, we have $\delta \in I \stackrel{\text{def}}{=} (0, 1 - \sum_i \binom{n_i}{2}/\binom{n}{2})$. Viewing the expression above as a linear function of δ, it suffices to check the endpoints of I.

If we let $\delta = 0$, then

$$i(R;H) \le \sum_i g_k(n_i) + \prod_i n_i \le g_k(n)$$

and we are done. The last inequality holds since $g_k(n)$ is the maximum over all partitions of n, possibly with empty parts (as long as no part has size n), and $n_1 + \cdots + n_k = n$ is one such partition.

If $\delta = 1 - \sum_i \binom{n_i}{2}/\binom{n}{2}$ then we get

$$i(R;H) \le \sum_i g_k(n_i) + \frac{\sum n_i n_j}{k-2} \left(\frac{z_{k,n}}{k-2}\right)^{k-2} (n-1)^{k-2}. \tag{28}$$

In order to show that the RHS in (28) is at most $g_k(n)$, we will use explicit upper bounds on $g_k(n_i)$ and $z_{k,n}$ and lower bounds on $g_k(n)$. Our first step is to state the following nontrivial lower bounds for $p(n,k)$. A proof is presented in the Appendix.

Lemma 4.4. *For integers $k \ge 3$ and $n > k(k-1)$,*

$$\left(\frac{n}{k}\right)^k (1 - e_k(n)) \le p(n,k) \le \left(\frac{n}{k}\right)^k,$$

where $e_k(n) = (4/27)(k^3/n^2)$.

We now give a bound on $z_{k,n}$.

Lemma 4.5. *For $k \ge 4$, $n > k(k-1)$, and $m = n - \lceil n/k \rceil$, let $z'_{k,n}$ be the largest real number $z \in (0, 1/2)$ that satisfies*

$$z^{k-1} + (1-z)^{k-1} \ge \frac{(k-1)^{k-1}}{k^{k-1}} \left(1 - \frac{(4/27)(k-1)^3}{m^2}\right). \tag{29}$$

Then $z_{k,n} \le z'_{k,n}$. Furthermore, $z_{k,n} < 0.2611$ if either $k = 4$, $n \ge 100$ or $k \ge 5$.

Proof. We begin by recalling that $d_{min} > g_k(n)-g_k(n-1) \geq p(m,k-1)$ where the second inequality holds by Lemma 4.2. Recall that $z_{k,n} = \max_y Z(y)/(n-1)$ and let $i \in [k]$ so that $x \in V_i$ achieves this maximum. Then by (19) we have

$$p(m,k-1) < d_{min} \leq d(x) = \sum_{\ell=1}^{k} d_\ell(x) \leq \sum_{\ell=1}^{k} p(|N_\ell(x)|, k-1). \tag{30}$$

Let $j \neq i$ be such that $z_{k,n} = Z(x)/(n-1) = |N_j(x)|/(n-1)$. Then, writing $z = z_{k,n}$, (30) continues as

$$\begin{aligned} p(m,k-1) &< \sum_{\ell=1}^{k} d_\ell(x) \leq \sum_{\ell=1}^{k} p(|N_\ell(x)|, k-1) \leq p(|N_j(x)|, k-1) + p(n-1-|N_j(x)|, k-1) \\ &= p(z(n-1), k-1) + p((1-z)(n-1), k-1). \end{aligned} \tag{31}$$

Using Lemma 4.4 this gives

$$\left(\frac{m}{k-1}\right)^{k-1}(1-e_{k-1}(m)) < \left(\frac{z(n-1)}{k-1}\right)^{k-1} + \left(\frac{(1-z)(n-1)}{k-1}\right)^{k-1}.$$

Since $\lceil n/k \rceil \leq (n+k-1)/k$, we have $m/(n-1) \geq (k-1)/k$ and the expression for $e_{k-1}(m)$ in Lemma 4.4 gives (29).

The RHS of (29) increases with n and it is easy to see that it is $> 2^{2-k}$ (the value of the LHS at $z = 1/2$) already when $n = k(k-1)+1$. Hence the corresponding equation has two roots $0 < z' < 1/2 < z'' < 1$ in the interval $(0,1)$ and $z_{k,n} \notin I_{k,n} = (z', z'')$ which is an interval symmetric around $1/2$. Since $I_{k,n+1} \supset I_{k,n}$, we conclude (for $k \geq 5$) that $z_{k,n} \notin I_{k,k(k-1)+1}$. Direct calculation shows that $I_{4,100} \supset (0.2611, 0.7389)$, $I_{5,21} \supset (0.2611, 0.7389)$, and it is an easy exercise to see that the intervals $I_{k,k(k-1)+1}$ only grow with k. To complete the proof, we only need to show that $z_{k,n} \leq 1/2$ for $k \geq 4$ and $n > k(k-1)$.

Suppose for the sake of contradiction that $Z(x) = |N_j(x)| \geq (n-1)/2$ (recall that $x \in V_i$). Then, since $x \in V_i$, we have

$$d_j(x) \leq d_i(x) \leq p((1-z)(n-1), k-1)$$

and hence instead of the bound $d_j(x) \leq p(z(n-1), k-1)$ we could use in (31) this better bound. That would give us

$$p(m,k-1) < 2p((1-z)(n-1), k-1) \leq 2p(\lfloor (n-1)/2 \rfloor, k-1).$$

This, however is false e.g. since, as we argued above, $1/2 \in I_{k,n}$. This contradiction shows that in fact $z_{k,n} \leq 1/2$ and completes the proof of Lemma 4.5 □

Our next lemma provides bounds for $g_k(n)$. We will give a proof in the Appendix.

Lemma 4.6. *For $k \geq 4$ and $n > k(k-1)$*

$$\frac{n^k - k^3 n^{k-2}}{k^k - k} \leq g_k(n) \leq \frac{n^k - n}{k^k - k}.$$

Our final task is to provide a nontrivial upper bound on each n_i. Write $p_i = n_i/n$ and $e'_k(n) = k^3/n^2$. Assume w.l.o.g. that $p_1 = \max p_i$.

Lemma 4.7. *Let $k \geq 4$ and $n > \begin{cases} k(k-1) & \text{if } k \geq 5 \\ 100 & \text{if } k = 4 \end{cases}$. Then we can assume w.l.o.g. that $p_1 < 0.86$.*

Proof. Let $p_1 = 1 - c_k$ and assume w.l.o.g. that $p_k = \min_{i>1} p_i$. Then our assumption implies that $p_k \leq c_k/(k-1)$. We consider

$$i(R;H) = \sum_{v\in V} d_k(v) = \sum_{v\notin V_k} d_k(v) + \sum_{v\in V_k} d_k(v). \tag{32}$$

Note that for $v \notin V_k$, $|N_k(v)| \leq z_{k,n}(n-1)$ and so $d_k(v) \leq p(z_{k,n}(n-1), k-1)$. For $v \in V_k$, we will use the weaker bound $d_k(v) \leq p(n-1, k-1)$. As we may assume that $i(R;H) \geq g_k(n)$ (otherwise we are done by induction), Lemma 4.6 and (32) give

$$\frac{n^k(1-e'_k(n))}{k^k - k} \leq g_k(n) \leq i(R;H) \leq n\, p(z_{k,n}(n-1), k-1) + \frac{c_k n}{k-1} p(n-1, k-1). \tag{33}$$

Dividing by n^k and using $p(n-1,k-1) < p(n,k-1) \leq (n/(k-1))^{k-1}$ we obtain

$$c_k \geq \frac{(k-1)^k}{k^k-k}(1-e'_k(n)) - (k-1)z_{k,n}^{k-1} \geq \frac{(k-1)^k}{k^k-k}\left(1 - \frac{k^3}{(k(k-1)+1)^2}\right) - (k-1)z_{k,n}^{k-1}.$$

By the last part of Lemma 4.5, $z_{k,n} \leq 0.27$. This shows that $c_4 > 0.14$ or $p_1 = 1 - c_4 < 0.86$, and it is an easy matter to see that the bound only improves as k increases. □

We are now ready to complete the proof. Recall that our main equation is

$$i(R;H) \leq \sum_i g_k(n_i) + \frac{\sum n_i n_j}{k-2}\left(\frac{z_{k,n}}{k-2}\right)^{k-2}(n-1)^{k-2}, \tag{34}$$

and we are to show that the RHS is at most $g_k(n)$.

We treat the case $k = 4,\ n \leq 100$ by exhaustive search through all partitions. A simple Maple worksheet verifying this fact (as well as a few other numerical facts that we state in our proof) can be found at the web pages `http://homepages.math.uic.edu/~mubayi/papers/ErdosHajnalmw.pdf` and `http://people.cs.uchicago.edu/~razborov/files/ErdosHajnal.mw`. Thus, in what follows we always assume that $k = 4$ entails $n \geq 100$. In particular, we can utilize the conclusions of Lemmas 4.5 and 4.7:

$$z_{k,n} < 0.2611,\ \ p_1 < 0.86.$$

Dividing by $n^k/(k^k - k)$ and using Lemma 4.6, we see that it suffices to prove

$$L \stackrel{\text{def}}{=} \sum_i p_i^k + A\sum_{ij} p_i p_j \leq 1 - e'_k(n)$$

where as before $e'_k(n) = k^3/n^2$ and

$$A = A(k) \stackrel{\text{def}}{=} \frac{(k^k - k) \cdot 0.2611^{k-2}}{(k-2)^{k-1}}.$$

Since $p_1 \geq p_i$ for all i and $\sum_i p_i = 1$,

$$L \leq p_1^{k-2} \left(\sum_i p_i^2 \right) + A \sum_{ij} p_i p_j = p_1^{k-2} \left(1 - 2 \sum_{ij} p_i p_j \right) + A \sum_{ij} p_i p_j = p_1^{k-2} + (A - 2p_1^{k-2}) \sum_{ij} p_i p_j.$$

If $A < 2p_1^{k-2}$, then $k > 4$ since $A(4) > 2.14$ and the coefficient of $\sum_{ij} p_i p_j$ is negative. Lemma 4.7 then gives

$$L < p_1^{k-2} \leq p_1^3 < (0.86)^3 < 0.7 < 1 - e'_4(21) \leq 1 - e'_k(n). \tag{35}$$

Now assume that $A \geq 2p_1^{k-2}$. Then, using $\sum_{ij} p_i p_j \leq (k-1)/2k$ it is enough to show

$$p_1^{k-2} + \frac{k-1}{2k}(A - 2p_1^{k-2}) < 1 - e'_k(n). \tag{36}$$

For the same reasons as at the end of the proof of Theorem 3.1, we must split further analysis into two cases.

If $k \geq 5$, we apply the trivial bound $\frac{k-1}{2k} < \frac{1}{2}$ that reduces (36) to merely

$$A < 2(1 - e'_k(n)).$$

This holds since

$$A(k) \leq A(5) < 1 < 2(1 - e'_5(21)) < 2(1 - e'_k(n)).$$

For $k = 4$, (36) becomes

$$\frac{p_1^2}{4} + \frac{3}{8}A(4) < 1 - e'_4(n).$$

In this case $p_1 < 0.86$, $A(4) < 2.15$ and

$$\frac{p_1^2}{4} + \frac{3}{8}A(4) < 0.992 < 1 - e'_4(100).$$

The proof of Theorem 2.1 is complete. □

5 Appendix

Here we give the proofs of Lemmas 4.2, 4.4 and 4.6.

5.1 Proof of Lemma 4.2

We only consider partitions $n_1 + \cdots + n_k = n$ where $0 \le n_i < n$ for all i; the partition is equitable if $|n_i - n_j| \le 1$ for all $i \ne j$. For a vertex v in a hypergraph H, we write $d_H(v)$ for the degree of v in H. We will use the notation H for the edge set of H.

Definition 5.1. For $k \ge 3$ and $n \ge 0$, let $\mathcal{G}_k(n)$ be the family of n-vertex k-graphs defined inductively as follows: For $n < k$, $\mathcal{G}_k(n)$ comprises the single n-vertex k-graph with no edge. For $n \ge k$, the vertex set V of any $G \in \mathcal{G}_k(n)$ is partitioned as $V_1 \cup \ldots \cup V_k$ with $n_i := |V_i|$ and $0 \le n_i < n$ for all i. For the edge set, $G[V_i] \in \mathcal{G}_k(n_i)$ for each i, and in addition G contains all edges that have one point in each V_i. Call $V_1 \cup \ldots \cup V_k$ the *defining partition* for G, or simply, *the partition* for G.

Definition 5.2. Let $H_k(n) \in \mathcal{G}_k(n)$ be the following k-graph. For $n < k$, $H_k(n)$ is the unique member of $\mathcal{G}_k(n)$, and for $n \ge k$, the defining partition $V_1 \cup \ldots \cup V_k$ of $H_k(n)$ is an equitable partition ($||V_i| - |V_j|| \le 1$ for all $i \ne j$) and for each i, the subgraph induced by V_i is isomorphic to $H_k(|V_i|)$. We let $h_k(n) := |H_k(n)|$.

Our proof of Lemma 4.2 will use induction on n and so we need one more definition.

Definition 5.3. A vertex v of $G \in \mathcal{G}_k(n)$ is *G-good* if the following holds: for $n < k$ every vertex is G-good. For $n \ge k$, if $V_1 \cup \ldots \cup V_k$ is the partition for G, and $|V_i| \ge |V_j|$ for all j, then v is G-good if $v \in V_i$ and v is $G[V_i]$-good. In other words, a vertex is G-good if it lies in a largest part V_i in the partition for G and the same is true inductively within V_i.

Removing any vertex v from $G \in \mathcal{G}_k(n)$ results in a k-graph $G - v \in \mathcal{G}_k(n-1)$. Moreover, if v is $H_k(n)$-good, then

$$H_k(n) - v \cong H_k(n-1). \tag{37}$$

Indeed, if we remove v from $H_k(n)$, then the partition for $H_k(n)$, after removal of v, is still equitable and the same remains true of all inductively defined partitions. Now (37) shows that every two $H_k(n)$-good vertices have the same degree and hence we may define $\delta_k(n) = d_{H_k(n)}(v)$ where v is any $H_k(n)$-good vertex. Observe that

$$\delta_k(n) = d_{H_k(n)}(v) = d_{H_k(\lceil n/k \rceil)}(v) + p(n - \lceil n/k \rceil, k-1) = \delta_k(\lceil n/k \rceil) + p(n - \lceil n/k \rceil, k-1). \tag{38}$$

Finally, (37) gives

$$h_k(n-1) + \delta_k(n) = h_k(n). \tag{39}$$

Lemma 5.4. *Let $G \in \mathcal{G}_k(n)$ and v be G-good. Then $d_G(v) \le \delta_k(n)$.*

Proof. Proceed by induction on n. The cases $n < k$ are trivial since $d_G(v) = 0 = \delta_k(n)$. Let $V_1 \cup \ldots \cup V_k$ be the partition for G and $n_i := |V_i|$ with $n_1 \ge n_2 \ge \cdots \ge n_k$ and assume wlog that $v \in V_1$. Let $X_1 \cup \cdots \cup X_k$ be the partition for $G[V_1] \in \mathcal{G}_k(n_1)$, $x_i := |X_i|$ with $b := x_1 \ge \cdots \ge x_k$ and assume wlog that $v \in X_1$. Note that $b \ge \lceil n_1/k \rceil$. Let

$$a_1 := |V_2 \cup \cdots \cup V_k| = n - n_1, \qquad a_2 := |V_1| - |X_1| = n_1 - b.$$

Since v is $G[X_1]$-good, $G[X_1] \in \mathcal{G}_k(b)$, and $b < n$, induction implies $d_{G[X_1]}(v) \le \delta_k(b)$ and hence

$$d_G(v) = d_{G[V_1]}(v) + \prod_{j=2}^{k} n_j = d_{G[X_1]}(v) + \prod_{\ell=2}^{k} x_\ell + \prod_{j=2}^{k} n_j \le \delta_k(b) + p(a_1, k-1) + p(a_2, k-1). \quad (40)$$

Case 1. $b \le \lceil n/k \rceil$**.** Note that $a_i + b \ge \lceil n/k \rceil$ for $i = 1, 2$ since $a_1 + b = (n - n_1) + b \ge (n - n_1) + n_1/k \ge n/k$ and $a_2 + b = n_1 \ge n/k$. For fixed b and n, $a_1 + a_2 = n - b$ is also fixed, so by Lemma 4.3c), $p(a_1, k-1) + p(a_2, k-1)$ is uniquely maximized when a_1 or a_2 is as small possible, namely $\{a_1, a_2\} = \{\lceil n/k \rceil - b, n - \lceil n/k \rceil\}$ where we use the assumption $\lceil n/k \rceil - b \ge 0$. Consequently,

$$d_G(v) \le \delta_k(b) + p(\lceil n/k \rceil - b, k-1) + p(n - \lceil n/k \rceil, k-1). \quad (41)$$

If $b = \lceil n/k \rceil$, then (38) and (41) give $d_G(v) \le \delta_k(n)$ and we are done, so assume that $b < \lceil n/k \rceil$. Consider $K \in \mathcal{G}_k(\lceil n/k \rceil)$ whose defining partition has largest part B of size b and all other $k-1$ parts form an equitable partition of $\lceil n/k \rceil - b$. Since $b = x_1 \ge \lceil n_1/k \rceil \ge \lceil \lceil n/k \rceil / k \rceil$, B is indeed a largest part. Further, let $K[B] \cong H_k(b)$ (for all other parts C, choose $K[C]$ arbitrarily) and let w be a $K[B]$-good vertex. Then $d_{K[B]}(w) = \delta_k(b)$ and w is also K-good, so by induction,

$$\delta_k(b) + p(\lceil n/k \rceil - b, k-1) = d_{K[B]}(w) + p(\lceil n/k \rceil - b, k-1) = d_K(w) \le \delta_k(\lceil n/k \rceil).$$

Continuing (41) we obtain

$$d_G(v) \le \delta_k(\lceil n/k \rceil) + p(n - \lceil n/k \rceil, k-1) = \delta_k(n)$$

where we use (38) for the last equality.

Case 2. $b > \lceil n/k \rceil$**.** In this case (40) and Lemma 4.3 yield

$$d_G(v) \le \delta_k(b) + p(a_1, k-1) + p(a_2, k-1) \le \delta_k(b) + p(a_1 + a_2, k-1) = \delta_k(b) + p(n - b, k-1).$$

Consider $G' \in \mathcal{G}_k(n)$ whose defining partition has largest part B' of size b and all other $k-1$ parts form an equitable partition of $n - b$. Since $b \ge \lceil n/k \rceil$, B' is indeed the largest part. Further, let $G'[B'] \cong H_k(b)$ and let $X'_1 \cup \ldots \cup X'_k$ be the partition for $G'[B']$ with $|X'_1| = \lceil b/k \rceil$. Let $v' \in X'_1$ be a $G'[B']$-good vertex. Since B' is the largest part in the partition for G', v' is also G'-good. Since $|X'_1| = \lceil b/k \rceil \le \lceil n/k \rceil$, by the proof in Case 1 (with (G', X'_1, v') playing the role of (G, X_1, v)) we conclude that $d_{G'}(v') \le \delta_k(n)$. On the other hand, since v' is $G'[B']$-good and $G'[B'] \cong H_k(b)$,

$$d_{G'}(v') = d_{G'[B']}(v') + p(n - b, k-1) = d_{H_k(b)}(v') + p(n - b, k-1) = \delta_k(b) + p(n - b, k-1).$$

We therefore have

$$d_G(v) \le \delta_k(b) + p(n - b, k-1) = d_{G'}(v') \le \delta_k(n). \quad \square$$

Proof of Lemma 4.2. We are to show that $g_k(n) = h_k(n)$. In other words, for each $G \in \mathcal{G}_k(n)$ we must show that $|G| \le h_k(n)$. We proceed by induction on n. The cases $n \le k$ are trivial, so assume $n > k$. Pick a G-good vertex v. Then $G - v \in \mathcal{G}_k(n-1)$ and by induction, Lemma 5.4, and (39),

$$|G| = |G - v| + d_G(v) \le h_k(n-1) + \delta_k(n) = h_k(n). \quad \square$$

5.2 Proof of Lemma 4.4

We begin with the following inequality.

Lemma 5.5. *Let a, b, k, n be positive integers with $k = a + b \geq 2$, and $n \geq \max\{a, b\}$. Then*

$$\left(1 + \frac{a}{n}\right)^b \left(1 - \frac{b}{n}\right)^a \geq 1 - \frac{\max\{ab^2, ba^2\}}{n^2} \geq 1 - \frac{(4/27)k^3}{n^2}.$$

Proof. Suppose first that $a \leq b$. Consider a copies of the number $1 + b/n$ and $b - a$ copies of the number 1. The arithmetic mean of these numbers is $1 + a/n$ hence the AMGM inequality gives

$$\left(1 + \frac{a}{n}\right)^b \geq \left(1 + \frac{b}{n}\right)^a \cdot 1^{b-a}$$

and Bernoulli's estimate yields

$$\left(1 + \frac{a}{n}\right)^b \left(1 - \frac{b}{n}\right)^a \geq \left(1 - \frac{b^2}{n^2}\right)^a \geq 1 - \frac{ab^2}{n^2}.$$

If $a \geq b$, then a similar argument applies by taking b copies of $1 - a/n$ and $a - b$ copies of 1. □

Proof of Lemma 4.4. We are to show that for $k \geq 3$ and $n > k(k-1)$,

$$\left(\frac{n}{k}\right)^k (1 - e_k(n)) \leq p(n, k) \leq \left(\frac{n}{k}\right)^k,$$

where $e_k(n) = (4/27)(k^3/n^2)$. The upper bound for $p(n, k)$ is trivial so we only prove the lower bound. Let $n \equiv t \pmod{k}$ where $0 \leq t < k$. Then

$$p(n, k) = \left(\frac{n}{k} + \frac{k-t}{k}\right)^t \left(\frac{n}{k} - \frac{t}{k}\right)^{k-t} = \left(\frac{n}{k}\right)^k \left(1 + \frac{k-t}{n}\right)^t \left(1 - \frac{t}{n}\right)^{k-t}.$$

Now apply Lemma 5.5 with $a = k - t$ and $b = t$. □

5.3 Proof of Lemma 4.6

We are to show that for $k \geq 4$ and $n > k(k-1)$

$$\frac{n^k - k^3 n^{k-2}}{k^k - k} \leq g_k(n) \leq \frac{n^k - n}{k^k - k}.$$

The upper bound is by induction on n and we prove it for all $n \geq 1$. The base cases $n \leq k$ are obvious so let $n > k$. For the induction step, apply Lemma 4.2 and take the equitable partition $n = \sum_i n_i$ that achieves the definition of $g_k(n)$. Then each $n_i < n$ and by induction,

$$g_k(n) = \sum_i g_k(n_i) + \prod_i n_i \leq \sum_i \frac{n_i^k - n_i}{k^k - k} + \prod_i n_i = \sum_i \frac{n_i^k}{k^k - k} - \frac{n}{k^k - k} + \prod_i n_i \leq \frac{n^k - n}{k^k - k}$$

where the last inequality (after dividing by $n^k/(k^k-k)$) is (17).

For the lower bound, we take an equitable partition $\sum n_i = n$ and proceed by induction on n. Let us first assume that $k \geq 5$. We will actually prove that

$$(k^k - k)\, g_k(n) \geq \begin{cases} n^k - k^4 n^{k-2} & \text{if } n \leq k(k-1) \\ n^k - k^3 n^{k-2} & \text{if } n > k(k-1). \end{cases}$$

Note that the first bound is trivial for $n \leq k(k-1)$ since it is negative. For $n > k(k-1)$, Lemma 4.4 and induction imply that $g_k(n)$ is at least

$$\sum_i g_k(n_i) + p(n,k) \geq \sum_i \frac{n_i^k - k^4 n_i^{k-2}}{k^k - k} + \left(\frac{n}{k}\right)^k (1 - e_k(n)) \geq \frac{n^k}{k^k - k} - \left(\frac{\sum k^4 n_i^{k-2}}{k^k - k} + \frac{n^k e_k(n)}{k^k}\right)$$

where we used Jensen's inequality to obtain

$$\sum \frac{n_i^k}{k^k - k} + \left(\frac{n}{k}\right)^k \geq \left(\frac{n}{k}\right)^k \left(\frac{k}{k^k - k} + 1\right) = \frac{n^k}{k^k - k}.$$

Since $n > k(k-1)$ and the n_i's are an equitable partition,

$$\sum_{i=1}^{k} n_i^{k-2} < \sum_{i=1}^{k} (n/k+1)^{k-2} = k(n/k+1)^{k-2} = \frac{n^{k-2}}{k^{k-3}}(1+k/n)^{k-2} < \frac{n^{k-2}}{k^{k-3}} \cdot e. \tag{42}$$

Using $k \geq 5$, we now have

$$\frac{\sum k^4 n_i^{k-2}}{k^k - k} \leq e \cdot \frac{k^4}{k^k - k} \cdot \frac{n^{k-2}}{k^{k-3}} < \frac{18}{27} \frac{k^3 n^{k-2}}{k^k - k}. \tag{43}$$

Hence

$$\frac{\sum k^4 n_i^{k-2}}{k^k - k} + \frac{n^k e_k(n)}{k^k} < \frac{(18/27) k^3 n^{k-2}}{k^k - k} + \frac{n^k (4/27)(k^3/n^2)}{k^k - k} < \frac{k^3 n^{k-2}}{k^k - k} \tag{44}$$

and the proof is complete.

Now we assume that $k = 4$. In this case we show by induction on n that for all $n > 0$,

$$g_4(n) \geq \frac{n^k - k^3 n^{k-2}}{k^k - k} = \frac{n^4 - 64n^2}{252}.$$

The cases $n \leq 8$ are trivial since the RHS is nonpositive so assume that $n \geq 9$. The cases $9 \leq n \leq 12$ can be checked by direct computation, so assume that $n > 12$ and we can apply the bounds in Lemma 4.4. We proceed as in the proof for $k \geq 5$ except that all occurrences of the k^4 term are replaced by $k^3 = 4^3$. Now (42) becomes

$$\sum_{i=1}^{4} n_i^2 < \frac{n^2}{4}(1+4/n)^2 \leq \frac{n^2}{4}(1+4/9)^2 < (2.1)\frac{n^2}{4}.$$

and, since $2.1 < 72/27$, (43) becomes

$$\frac{\sum 4^3 n_i^2}{252} \leq \frac{4^3}{252} \cdot (2.1) \cdot \frac{n^2}{4} < \frac{18}{27}\frac{4^3 n^2}{252}.$$

Finally, (44) becomes

$$\frac{\sum 4^3 n_i^2}{252} + \frac{n^4 e_4(n)}{4^4} < \frac{(18/27)4^3 n^2}{252} + \frac{n^4(8/27)(4^3/n^2)}{252} < \frac{64n^2}{252}$$

and the proof is complete. □

Acknowledgment

We are grateful to David Conlon and Sergey Tarasov for several useful remarks and to a referee for extremely careful and expert reading of the paper and providing a shorter proof of Lemma 5.5.

References

[1] J. Balogh, P. Hu, B. Lidický, F. Pfender: Maximum density of induced 5-cycle is achieved by an iterated blow-up of 5-cycle, *European J. Combin.* 52 (2016), 47-58

[2] J. Balogh, P. Hu, B. Lidický, F Pfender, J. Volec, M. Young, Michael, Rainbow triangles in three-colored graphs. *J. Combin. Theory Ser. B* 126 (2017), 83–113.

[3] V. Bhat, V. Rödl, Note on upper density of quasi-random hypergraphs *Electron. J. Combin.* 20 (2013), no. 2, Paper 59, 8 pp.

[4] J. Brown, A. Sidorenko, The inducibility of complete bipartite graphs. *J. Graph Theory* 18 (1994), no. 6, 629–645.

[5] F. R. K. Chung, Open problems of Paul Erdős in graph theory *J. Graph Theory* **25** (1997), 3–36.

[6] D. Conlon, J. Fox, and B. Sudakov, Hypergraph Ramsey numbers, *J. Amer. Math. Soc.* **23** (2010), 247–266.

[7] D. Conlon, H. Hán, Y. Person, M. Schacht, Weak quasi-randomness for uniform hypergraphs, *Random Structures Algorithms* 40 (2012), no. 1, 1–38.

[8] L. Coregliano, A. Razborov, Semantic Limits of Dense Combinatorial Objects, arXiv:1910.08797 [math.CO], to appear in *Russian Mathematical Surveys.*

[9] D. Duffus, H. Lefmann, and V. Rödl, Shift graphs and lower bounds on Ramsey numbers $r_k(l;r)$, *Discrete Math.* **137** (1995), 177–187.

[10] M. Eliáš and J. Matoušek, Higher-order Erdős-Szekeres theorems, *Adv. Math.* **244** (2013), 1–15.

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

[12] P. Erdős, A. Hajnal, On Ramsey like theorems, problems and results, in Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst., Oxford, 1972), pp. 123–140, Inst. Math. Appl., Southhend-on-Sea, 1972.

[13] P. Erdős, R. Rado, Combinatorial theorems on classifications of subsets of a given set, *Proc. Lond. Math. Soc.* **3** (1952), 417–439.

[14] C. Even-Zohar, N. Linial, A note on the inducibility of 4-vertex graphs, *Graphs Combin.* 31 (2015), no. 5, 1367–1380.

[15] J Fox, H. Huang, C. Lee, private communication.

[16] J. Fox, J. Pach, B. Sudakov, and A. Suk, Erdős-Szekeres-type theorems for monotone paths and convex bodies, *Proc. Lond. Math. Soc.* **105** (2012), 953–982.

[17] P. Frankl, V. Rödl, Hypergraphs do not jump, *Combinatorica* 4(2) (1984), 149–159.

[18] H. Hatami, J. Hirst, S. Norine, The inducibility of blow-up graphs. *J. Combin. Theory Ser. B* 109 (2014), 19–212.

[19] H. Hatami, J. Hladky, D. Kral, S. Norin, A. Razborov, On the Number of Pentagons in Triangle-Free Graphs, *J. Combin. Theory Ser. B* **120** (2013), 722–732.

[20] J. Hirst, The inducibility of graphs on four vertices. *J. Graph Theory* 75 (2014), no. 3, 231–243.

[21] H. Huang, On the maximum induced density of directed stars and related problems, *SIAM J. Discrete Math.*, 28-1 (2014), 92-98.

[22] Y. Kohayakawa, B. Nagle, V. Rödl, and M. Schacht, Weak hypergraph regularity and linear hypergraphs, *J. Combin. Theory Ser. B* **100** (2010), 151–160.

[23] T . Kövári, V. T. Sós, P. Turán, On a problem of K. Zarankiewicz, *Colloquium Math* 3 (1954), 50-57.

[24] D. Kral, S. Norine. J. Volec, A bound on the inducibility of cycles, *J. Combin. Theory, Ser. A* **161** (2019), 359-363

[25] J. Lenz and D. Mubayi, The poset of hypergraph quasirandomness, *Random Structures Algorithms* **46** (2015), 762–800.

[26] J. Lenz and D. Mubayi, Eigenvalues and linear quasirandom hypergraphs, *Forum Math. Sigma* **3** (2015), e2, 26 pp.

[27] N. Linial and A. Morgenstern, On high-dimensional acyclic tournaments, *Discrete Comput. Geom.* **50** (2013), 1085—1100.

[28] Mathematics of Ramsey theory, Edited by Jaroslav Nešetřil and Vojtěch Rödl. Algorithms and Combinatorics, 5. Springer-Verlag, Berlin, 1990. xiv+269 pp. ISBN: 3-540-18191-1

[29] K.G. Milans, D. Stolee, and D. West, Ordered Ramsey theory and track representations of graphs, *J. Comb.* 6 (2015), no. 4, 445–456.

[30] G. Moshkovitz and A. Shapira, Ramsey-theory, integer partitions and a new proof of the Erdős-Szekeres theorem, *Adv. Math.* **262** (2014), 1107–1129.

[31] D. Mubayi, A. Suk, The Erdos-Hajnal hypergraph Ramsey problem, to appear in *J. Eur. Math. Soc. (JEMS).*

[32] N. Pippenger, M. Golumbic, The inducibility of graphs. *J. Combin. Theory Ser. B* 19 (1975), no. 3, 189–203.

[33] C. Reiher, V. Rödl, M. Schacht, Hypergraphs with vanishing Turán density in uniformly dense hypergraphs *J. Lond. Math. Soc.* (2) 97 (2018), no. 1, 77–97.

[34] A. Razborov, Flag algebras, *J. Symbolic Logic* **72** (2017), 1239-1282.

[35] A. Razborov, On the Caccetta-Haggkvist conjecture with forbidden subgraphs, *J. Graph Theory* **74** (2013), 236–248.

[36] K. Sperfeld. The inducibility of small oriented graphs. Technical Report 1111.4813 [math.CO], arXiv, 2011.

[37] R. Yuster. On the exact maximum induced density of almost all graphs and their inducibility. Technical Report 1801.01047 [math.CO], arXiv, 2018.