Math. Proc. Camb. Phil. Soc. (2026), **180**, 191–218
doi:10.1017/S0305004125101692
First published online 2 October 2025

Inducibility of rainbow graphs

BY EMILY CAIRNCROSS, CLAYTON MIZGERD AND DHRUV MUBAYI
Department of Mathematics, Statistics and Computer Science, University of Illinois Chicago, IL 60607, *U.S.A.*
e-mails: emilyc10@uic.edu, cmizge2@uic.edu, mubayi@uic.edu

(*Received* 24 *July* 2024; *revised* 31 *May* 2025; *accepted* 05 *June* 2025)

Abstract

We prove that there is an absolute constant $C > 0$ such that every k-vertex connected rainbow graph R with minimum degree at least $C \log k$ has inducibility $k!/(k^k - k)$. The same result holds if $k \geq 11$, and R is a clique. This answers a question posed by Huang, that is a generalisation of an old problem of Erdös and Sós. It remains open to determine the minimum k for which this is true.

2020 Mathematics Subject Classification: 05C35

1. *Introduction*

Fix a graph F on k vertices and another graph G on $n > k$ vertices. Write $I(F, G)$ for the number of k-subsets $S \subset V(G)$ such that $G[S] \cong F$ and let

$$\varrho(F, G) := \frac{I(F, G)}{\binom{n}{k}}.$$

Many foundational questions in extremal graph theory deal with estimating $\varrho(F, G)$ for various choices of F and G. One central question is to determine the minimum value when F is a clique and G has a specified edge density [**13**, **15**, **16**], but there are also many fundamental questions about the maximum value regardless of edge density. This is the direction we take here.

Let $I(F, n)$ be the maximum of $I(F, G)$ over all n vertex graphs G. A standard averaging argument implies that

$$\operatorname{ind}(F, n) := \frac{I(F, n)}{\binom{n}{k}} \leq \frac{I(F, n-1)}{\binom{n-1}{k}} = \operatorname{ind}(F, n-1).$$

Thus, $\operatorname{ind}(F, n)$ is a decreasing sequence bounded below by zero, so it has a limit. Define the *inducibility* of F to be

$$\operatorname{ind}(F) := \lim_{n \to \infty} \operatorname{ind}(F, n).$$

The *iterated balanced blow-up of a* graph F is a family $\mathcal{G}_F(n)$ of graphs on n vertices defined inductively as follows. Label $V(F)$ with $[k] := \{1, \ldots, k\}$. For $n < k$, the family

© The Author(s), 2025. Published by Cambridge University Press on behalf of Cambridge Philosophical Society. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

$\mathcal{G}_F(n)$ contains only the empty graph on n vertices. For $n \geq k$, for any $G \in \mathcal{G}_F(n)$, we have a partition $V(G) = V_1 \cup \cdots \cup V_k$ with the following properties:

(1) for all $i, j \in [k]$, $\big||V_i| - |V_j|\big| \leq 1$;
(2) for all $i \in [k]$, the induced subgraph $G[V_i] \in \mathcal{G}_F(|V_i|)$;
(3) for all $v \in V_i, w \in V_j$ with $i \neq j$, we have $vw \in E(G)$ if and only if $ij \in E(F)$.

In many interesting cases, the construction above achieves the inducibility of F and we now define this formally (our definition is slightly different than that in [11]).

Definition 1·1. A graph F is a *fractaliser* if

$$\operatorname{ind}(F) = \lim_{n\to\infty} \max_{G\in\mathcal{G}_F(n)} \varrho(F, G).$$

In other words, the iterated balanced blow-up of F achieves the inducibility.

The subgraph induced by every k-set comprising exactly one vertex in each V_i is isomorphic to F. Consequently, for every $G \in \mathcal{G}_F(n)$,

$$I(F, G) \geq \sum_{i=1}^{k} I(F, G[V_i]) + \prod_{i=1}^{k} |V_i|.$$

Together with a standard computation (see, e.g. [12]), this yields

$$\operatorname{ind}(F) \geq \lim_{n\to\infty} \max_{G\in\mathcal{G}_F(n)} \varrho(F, G) \geq \frac{k!}{k^k - k}. \tag{1·1}$$

In most cases we consider, the fact that F is a fractaliser will imply further that $\operatorname{ind}(F) = k!/(k^k - k)$. Note that this is not always the case, for example $F = K_k$ is a fractaliser with $\operatorname{ind}(K_k) = 1$.

The fundamental conjecture in this area, due to Pippenger and Golumbic [14], states that for $k \geq 5$, the cycle C_k is a fractaliser and satisfies $\operatorname{ind}(C_k) = k!/(k^k - k)$. This conjecture has been resolved for $k = 5$ by Balogh, Hu, Lidický, and Pfender [1] (see also [11]), but remains open for all $k \geq 6$. Král, Norin and Volec [10] showed that $I(C_k, n) \leq 2n^k/k^k$. More generally, Fox, Huang and Lee [6] and Yuster [17] independently proved that random graphs are fractalisers asymptotically almost surely. Fox, Sauermann and Wei [7] further proved that random Cayley graphs of abelian groups with small number of vertices removed are almost surely fractalisers.

We now consider these notions on coloured and directed structures. A *tournament* is an orientation of a complete graph. An edge-colouring of a graph or tournament G is a function $\chi : E(G) \to T$ where T is a set of colours; we say that G is T-coloured. A coloured graph or tournament G is *rainbow* if χ is injective. Two coloured graphs (or tournaments) G and H are *isomorphic*, written $G \cong H$, if there exists a bijection $\varphi : V(G) \to V(H)$ such that the colours (and orientations) of all edges are preserved under φ. If F is a coloured tournament or coloured complete graph, then $\operatorname{ind}(F)$ is defined identically as in the graph case, but with these altered definitions of graph isomorphism; naturally, the underlying graph G should have the colours or orientations corresponding to F. If F is an arbitrary coloured graph, then we can colour all missing edges with a single new colour and view F as a coloured complete graph. Consequently, we can define *fractaliser* for all these structures.

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

There are very few results on the inducibility of coloured, oriented or directed structures. The first exact result which involved an iterated construction was due to Huang [8] who determined the inducibility of the directed star. Later, in order to solve an old conjecture of Erdös and Hajnal [5] in hypergraph Ramsey theory, the third author and Razborov [12] proved the following result for $k \geq 4$ (the case $k = 3$ was proven earlier by Conlon, Fox and Sudakov [3]).

THEOREM 1·2 ([12]). *All rainbow tournaments R on $k \geq 4$ vertices are fractalisers. In particular,* $\mathrm{ind}(R) = k!/(k^k - k)$.

In this paper, we consider the question addressed by Theorem 1·2 in the undirected setting. The first conjecture in this setting is due to Erdös and Sós from the 1970s (see [5, equation (20)]), and implies, in particular, that a rainbow triangle is not a fractaliser. Their conjecture was proved by Balogh et. al. [2], who showed that a blow-up of a properly 3-edge-coloured K_4 (instead of a rainbow K_3) achieves the inducibility of the rainbow triangle. See also [4] for similar computations, but in terms of the number of edges of each colour instead of the number of vertices.

Huang [9] asked whether Theorem 1·2 can be extended to the undirected setting for cliques of size larger than three. This, in particular, would imply that the phenomenon conjectured by Erdös and Sós and proved in [2] (that K_k is not a fractaliser for $k = 3$) fails to hold for larger k. Our first result addresses Huang's question and proves that rainbow K_k are fractalisers for $k \geq 11$. Note that the result below applies to a specific rainbow colouring of a clique. The corresponding problem for the collection of all rainbow colourings is discussed in Section 5.

THEOREM 1·3. *All rainbow cliques R on $k \geq 11$ vertices are fractalisers. In particular,*

$$\mathrm{ind}(R) = \frac{k!}{k^k - k}.$$

We make the following observations regarding Theorem 1·3.

(1) Theorem 1·3 implies Theorem 1·2 for $k \geq 11$, since any construction of a tournament inducing ℓ rainbow copies of R yields a corresponding construction of a complete graph that induces at least ℓ rainbow (undirected) copies of R by ignoring orientations.

(2) Similarly, if a graph G on k vertices is known to have inducibility $k!/(k^k - k)$, then the rainbow k-clique is a fractaliser as well by the following argument. Let R be a rainbow k-clique and let $e_1, e_2, \ldots, e_m \in E(R)$ such that $(V(R), \{e_1, \ldots, e_m\})$ is a rainbow copy of G. Let $c_1, \ldots, c_m$ be the colours assigned to $e_1, \ldots, e_m$, respectively. Then any construction of an edge-coloured graph inducing ℓ rainbow copies of R induces at least ℓ copies of G by deleting all edges except those coloured by $c_1, \ldots, c_m$ and then ignoring the edge colours. It follows that $\mathrm{ind}\,(R) \leq \mathrm{ind}\,(G) = k!/(k^k - k)$, so R is a fractaliser. Thus, the result of [1] that C_5 is a fractaliser with $\mathrm{ind}\,(C_5) = 5!/(5^5 - 5)$ implies that the rainbow 5-clique is a fractaliser, and the result of [6] that random graphs are almost surely fractalisers implies that the rainbow k-clique is a fractaliser for large k.

(3) We believe our proof of Theorem 1·3 has been optimised and requires $k \geq 11$. As [1] showed that rainbow 5-cliques are fractalisers, and [2] showed that rainbow 3-cliques

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

are not fractalisers, it remains open to determine whether rainbow k_0-cliques are fractalisers only for $k_0 \in \{4, 6, 7, 8, 9, 10\}$.

Our proof of Theorem 1·3 follows the broad framework of the proof of Theorem 1·2 but there are several nontrivial technical difficulties that need to be addressed in the undirected setting. The difficulties arise due to the following reason: the role that each endpoint of an edge plays in a rainbow copy of a tournament is determined by the colour and orientation of the edge, but this is no longer true in the undirected setting. We overcome these obstacles by adding some new ideas, at the expense of requiring a slightly higher value of k. For example, our proof of Theorem 1·3 requires a bound on the colour degree of a vertex and this was not needed in [**12**].

For large values of k, we prove the following more general result which shows that the analog of Theorem 1·3 holds for much sparser graphs. The proof requires several major new ideas.

THEOREM 1·4. *There exists an absolute constant $C > 0$ such that all connected rainbow graphs R with k vertices and minimum degree at least $C \log k$ are fractalisers and satisfy*

$$\mathrm{ind}(R) = \frac{k!}{k^k - k}.$$

We make the following observations regarding Theorem 1·4.

(1) Theorem 1·4 implies Theorem 1·3 for large k, since R may be viewed as a rainbow k-clique with edges deleted. Let $c_1, c_2, \ldots, c_m$ be the colours assigned to the deleted edges. Any construction of a coloured complete graph inducing ℓ rainbow copies of the rainbow k-clique yields at least ℓ rainbow copies of R by deleting edges coloured $c_1, \ldots, c_m$.

(2) The requirement that R is connected in the statement of Theorem 1·4 is necessary, as disconnected rainbow graphs without isolated vertices are not fractalisers (see Section 4).

(3) We are not able to show that our requirement on minimum degree is tight, and this remains open.

Theorem 1·3 is proven in Section 2 and Theorem 1·4 is proven in Section 3. In Section 4, we justify the second observation above. In Section 5, we discuss the problem of maximising the number of rainbow cliques on k vertices in a large graph, where the colouring of the clique is not specified.

2. *Proof of Theorem 1·3*

We give the proof of Theorem 1·3 in the following subsections.

2·1. *Setup*

Fix $k \geq 11$ and $T = \binom{[k]}{2}$. Let R be a T-coloured rainbow k-clique with colouring function χ_R and for concreteness, put $V(R) := [k]$ and $\chi_R(ij) = \{i, j\}$ for all $i, j \in [k]$.

Set

$$a := \frac{k!}{k^k - k}.$$

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

Our aim is to prove that $\text{ind}(R) \le a$. To this end, fix $\gamma > 0$ and assume for contradiction $\text{ind}(R) = a + \gamma$. Next choose $0 < \varepsilon < \min\{\gamma, \text{ind}(R)\}/100$. Let c_0 be chosen so that $\text{ind}(R, n) \le \text{ind}(R) + \varepsilon$ for all $n > c_0$. Choose $M \ge \lceil 2k! c_0/\varepsilon \rceil$ such that

$$\frac{n^k}{(n)_k} < 1 + \varepsilon \tag{2·1}$$

and

$$a\left(\frac{n^{k-1}}{(k-1)!} - \binom{n-1}{k-1}\right) < \gamma\binom{n-1}{k-1} - \binom{n-2}{k-2} \tag{2·2}$$

for all $n > M$. This is possible since $\lim_{n\to\infty} n^k/(n)_k = 1$ and $n^{k-1}/(k-1)! - \binom{n-1}{k-1} = O(n^{k-2})$, while $\gamma\binom{n-1}{k-1} - \binom{n-2}{k-2} = \Omega\left(n^{k-1}\right)$ as $n \to \infty$. Suppose that $n > M$ is given and H is a T-coloured n-vertex graph with colouring function χ_H achieving $I(R, n)$. This implies

$$I(R, H) = I(R, n) = \text{ind}(R, n)\binom{n}{k},$$

where $a + \gamma = \text{ind}(R) \le \text{ind}(R, n) \le \text{ind}(R) + \varepsilon = a + \gamma + \varepsilon$.

Definition 2·1. For $q \ge 0$ and $t > 0$, let $p(q, t)$ be the maximum of $\prod_i q_i$ where $q_1 + \cdots + q_t = q$ and each $q_i \ge 0$ an integer.

The AMGM inequality yields $p(q, t) \le (q/t)^t$ and it is easy to see that

$$p(q, t)p(q', t') \le p(q + q', t + t') \tag{2·3}$$

for all $q, q' \ge 0$ and $t, t' > 0$ (see Appendix).

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. We will refer to this as the *degree of x in H*. Similarly, let $d(x, y)$ be the number of copies of R containing both x and y. For $i \in [k]$, let $N_i(x)$ be the set of $y \in V(H) \setminus \{x\}$ 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. Note that we do not have $N_j(x) \cap N_{j'}(x) = \emptyset$ for $j \ne j'$, but all edges between $N_j(x) \cap N_{j'}(x)$ and x have the same colour $\{j, j'\}$. However, $N_i(x)$ has a (unique) partition $\cup_{j\ne i} N_i^j(x)$ 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. Indeed, the partition is obtained based on the colour of a vertex to x. This gives

$$d(x) = \sum_{i=1}^{k} d_i(x) \le \sum_{i=1}^{k} \prod_{j\ne i} |N_i^j(x)| \le \sum_{i=1}^{k} p(|N_i(x)|, k-1). \tag{2·4}$$

We partition $V(H)$ into $V_1 \cup \cdots \cup V_k$, where

$$V_i = \{x \in V(H) : |N_i(x)| \ge |N_j(x)| \text{ for all } j \ne i\}.$$

If there is a tie, we break it arbitrarily. Set $n_i = |V_i|$ for all $i \in [k]$.

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

2·2. *Minimum degree*

Here we show that a standard technique in extremal graph theory can be used to prove that each vertex of H lies in at least the average number of copies of R (apart from a small error term).

LEMMA 2·2. *$d(x) \geq an^{k-1}/(k-1)!$ for all $x \in V(H)$.*

Proof. We write $d = b \pm c$ for the inequalities $b - c \leq d \leq b + c$. Denote the average degree of H by

$$d(H) := \frac{k \cdot I(R,H)}{n} = \text{ind}\,(R,n)\binom{n-1}{k-1}.$$

We claim that for every $x \in V(H)$

$$d(x) = d(H) \pm \binom{n-2}{k-2}. \tag{2·5}$$

This follows from a standard application of Zykov symmetrisation. Indeed, if the degrees of two vertices x and y differ by more than $\binom{n-2}{k-2}$, say $d(x) > d(y) + \binom{n-2}{k-2}$, then we can delete y and duplicate x, meaning we add a new vertex x' with $\chi_H(x'z) = \chi_H(xz)$ for all other vertices z, and $\chi_H(xx')$ can be arbitrary. This transformation increases the number of copies of R by at least

$$d(x) - d(y) - d(x,y) \geq d(x) - d(y) - \binom{n-2}{k-2} > 0,$$

contradicting the maximality $I(R,H) = I(R,n)$. Hence all degrees lie in an interval of length at most $\binom{n-2}{k-2}$ and (2·5) follows, since this interval must contain $d(H)$. In particular, the minimum degree is at least

$$d(H) - \binom{n-2}{k-2} = \text{ind}\,(R,n)\binom{n-1}{k-1} - \binom{n-2}{k-2}.$$

It follows from (2·2) that

$$\text{ind}\,(R,n)\binom{n-1}{k-1} - \binom{n-2}{k-2} \geq (a+\gamma)\binom{n-1}{k-1} - \binom{n-2}{k-2} > a\frac{n^{k-1}}{(k-1)!}$$

for $n > M$, completing the proof.

2·3. *Maximum colour degree*

Let

$$\alpha := \frac{\max_{x,i,j} d_{\{i,j\}}(x)}{n},$$

where the maximum is taken over all vertices $x \in V(H)$ and all colours $\{i,j\} \in T$ and $d_{\{i,j\}}(x)$ is the number of edges in H incident with x in colour $\{i,j\}$. We upper bound this value.

LEMMA 2·3. *$\alpha \leq 0.4$.*

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

Proof. Let x, i, j achieve this maximum, so that $d_{\{i,j\}}(x) = \alpha n$. Then $|N_i^j(x)| \le \alpha n$ and $N_i(x)$ has a partition $N_i^j(x) \bigcup \cup_{\ell \neq j} N_i^\ell(x)$ where every copy of R containing x with x playing the role of i has exactly one vertex in each $N_i^\ell(x)$ for all $\ell \in [k] \setminus \{i\}$. Further,

$$\left| \bigcup_{\ell \neq j} N_i^\ell(x) \right| \le n - d_{\{i,j\}}(x)$$

since a vertex incident to an edge coloured $\{i,j\}$ cannot play the role of $\ell \neq i,j$. Consequently,

$$d_i(x) \le |N_i^j(x)| \cdot p\left(n - d_{\{i,j\}}(x), k-2\right) \le \alpha\, n \cdot \left(\frac{(1-\alpha)n}{k-2}\right)^{k-2}.$$

The same upper bound holds for $d_j(x)$. For $\ell \notin \{i,j\}$, we have $N_\ell(x) \le n - d_{\{i,j\}}(x)$ since x is playing the role of ℓ, so we cannot include an edge incident to x of colour $\{i,j\}$ since the colour must include ℓ. Hence

$$d_\ell(x) \le p\left(n - d_{\{i,j\}}(x), k-1\right) \le \left(\frac{(1-\alpha)n}{k-1}\right)^{k-1} \le \left(\frac{(1-\alpha)n}{k-2}\right)^{k-1}.$$

Altogether this yields

$$d(x) \le 2\,\alpha\, n \left(\frac{(1-\alpha)n}{k-2}\right)^{k-2} + (k-2)\left(\frac{(1-\alpha)n}{k-2}\right)^{k-1} = (1+\alpha)\left(\frac{1-\alpha}{k-2}\right)^{k-2} n^{k-1}.$$

Suppose for contradiction that $\alpha > 0.4$. Since $k \ge 3$, $(1+\alpha)(1-\alpha)^{k-2}$ is a decreasing function of α for $\alpha \in (0.4, 1]$, and $d(x) \ge a n^{k-1}/(k-1)!$ by Lemma 2·2. Therefore

$$\frac{1}{k^{k-1}-1} = \frac{a}{(k-1)!} \le \frac{d(x)}{n^{k-1}} \le 1.4\left(\frac{0.6}{k-2}\right)^{k-2}. \tag{2·6}$$

However, this fails to hold for $k \ge 11$ (see Appendix), and we conclude that $\alpha \le 0.4$ as desired.

2·4. *The second largest neighbourhood*

For a vertex $x \in V(H)$, let $Z(x)$ be the second largest set in $\{N_1(x), \ldots, N_k(x)\}$ and define

$$z := z_{k,n} = \max_{x \in V(H)} \frac{|Z(x)|}{n}.$$

LEMMA 2·4. $z \le 0.5$.

Proof. Let x be such that $z = |Z(x)|/n$. Let $a_i = |N_i(x)|/n$ and assume by relabeling that $a_1 \ge a_2 = z \ge a_3 \ge \cdots \ge a_k$. Since $N_j(x) \cap N_{j'}(x) \cap N_{j''}(x) = \emptyset$ for any three distinct j, j', j' we have $\sum a_i \le 2$. Let $a_3 + \cdots + a_k = s \le 2 - (a_1 + z)$. Write $s = qz + r$ where $q \in \mathbb{Z}^{\ge 0}$ and $0 \le r < z$. If $x \le y$, then $x^{k-1} + y^{k-1} < (x-\rho)^{k-1} + (y+\rho)^{k-1}$ for $0 < \rho < x$ by convexity of x^{k-1} so successively increasing the largest a_i to z and decreasing the smallest a_j to 0 or r,

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

we obtain

$$\sum_{i=3}^{k} a_i^{k-1} \le qz^{k-1} + r^{k-1} \le qz^{k-1} + \frac{r}{z}z^{k-1} = \frac{s}{z}z^{k-1} \le \frac{2-(a_1+z)}{z}z^{k-1}.$$

Consequently,

$$\sum_{i=1}^{k} a_i^{k-1} = a_1^{k-1} + z^{k-1} + \sum_{i=3}^{k} a_i^{k-1} \le a_1^{k-1} + z^{k-1} + \frac{2-(a_1+z)}{z}z^{k-1}.$$

Since $a_1 \ge z$, taking the derivative shows that for any z, this expression is increasing with a_1. Using Lemma 2·3, we note that $a_1 + z \le 1 + \alpha \le 1.4$ since

$$|N_1(x)| + |Z(x)| = |N_1(x) \cup Z(x)| + |N_1(x) \cap Z(x)| \le n + d_{\{1,2\}}(x),$$

where $d_{\{1,2\}}(x) \le \alpha n \le 0.4 \cdot n$. Thus $a_1 \le 1.4 - z$ and $a_1 \le 1$, so

$$\sum_{i=1}^{k} a_i^{k-1} \le a_1^{k-1} + z^{k-1} + \frac{2-(a_1+z)}{z}z^{k-1} \le \min\{1.4-z, 1\}^{k-1} + z^{k-1} + \frac{0.6}{z}z^{k-1}.$$

Using (2·4) and Lemma 2·2 yields

$$\frac{1}{k^{k-1}-1} \le \frac{d(x)}{n^{k-1}} \le \sum_{i=1}^{k}\left(\frac{a_i}{k-1}\right)^{k-1} \le \frac{\min\{1.4-z,1\}^{k-1} + z^{k-1} + \frac{0.6}{z}z^{k-1}}{(k-1)^{k-1}}.$$

Multiplying by $(k-1)^{k-1}$ and using the fact that

$$\frac{(k-1)^{k-1}}{k^{k-1}-1} \ge \frac{(k-1)^{k-1}}{k^{k-1}} = \left(1-\frac{1}{k}\right)^{k-1} > \frac{1}{e} \tag{2·7}$$

for $k > 1$, we obtain

$$\frac{1}{e} < (\min\{1.4-z,1\})^{k-1} + z^{k-1} + 0.6 \cdot z^{k-2}.$$

As $z \le a_1$ and $z + a_1 < 1.4$, we have $z < 0.7$. Thus, the RHS is nonincreasing with k and we may consider only the $k = 11$ case. Numerical calculations show that for $z \in [0.5, 0.7]$, we have $(1.4-z)^{10} + z^{10} + 0.6z^9 < 1/e$, so we conclude that $z < 0.5$.

2·5. *One large part*

We now take care of the situation when one of the V_i's is very large.

LEMMA 2·5. *$|V_i| \le (1 - 1/3k)n$ for all $i \in [k]$.*

Proof. By contradiction, w.l.o.g. suppose that $|V_1| > (1-1/3k)n$. If $x \in V_1$, then $|N_1(x)| \ge |N_i(x)|$ for all $i > 1$ so $|N_2(x)| \le |Z(x)| \le zn$. Using (2·4) we have

$$a\binom{n}{k} \le I(R,H) = \sum_{x \in V(H)} d_2(x) \le |V_1|p(zn, k-1) + \frac{n}{3k}p(n, k-1)$$

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

and we further see that

$$|V_1|p(zn,k-1)+\frac{n}{3k}p(n,k-1)<\left(z^{k-1}+\frac{1}{3k}\right)\frac{n^k}{(k-1)^{k-1}}.$$

Using our lower bound on M in (2·1), we get

$$\left(\frac{(k-1)^{k-1}}{k^k-k}\right)<(1+\varepsilon)\left(z^{k-1}+\frac{1}{3k}\right). \tag{2·8}$$

This fails to hold for $k\geq 11$ (see Appendix). We conclude that $|V_i|\leq(1-1/3k)n$ for all $i\in[k]$.

2·6. *Counting the copies of R in H*

Here we describe the broad framework we will use to count copies of R in H. This is the same as in [12], though there are subtle differences which arise since we are in the undirected setting.

Call a copy f of R in H *transversal* if it includes exactly one vertex in V_i for all $i\in[k]$. We 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 colouring 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 (including those transversal copies where some vertex is 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

$$h_m=\sum_j I(R,H[V_j])\leq\sum_j I(R,n_j). \tag{2·9}$$

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 colouring does not coincide with the natural one given by the vertex partition $V_1\cup\ldots\cup V_k$. Then

$$h_g=\prod_i n_i-\Delta \tag{2·10}$$

and we need to bound Δ from below.

Note that the colour 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 $\{v_i,v_j\}$ where $v_i\in V_i$, $v_j\in V_j$, $i\neq j$ such that $\chi_H(v_iv_j)\neq\chi_R(ij)=\{i,j\}$. 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 $i<j$,

$$\Delta\geq|D_{ij}|\prod_{\ell\neq i,j}n_\ell=\delta_{ij}\binom{n}{2}\prod_{\ell\neq i,j}n_\ell=\delta_{ij}\binom{n}{2}\frac{\prod_{\ell=1}^k n_\ell}{n_in_j}.$$

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

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

$$\Delta\left(\sum_{1\le i<j\le k} n_i n_j\right) \ge \delta\binom{n}{2}\prod_{\ell=1}^{k} n_\ell.$$

This with along with (2·10) gives

$$h_g \le \prod_{\ell=1}^{k} n_\ell\left(1 - \frac{\delta\binom{n}{2}}{\sum_{1\le i<j\le k} n_i n_j}\right) = \prod_{\ell=1}^{k} n_\ell\left(1 - \frac{\delta\binom{n}{2}}{\binom{n}{2} - \sum_i \binom{n_i}{2}}\right). \tag{2·11}$$

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. 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 $i \notin \chi_H(vw)$. This means that v must play the role of i' in f for some $i' \ne i$, so the colours on all $k-1$ pairs (v, x) with $x \in f$ contain i'; in particular v is incident to $k-2$ pairs in f whose colour does not contain i. If $v \in V_i$ and $w \in V_j$, then say that (v, w, f) is 1-sided if $|\chi_H(vw) \cap \{i, j\}| = 1$ and (v, w, f) is 2-sided if $|\chi_H(vw) \cap \{i, j\}| = 0$.

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 miscoloured vertex which yields at least $k-2$ pairs from D in f. If f is not transversal, choose $j \in [k]$ such that $|f \cap V_j|$ is largest. Set $\mathcal{C} := f \cap V_j$ and observe that at least $|\mathcal{C}| - 1$ of the vertices in $\mathcal{C}$ are miscoloured. Also, note that $2 \le |\mathcal{C}| \le k-1$ since f is not contained in one colour class and we have assumed f is not transversal.

If exactly $|\mathcal{C}| - 1$ vertices in $\mathcal{C}$ are miscoloured, then every edge vw where $v \in \mathcal{C}$ is miscoloured and $w \in f \setminus \mathcal{C}$ is in D. Since $|f \setminus \mathcal{C}| = k - |\mathcal{C}|$, this yields at least $(|\mathcal{C}| - 1)(k - |\mathcal{C}|) \ge k-2$ pairs from D in f. On the other hand, if all $|\mathcal{C}|$ vertices in $\mathcal{C}$ are miscoloured, then there is a unique vertex $u \in f \setminus \mathcal{C}$ that plays the role of j in f. Every edge vw where $v \in \mathcal{C}$ and $w \in f \setminus (\mathcal{C} \cup u)$ is in D, so if $|\mathcal{C}| \le k-2$, this yields at least $|\mathcal{C}|(k - |\mathcal{C}| - 1) \ge k-2$ pairs from D in f. If $|\mathcal{C}| = k-1$, then $f = \mathcal{C} \cup u$ where u plays vertex j in f but is in a different colour class, say the colour class corresponding to colour ℓ. There are $k-1$ edges between $\mathcal{C}$ and u, but only one can contain both k and ℓ, so at least $k-2$ edges from D are 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 $e = \{v, w\} \in D$ 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

$$S \ge 2(k-2)h_b. \tag{2·12}$$

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 $v \in V_i$, and $i \in \chi_H(vw)$; if v is not correct in e then $i \notin \chi_H(vw)$ and 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

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

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

The crucial observation is that

$$(v,w,f)\in J_i \qquad \Longrightarrow \qquad \{v,w\}\in D_i \qquad\qquad (i=1,2). \tag{2·13}$$

To bound S from above, we use (2·13) and consider first J_1 and D_1. We start by choosing vw in D_1 with wrong vertex v. Note that w is correct in vw since $vw\in D_1$. Let $v\in V_i, w\in V_j$. Then $\chi_H(vw)=\{j,\ell\}$ for some $\ell\neq i$ since v is wrong in e but w is correct in e. Thus for each triple $(v,w,f)\in J_1$, vertex v plays the role of j in f or v plays the role of ℓ in f; thus the total number of $(v,w,f)\in J_1$ for some f is at most $p(|N_j(v)|-1,k-2)+p(|N_\ell(v)|-1,k-2)$. Summing over all $vw\in D_1$, we get

$$|J_1|\le \sum_{vw\in D_1} p(|N_j(v)|-1,k-2)+p(|N_\ell(v)|-1,k-2)\le 2|D_1|p(zn,k-2).$$

The bound for J_2 is similar. Choose $vw\in D_2$ with $v\in V_i, w\in V_j$. Let $\chi_H(vw)=\{\ell_1,\ell_2\}$ where $\{\ell_1,\ell_2\}\cap\{i,j\}=\emptyset$. Since vw is two-sided, we see that $(v,w,f)\in J_2$ exactly when $(w,v,f)\in J_2$. Consequently,

$$\begin{aligned}|J_2| &\le \sum_{vw\in D_2} p(|N_{\ell_1}(v)|-1,k-2)+p(|N_{\ell_2}(v)|-1,k-2)\\ &\quad + p(|N_{\ell_1}(w)|-1,k-2)+p(|N_{\ell_2}(w)|-1,k-2)\\ &\le 4|D_2|p(zn,k-2).\end{aligned}$$

This gives

$$S=2|J_1|+|J_2|\le 4\,|D|\,p(zn,k-2)\le 4\,\delta\binom{n}{2}\left(\frac{z}{k-2}\right)^{k-2}n^{k-2}. \tag{2·14}$$

Finally, (2·12) and (2·14) give

$$h_b\le \frac{S}{2(k-2)}\le \frac{2\delta\binom{n}{2}}{k-2}\left(\frac{z}{k-2}\right)^{k-2}n^{k-2}. \tag{2·15}$$

Using (2·9), (2·11) and (2·15) we have that

$$I(R,n)\le \sum_i I(R,n_i)+\prod_\ell n_\ell\left(1-\frac{\delta\binom{n}{2}}{\binom{n}{2}-\sum_i\binom{n_i}{2}}\right)+\frac{2\delta\binom{n}{2}}{k-2}\left(\frac{z}{k-2}\right)^{k-2}n^{k-2}. \tag{2·16}$$

Our final task is to upper bound the RHS.

Since $\delta\binom{n}{2}\le\sum_{i\neq j}n_in_j=\binom{n}{2}-\sum_i\binom{n_i}{2}$, we have $\delta\in I\stackrel{\text{def}}{=}\left[0,1-\sum_i\binom{n_i}{2}/\binom{n}{2}\right]$. Viewing (2·16) as a linear function of δ, it suffices to check the endpoints of I.

2·7. *The extremal case*

CLAIM 2·6. *If $\delta=0$, then* $\mathrm{ind}\,(R)\le a$.

Proof. If $\delta=0$, then (2·16) implies that

$$I(R,n)\le \sum_{i=1}^{k} I(R,n_i)+\prod_{i=1}^{k} n_i. \tag{2·17}$$

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

Let $p_i := n_i/n$. Using $\max_i p_i \le 1 - 1/3k$ by Lemma 2·5, convexity of x^k, and $k \ge 11$ we obtain

$$\sum_{i=1}^{k} p_i^k \le \left(1 - \frac{1}{3k}\right)^k + \left(\frac{1}{3k}\right)^k \le e^{-1/3} + 33^{-11} < 0.72. \tag{2·18}$$

We begin by bounding the summation in (2·17). By relabeling if necessary, let $n_1 \le \cdots \le n_\ell \le c_0 < n_{\ell+1} \le \cdots \le n_k$ where $\ell \ge 0$. We have that

$$\sum_{i=1}^{k} I(R, n_i) \le \ell \binom{c_0}{k} + \sum_{i=\ell+1}^{k} I(R, n_i) \le \ell \binom{c_0}{k} + (\operatorname{ind}(R) + \varepsilon) \sum_{i=\ell+1}^{k} \binom{n_i}{k}. \tag{2·19}$$

Observe that $\binom{n_i}{k} = \binom{p_i n}{k} < p_i^k \binom{n}{k}$ since $p_i < 1$. Dividing (2·19) by $\binom{n}{k}$ yields

$$\frac{1}{\binom{n}{k}} \sum_{i=1}^{k} I(R, n_i) \le \ell \frac{\binom{c_0}{k}}{\binom{n}{k}} + (\operatorname{ind}(R) + \varepsilon) \sum_{i=\ell+1}^{k} p_i^k. \tag{2·20}$$

Suppose $\ell \ge 1$. Using our bounds on ε and M and (2·18), we can further bound

$$\frac{1}{\binom{n}{k}} \sum_{i=1}^{k} I(R, n_i) \le \ell \frac{\binom{c_0}{k}}{\binom{n}{k}} + (\operatorname{ind}(R) + \varepsilon) \sum_{i=\ell+1}^{k} p_i^k < 0.74 \operatorname{ind}(R) \tag{2·21}$$

and bound the product term

$$\frac{1}{\binom{n}{k}} \prod_{i=1}^{k} n_i \le \frac{1}{\binom{n}{k}} c_0 n^{k-1} < \frac{2k! c_0}{n} < \varepsilon.$$

This yields $\operatorname{ind}(R, n) \le 0.74 \operatorname{ind}(R) + \varepsilon < \operatorname{ind}(R)$, a contradiction. Thus $\ell = 0$, so using (2·20) we may rewrite (2·17) as

$$\operatorname{ind}(R, n) \le (\operatorname{ind}(R) + \varepsilon) \sum_{i=1}^{k} p_i^k + \frac{1}{\binom{n}{k}} \prod_{i=1}^{k} n_i. \tag{2·22}$$

Isolating the product term and recalling the definition of a, as well as our lower bound on M,

$$\frac{1}{\binom{n}{k}} \prod_{i=1}^{k} n_i = \frac{n^k}{\binom{n}{k}} \prod_{i=1}^{k} p_i \le (a + \varepsilon)(k^k - k) \prod_{i=1}^{k} p_i.$$

Plugging this into (2·22) and recalling $\operatorname{ind}(R) = a + \gamma$,

$$\operatorname{ind}(R, n) \le (a + \varepsilon) \left(\sum_{i=1}^{k} p_i^k + (k^k - k) \prod_{i=1}^{k} p_i \right) + \gamma \sum_{i=1}^{k} p_i^k \le (a + \varepsilon) + 0.72\gamma.$$

The first bound $\sum p_i^k + (k^k - k) \prod p_i \le 1$ is well-known (see, e.g. (17) in [**12**]) and the second bound comes from (2·18). This gives the contradiction

$$a + \gamma = \operatorname{ind}(R) \le \operatorname{ind}(R, n) \le a + 0.72\gamma + \varepsilon$$

since $\varepsilon < \gamma/100$.

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

2·8. *The absurd case*

Now, we consider the other endpoint of I.

CLAIM 2·7. *If* $\delta = 1 - \sum_i \binom{n_i}{2}/\binom{n}{2}$, *then* $\text{ind}\,(R) \le a$.

Proof. If $\delta = 1 - \sum_i \binom{n_i}{2}/\binom{n}{2}$, then (2·16) implies that

$$I(R,n) \le \sum_{i=1}^{k} I(R,n_i) + \frac{2\sum_{i\neq j} n_i n_j}{k-2}\left(\frac{z}{k-2}\right)^{k-2} n^{k-2}. \tag{2·23}$$

We first bound the second term. Dividing by $\binom{n}{k}$ and again letting $p_i := n_i/n$, we reorganise

$$\frac{2}{\binom{n}{k}}\cdot\frac{\sum n_i n_j}{k-2}\left(\frac{z}{k-2}\right)^{k-2} n^{k-2} = 2\cdot\frac{k^k-k}{(k-2)^{k-1}}\cdot\frac{n^k}{(n)_k}\cdot\left(\sum_{i\neq j} p_i p_j\right) z^{k-2} a.$$

Observe that $(k^{k-1}-1)/(k-2)^{k-1}$ decreases to e^2. In particular, for $k \ge 11$, we have $(k^k-k)/(k-2)^{k-1} \le 7.5k$. For $n > M$, we have $n^k/(n)_k < 1+\varepsilon$. Finally, $\sum_{i\neq j} p_i p_j = (1 - \sum p_i^2)/2 \le (1-1/k)/2$ as $\sum p_i^2$ is minimised when $p_i = 1/k$ for all i. Thus

$$\frac{2}{\binom{n}{k}}\cdot\frac{\sum n_i n_j}{k-2}\left(\frac{z}{k-2}\right)^{k-2} n^{k-2} \le 7.5(1+\varepsilon)(k-1)z^{k-2}a < 0.25a$$

for $k \ge 11$ as $(k-1)z^{k-2}$ is decreasing in k and $(11-1)z^{11-2} < 10\cdot 2^{-9} < 1/50$. Using this and (2·21) in (2·23), and $a <$ ind (R) gives

$$\text{ind}\,(R,n) \le 0.74\ \text{ind}\,(R) + 0.25a < 0.99\ \text{ind}\,(R).$$

This contradiction completes the proof of the claim and the theorem.

3. *Proof of Theorem 1·4*

We give the proof of Theorem 1·4 in the following subsections.

3·1. *Setup*

Fix k and $R = ([k], E)$ a rainbow coloured graph with minimum degree at least $\eta(k-1)$ where $\eta > C\log k/(k-1)$. We may assume that k is sufficiently large by making C sufficiently large so that the theorem is vacuous for small k. In particular, we will assume $k \ge 11$ so that we may use the same bounds as the previous section. It is notationally convenient to set $T = E \cup \{\emptyset\}$ and view R as a T-coloured complete graph $([k], \binom{[k]}{2})$ with colouring function χ_R defined as follows:

$$\chi_R(ij) = \begin{cases} \{i,j\} & ij \in E \\ \emptyset & ij \notin E. \end{cases}$$

Our goal is to prove that $\text{ind}\,(R) \le a$. To this end, fix $\gamma > 0$ and assume for contradiction $\text{ind}\,(R) = a + \gamma$. Next choose ε, c_0, M as in Section 2.1.

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

Suppose that $n > M$ is given and H is a T-coloured n-vertex graph with colouring function χ_H achieving $I(R, n)$. This implies

$$I(R, H) = I(R, n) = \text{ind}\,(R, n)\binom{n}{k},$$

where $a + \gamma = \text{ind}\,(R) \leq \text{ind}\,(R, n) \leq \text{ind}\,(R) + \varepsilon = a + \gamma + \varepsilon$.

Let $d_i(x)$, $d(x)$, $d(x, y)$, $d_{\{i,j\}}(x)$, $N_i(x)$, and $N_i^j(x)$ be defined as in Section 2. Note that we do not have that all vertices in $N_j(x) \cap N_{j'}(x)$ for $j \neq j'$ have the same colour to x as it may be the case that $\chi_H(xy) = \emptyset$ and $\chi_H(xy') = \{j, j'\}$ for distinct $y, y' \in N_j(x) \cap N_{j'}(x)$. We also do not have that $\cup_{j \neq i} N_i^j(x)$ is a partition of $N_i(x)$ as it may be the case that $y \in N_i^j(x) \cap N_i^{j'}(x)$ for some $j \neq j'$ satisfying $\chi_R(ij) = \chi_R(ij') = \emptyset$ and $y \in V(H)$ satisfying $\chi_H(xy) = \emptyset$. Thus we must develop new techniques to prove a version of (2·4) from Section 2.1 to obtain bounds on $d_i(x)$. This is the content of Section 3.2.

As in Section 2.1, we partition $V(H)$ into $V_1 \cup \cdots \cup V_k$, $n_i = |V_i|$, where

$$V_i = \{x \in V(H) : |N_i(x)| \geq |N_j(x)| \text{ for all } j \neq i\}.$$

If there is a tie, we break it arbitrarily.

3·2. *Partitioning argument*

Let the *distance* between two vertices v and w in a graph G, denoted $\text{dist}_G(v, w)$, be the number of edges in the shortest path between v and w in G. In our setting, a path cannot use an edge e with $\chi(e) = \emptyset$. Then, define

$$\epsilon_G(v) := \max_{w \in V(G)} \text{dist}_G(v, w),$$

the *eccentricity* of v in G. Note that the diameter $\text{diam}\,(G) = \max_{v \in V(G)} \epsilon_G(v)$. For convenience, let $\epsilon(i) := \epsilon_R(i)$ for all $i \in [k]$.

Let $B(x)$ be x in H. For $r \in \mathbb{N}$, let $k_r(i)$ be the number of vertices in R at distance r from i. Recall that $d_{\{i,j\}}(x)$ is the number of edges in H incident with x in colour $\{i, j\}$

LEMMA 3·1. *Let $i, j \in [k]$ with $\{i, j\} \in T$ and $x \in V(H)$. Then*

$$(a)\quad d_i(x) \leq \left(\frac{|B(x)|}{k_1(i)}\right)^{k_1(i)} \left(\frac{n - |B(x)|}{k - k_1(i) - 1}\right)^{k - k_1(i) - 1}$$

$$(b)\quad d_i(x) \leq \left(\frac{|N_i(x)|}{k-1}\right)^{k-1}$$

$$(c)\quad d_i(x) \leq d_{\{i,j\}}(x) \cdot \left(\frac{n - d_{\{i,j\}}(x)}{k-2}\right)^{k-2}.$$

Further, the number of copies of R in H containing vertices $x, y \in V(H)$ such that $x \in V(H)$ plays the role of vertex $i \in [k]$ in R is at most

$$\left(\frac{|N_i(x)|}{k-2}\right)^{k-2}.$$

Proof. We start by proving the three upper bounds on $d_i(x)$. To count the number of copies of R in H where x plays the role of i, we will recursively partition $N_i(x)$. First, we pick

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

$k_1(i)$ vertices from $B(x) \cap N_i(x) \subset V(H)$ to play the role of the vertices adjacent to i in R. Notice that we may partition $B(x) \cap N_i(x)$ into $k_1(i)$ parts based on the colour of each vertex to x as it uniquely determines its possible role in a copy of R. Set $B_1 := B(x) \cap N_i(x)$. We now recursively define B_r for all $r \in [\epsilon(i)]$. Let $2 \leq r \leq \epsilon(i)$, let $m := k_1(i) + \cdots + k_{r-1}(i)$, and suppose that we have chosen $y_1, y_2, \ldots, y_m \in V(H)$ to play the roles of all vertices at distance $r-1$ or less from i in R, where $y_1, \ldots, y_{k_{r-1}(i)}$ play the roles of vertices at distance exactly $r-1$ from i. Then

$$\begin{aligned} B_r &:= B_r(x, B_1, B_2, \ldots, B_{r-1}, y_1, \ldots, y_{k_{r-1}(i)}) \\ &= N_i(x) \cap \left(B(y_1) \cup \cdots \cup B(y_{k_{r-1}(i)})\right) \setminus (x \cup B_1 \cup \cdots \cup B_{r-1}) \,. \end{aligned}$$

Here, B_r is the set of vertices in H that can play the role of vertices at distance r from i in R, given that we have already selected all vertices at distance at most $r-1$ from i.

Note that by definition, $B_r \cap B_\ell = \emptyset$ for all $r, \ell \in [\epsilon(i)]$ and $\bigcup B_r \subseteq N_i(x)$. For the remainder of the proof, we write $k_r := k_r(i)$ for all $r \in [\epsilon(i)]$ for convenience.

Each vertex $v \in B_r$ has an edge to at least one of $y_1, \ldots, y_{k_{r-1}}$. The colour of this edge uniquely determines the role that v may play in a copy of R, so this allows us to uniquely partition B_r into k_r parts. We note that it may be the case that v cannot legally play any role, but that only decreases the number of possible copies of R, so we may assume that this does not occur. Let $P_r := P(B_r, k_r)$ be the set of tuples $\vec{y} \in B_r^{k_r}$ with one vertex from each part of B_r, so

$$|P_r| \leq p(|B_r|, k_r).$$

This gives

$$\begin{aligned} d_i(x) &\leq \sum_{\vec{y}_1 \in P_1} \cdots \sum_{\vec{y}_{\epsilon(i)-1} \in P_{\epsilon(i)-1}} p(|B_{\epsilon(i)}|, k_{\epsilon(i)}) \\ &\leq \sum_{\vec{y}_1 \in P_1} \cdots \sum_{\vec{y}_{\epsilon(i)-1} \in P_{\epsilon(i)-1}} p\left(|N_i(x)| - \sum_{r=1}^{\epsilon(i)-1} |B_r|, k_{\epsilon(i)}\right). \end{aligned}$$

Using (2·3) from Section 2.1 we see that

$$\begin{aligned} &\sum_{\vec{y}_{\epsilon(i)-1} \in P_{\epsilon(i)-1}} p\left(|N_i(x)| - \sum_{r=1}^{\epsilon(i)-1} |B_r|, k_{\epsilon(i)}\right) \\ &\qquad \leq p\left(|B_{\epsilon(i)-1}|, k_{\epsilon(i)-1}\right) \cdot p\left(|N_i(x)| - \sum_{r=1}^{\epsilon(i)-1} |B_r|, k_{\epsilon(i)}\right) \\ &\qquad \leq p\left(|N_i(x)| - \sum_{r=1}^{\epsilon(i)-2} |B_r|, k_{\epsilon(i)-1}(i) + k_{\epsilon(i)}\right). \end{aligned}$$

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

Using $\sum_{r=1}^{\epsilon(i)} k_r = k-1$, we obtain

$$d_i(x) \le \sum_{\vec{y}_1 \in P_1} \cdots \sum_{\vec{y}_{\epsilon(i)-2} \in P_{\epsilon(i)-2}} p\left(|N_i(x)| - \sum_{r=1}^{\epsilon(i)-2} |B_r|, k_{\epsilon(i)-1} + k_{\epsilon(i)}\right)$$
$$= \sum_{\vec{y}_1 \in P_1} \cdots \sum_{\vec{y}_{\epsilon(i)-2} \in P_{\epsilon(i)-2}} p\left(|N_i(x)| - \sum_{r=1}^{\epsilon(i)-2} |B_r|, (k-1) - \sum_{r=1}^{\epsilon(i)-2} k_r\right).$$

Continuing this process, we obtain, for each $1 \le \ell \le \epsilon(i)-1$,

$$d_i(x) \le \sum_{\vec{y}_1 \in P_1} \cdots \sum_{\vec{y}_\ell \in P_\ell} p\left(|N_i(x)| - \sum_{r=1}^{\ell} |B_r|, (k-1) - \sum_{r=1}^{\ell} k_r\right).$$

When $\ell = 1$ this becomes

$$\begin{aligned} d_i(x) &\le \sum_{\vec{y}_1 \in P_1} p\left(|N_i(x)| - |B_1|, k-1-k_1\right) \\ &\le p(|B_1|, k_1) \cdot p\left(|N_i(x)| - |B_1|, k-k_1-1\right). \end{aligned} \tag{3·1}$$

As $|B(x)| \ge |B_1|$ and $n - |B(x)| \ge |N_i(x)| - |B_1|$ for all $i \in [k]$ by definition,

$$p(|B_1|, k_1) \cdot p\left(|N_i(x)| - |B_1|, k-k_1-1\right) \le \left(\frac{|B(x)|}{k_1}\right)^{k_1} \left(\frac{n-|B(x)|}{k-k_1-1}\right)^{k-k_1-1},$$

so (a) holds. Alternatively, (2·3) also yields

$$p(|B(x)|, k_1) \cdot p\left(|N_i(x)| - |B(x)|, k-k_1-1\right) \le p(|N_i(x)|, k-1) \le \left(\frac{|N_i(x)|}{k-1}\right)^{k-1},$$

so (b) holds.

For (c), let $j \in [k]$ such that $ij \in E$. We bound $d_i(x)$ as before, but we choose the vertex y that plays role j separately. We see that

$$|P_1| \le d_{\{i,j\}}(x) \cdot p(|B_1| - d_{\{i,j\}}(x), k_1 - 1).$$

This combined with (3·1) and (2·3) gives

$$\begin{aligned} d_i(x) &\le \sum_{\vec{y}_1 \in P_1} p\left(|N_i(x)| - |B_1|, k-k_1-1\right) \\ &\le d_{\{i,j\}}(x) \cdot p(|B_1| - d_{\{i,j\}}(x), k_1 - 1) \cdot p\left(|N_i(x)| - |B_1|, k-k_1-1\right) \\ &\le d_{\{i,j\}}(x) \cdot p(n - d_{\{i,j\}}(x), k-2) \\ &\le d_{\{i,j\}}(x) \cdot \left(\frac{n-d_{\{i,j\}}(x)}{k-2}\right)^{k-2}. \end{aligned}$$

It remains to prove the last sentence of the lemma. We proceed as before except that, for $\ell \in [\epsilon(i)]$ such that $y \in B_\ell$, we require that y is chosen. This means that instead of choosing $k_\ell(i)$ vertices from B_ℓ, we only need to choose $k_\ell(i) - 1$ vertices from B_ℓ as we have already

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

chosen y. Following the same procedure as before, we see that

$$\begin{aligned}
d_i(x) &\leq \sum_{\vec{y}_1 \in P_1} \cdots \sum_{\vec{y}_\ell \in P_\ell} p\left(|B_\ell|, k_\ell - 1\right) \cdot p\left(|N_i(x)| - \sum_{r=1}^{\ell+1} |B_r|, (k-1) - \sum_{r=1}^{\ell+1} k_r\right) \\
&\leq \sum_{\vec{y}_1 \in P_1} \cdots \sum_{\vec{y}_\ell \in P_\ell} p\left(|N_i(x)| - \sum_{r=1}^{\ell} |B_r|, (k-2) - \sum_{r=1}^{\ell} k_r\right) \\
&\vdots \\
&\leq \sum_{\vec{y}_1 \in P_1} p\left(|N_i(x)| - |B_1|, k - 2 - k_1\right) \\
&\leq p(|B_1|, k_1) \cdot p(|N_i(x)| - |B_1|, k - 2 - k_1) \\
&\leq p(|N_i(x)|, k-2) \\
&\leq \left(\frac{|N_i(x)|}{k-2}\right)^{k-2}.
\end{aligned}$$

This completes the proof.

3·3. *Minimum degree*

As in Section 2.2, we wish to show that each vertex of H lies in approximately the average number of copies of R.

LEMMA 3·2. *$d(x) \geq an^{k-1}/(k-1)!$ for all $x \in V(H)$.*

This follows from an identical Zykov symmetrisation argument as used in the proof of Lemma 2·2. Note that we have assumed the same inequalities for M as we did in Section 2.1.

3·4. *Maximum colour and non-edge degrees*

The following two claims are used in the proof of Lemma 3·5 to bound the size of the second largest neighborhood.

Let

$$\alpha := \frac{\max_{x,i,j} d_{\{i,j\}}(x)}{n}$$

where the maximum is taken over all vertices $x \in V(H)$ and all colours $\{i,j\} \in T$. We upper bound this value.

CLAIM 3·3. $\alpha < \eta/4$.

Proof. Let x achieve this maximum, so that $d_{\{i,j\}}(x) = \alpha n$ for some $\{i,j\} \in T$. By Lemma 3·1(c) and $\alpha \leq 1$, we get

$$\max\{d_i(x), d_j(x)\} \leq \alpha n \left(\frac{(1-\alpha)n}{k-2}\right)^{k-2} \leq n^{k-1} \left(\frac{1-\alpha}{k-2}\right)^{k-2}.$$

For any other vertex $\ell \neq i,j$, we have $|N_\ell(x)| \leq (1-\alpha)n$, since ℓ is adjacent to no edges of colour $\{i,j\}$ in R. Thus by Lemma 3·1(b), we get

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

$$d_\ell(x) \le \left(\frac{|N_\ell(x)|}{k-1}\right)^{k-1} \le \left(\frac{(1-\alpha)n}{k-1}\right)^{k-1} \le n^{k-1}\left(\frac{1-\alpha}{k-2}\right)^{k-2}.$$

The last inequality comes as decreasing the denominator increases the fraction, and the base is less than 1, so decreasing the exponent increases the result. Summing over all indices in $[k]$ and using Lemma 3·2, we get

$$\frac{1}{k^{k-1}-1} = \frac{a}{(k-1)!} \le \frac{d(x)}{n^{k-1}} \le k\left(\frac{1-\alpha}{k-2}\right)^{k-2}.$$

Rearranging yields

$$\frac{1}{k}\cdot\frac{(k-2)^{k-2}}{k^{k-1}-1} \le (1-\alpha)^{k-2}.$$

We see that

$$\frac{(k-2)^{k-2}}{k^{k-1}-1} \ge \frac{(k-2)^{k-2}}{k^{k-1}} = \frac{1}{k}\left(1-\frac{2}{k}\right)^{k-2} > \frac{1}{e^2k},$$

so

$$\frac{1}{e^2k^2} < (1-\alpha)^{k-2} \le \exp(-(k-2)\alpha).$$

Assume for contradiction that $\alpha \ge \eta/4 > C\log k/(4(k-1))$. Then

$$\frac{1}{e^2k^2} < k^{-C(1-1/(k-1))/4} < k^{-0.9C/4},$$

since $k \ge 11$. For $C > 10$, this gives a contradiction for sufficiently large k.

Let

$$\beta := \frac{\max_x d_\emptyset(x)}{n},$$

where the maximum is taken over all vertices $x \in V(H)$ and $d_\emptyset(x)$ is the number of edges in H incident with x in colour $\emptyset$ (non-edges). Note that we may assume that R has at least one non-edge, since otherwise the proof from Section 2 suffices. Thus we may also assume that H has at least one non-edge, so $\beta > 0$. We upper bound β.

CLAIM 3·4. $\beta < 1-\eta/2$.

Proof. Assume for contradiction that $\beta \ge 1-\eta/2$. Let x achieve this maximum so that $d_\emptyset(x) = \beta n$. This implies that $B(x) = (1-\beta)n$. For any $i \in V(R)$, Lemma 3·1(a) gives

$$\begin{aligned} d_i(x) &\le \left(\frac{B(x)}{k_1(i)}\right)^{k_1(i)}\left(\frac{n-B(x)}{k-k_1(i)-1}\right)^{k-k_1(i)-1} \\ &= \left(\frac{(1-\beta)n}{k_1(i)}\right)^{k_1(i)}\left(\frac{\beta n}{k-k_1(i)-1}\right)^{k-k_1(i)-1} \\ &= n^{k-1}\frac{1}{k_1(i)^{k_1(i)}}(1-\beta)^{k_1(i)}\beta^{k-k_1(i)-1}\left(\frac{1}{k-1-k_1(i)}\right)^{k-1-k_1(i)}. \end{aligned}$$

For this section, we take the convention $0^0 = 1$ to handle the case that $k_1(i) = k-1$.

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

Let $q = k_1(i)/(k-1) \in (0,1]$. Then

$$d_i(x) \leq \left(\frac{n}{k-1}\right)^{k-1} \left(\frac{(1-\beta)^q \beta^{1-q}}{q^q (1-q)^{1-q}}\right)^{k-1}. \tag{3·2}$$

We will first bound the term

$$\frac{(1-\beta)^q \beta^{1-q}}{q^q (1-q)^{1-q}}. \tag{3·3}$$

Regarding (3·3) as a function of β, we see that the derivative

$$\frac{\partial}{\partial \beta}\left(\frac{(1-\beta)^q \beta^{1-q}}{q^q (1-q)^{1-q}}\right) = \frac{(1-\beta)^{q-1}\beta^{-q}}{q^q (1-q)^{1-q}}((1-q)-\beta)$$

is negative for $\beta > 1-q$ since the fraction is nonnegative. Recall that $k_1(i) = \deg_R(i) \geq \eta(k-1)$ by assumption, so $q > \eta$. We have also assumed for contradiction that $\beta \geq 1 - \eta/2 > 1 - \eta > 1 - q$. Thus decreasing β to $1 - \eta/2$ will only increase (3·3), i.e.

$$\frac{(1-\beta)^q \beta^{1-q}}{q^q (1-q)^{1-q}} \leq \frac{(\eta/2)^q (1-\eta/2)^{1-q}}{q^q (1-q)^{1-q}}. \tag{3·4}$$

We now have a function purely of q. Taking the derivative, we get

$$\frac{\partial}{\partial q}\left(\frac{(\eta/2)^q (1-\eta/2)^{1-q}}{q^q (1-q)^{1-q}}\right) = \frac{1}{2}(1-q)^{q-1} q^{-q} (2-\eta)^{1-q} \eta^q \log\left(\frac{\eta(1-q)}{q(2-\eta)}\right),$$

where all terms are positive except the logarithm, which is negative for $q > \eta/2$. Thus (3·4) is decreasing with q for $q > \eta/2$, so we may take the further upper bound

$$\frac{(1-\beta)^q \beta^{1-q}}{q^q (1-q)^{1-q}} \leq \frac{(\eta/2)^\eta (1-\eta/2)^{1-\eta}}{\eta^\eta (1-\eta)^{1-\eta}} = 2^{-\eta}\left(1 + \frac{\eta}{2(1-\eta)}\right)^{1-\eta},$$

where $2^{-\eta}\left(1 + \frac{\eta}{2(1-\eta)}\right)^{1-\eta} \leq \exp(-(\log 2 - 1/2)\eta)$.

We now have an appropriate upper bound. Substituting into (3·2) and recalling that $\eta > C \log k/k$, we see that

$$d_i(x) \leq \left(\frac{n}{k-1}\right)^{k-1} \exp(-(\log 2 - 1/2)(k-1)\eta) < \left(\frac{n}{k-1}\right)^{k-1} k^{-C(\log 2 - 1/2)}.$$

Using Lemma 3·2, we get

$$\frac{1}{k^{k-1} - 1} = \frac{a}{(k-1)!} \leq \frac{d(x)}{n^{k-1}} \leq k\left(\frac{1}{k-1}\right)^{k-1} k^{-C(\log 2 - 1/2)}.$$

Rearranging terms and using the standard inequality (2·7) yields

$$\frac{1}{ek} \leq k^{-C(\log 2 - 1/2)}.$$

For $C > 1/(\log 2 - 1/2) \approx 5.18$, this yields a contradiction for sufficiently large k.

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

3·5. *The second largest neighborhood*

For a vertex $x \in V(H)$, let $Z(x)$ be the second largest set in $\{N_1(x), \dots, N_k(x)\}$ and define

$$z := z_{k,n} = \max_{x \in V(H)} \frac{|Z(x)|}{n}.$$

LEMMA 3·5. $z < 1 - \eta/8$.

Proof. Let $x \in V(H)$ such that $z = |Z(x)|/n$. Suppose $x \in V_i$ and $Z(x) = N_j(x)$ for distinct $i, j \in [k]$. Then we want to bound $|N_i(x) \cap Z(x)|$. Suppose $y \in N_i(x) \cap Z(x)$. If $xy \in E$, then $i \in \chi_H(xy)$ and $j \in \chi_H(xy)$, so $\chi_H(xy) = \{i, j\}$. Thus $|N_1(x) \cap Z(x)| \le d_{\{i,j\}}(x) + d_\emptyset(x)$. It follows that

$$|N_i(x)| + |Z(x)| = |N_i(x) \cup Z(x)| + |N_i(x) \cap Z(x)| \le n + d_{\{i,j\}}(x) + d_\emptyset(x),$$

where $d_{\{i,j\}}(x) \le \alpha n$ and $d_\emptyset(x) \le \beta n$. Thus, $|N_i(x)| + |Z(x)| < (2 - \eta/4)n$ by Claims 3·3 and 3·4 Since $|Z(x)| \le |N_i(x)|$, this gives $z < 1 - \eta/8$.

3·6. *One large part*

We now take care of the situation when one of the V_i's is very large.

LEMMA 3·6. $|V_i| \le (1 - 1/3k)n$ *for all* $i \in [k]$.

Proof. By contradiction, w.l.o.g. suppose that $|V_1| > (1 - 1/3k)n$. If $x \in V_1$, then $|N_1(x)| \ge |N_i(x)|$ for all $i > 1$ so $|N_2(x)| \le |Z(x)| \le zn$. Applying Lemma 3·1($b$) to $d_2(x)$ gives

$$a\binom{n}{k} \le I(R, H) = \sum_{x \in V(H)} d_2(x) \le |V_1| \left(\frac{zn}{k-1}\right)^{k-1} + \frac{n}{3k}\left(\frac{n}{k-1}\right)^{k-1}$$

and we further see that

$$|V_1| \left(\frac{zn}{k-1}\right)^{k-1} + \frac{n}{3k}\left(\frac{n}{k-1}\right)^{k-1} \le \left(z^{k-1} + \frac{1}{3k}\right) \frac{n^k}{(k-1)^{k-1}}.$$

Rearranging and using $M^k/(M)_k < 1.01$ as assumed in (2·1), we get

$$\frac{(k-1)^{k-1}}{k^k - k} \le 1.01 \left(z^{k-1} + \frac{1}{3k}\right).$$

Using the standard inequality (2·7) and then Lemma 3·5 gives

$$\left(\frac{1}{1.01e} - \frac{1}{3}\right)\frac{1}{k} < z^{k-1} < \left(1 - \frac{\eta}{8}\right)^{k-1} \le \exp(-C \log k/8) = k^{-C/8}.$$

For any $C > 8$ this fails to hold for sufficiently large k.

3·7. *Counting the copies of R in H*

The way we count copies of R in H is very similar to the previous section and to [**12**]. While we do not have as much information in this case, without a focus on optimising for small k, we allow ourselves to be less strict with the counting arguments.

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

Call a copy f of R in H *transversal* if it includes exactly one vertex in V_i for all $i \in [k]$. We 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 colouring 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 (including those transversal copies where some vertex is in an inappropriate V_i). Thus a transversal copy f is in H_b if and only if the unique map $\phi : [k] \to f$ with $\phi(i) \in V_i$ for all i is not a graph isomorphism from $R \to H[f]$. 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. As in Section 2.6, let D be the set of all pairs $\{v,w\}$ such that $v \in V_i$, $w \in V_j$, and $\chi_H(vw) \neq \chi_R(ij)$ where $i \neq j$. Let $\delta := |D|/\binom{n}{2}$. The identical reasoning as in Section 2.6 gives the first two bounds

$$h_m = \sum_{j=1}^{k} I(R, H[V_j]) \leq \sum_{j=1}^{k} I(R, n_j) \tag{3·5}$$

and

$$h_g \leq \prod_{\ell=1}^{k} n_\ell \left(1 - \frac{\delta\binom{n}{2}}{\sum_{1\leq i<j\leq k} n_i n_j}\right) = \prod_{\ell=1}^{k} n_\ell \left(1 - \frac{\delta\binom{n}{2}}{\binom{n}{2} - \sum_i \binom{n_i}{2}}\right). \tag{3·6}$$

Our next task is to upper bound h_b. This argument must be carried out differently. For a vertex $x \in V(H)$ 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. Let us enumerate the set J of ordered pairs (e,f) where $e \in D, f \in H_b$, and $e \subset f$. To simplify the argument and notation, we count pairs instead of triples as in Section 2.

We must show that each $f \in H_b$ contains an edge in D. If f is transversal, then as we have noted, the natural map is not a graph isomorphism. Thus there is some incorrectly coloured edge which is in D. If f is not transversal, there is some $i \in [k]$ such that $|f \cap V_i| \geq 2$. Note that $f \notin H_m$, so $|f \cap V_i| < k$. As R is connected, there exist $v \in V_i$, $u \in V_j$ for some $j \neq i$ such that vu is an edge in f. Since $|f \cap V_i| \geq 2$, choose also $w \in f \cap V_i$ with $w \neq v$. If $\chi_R(ij) = \emptyset$ then $vu \in D$. If $\chi_R(ij) = \{i,j\}$, then as $\chi_H(vu) = \chi_H(wu) = \{i,j\}$ would contradict that f is a copy of R in H, we must have that uv or uw in D. This gives us that

$$h_b \leq |J|.$$

To bound $|J|$ from above, we start by choosing some bad edge $vw \in D$. Let $f \subset V(H)$ such that $(vw,f) \in J$. Either $v \in V_i$ does not play the role of i or $w \in V_j$ does not play the role of j in f. Then $f \subset N_\ell(v) \cup \{v\}$ for some $\ell \neq i$ or $f \subset N_\ell(w) \cup \{w\}$ for some $\ell \neq j$. We have $|N_\ell(v)|, |N_\ell(w)| \leq zn$ by the definition of z and the partition $V_1 \cup \cdots \cup V_k = V(H)$. By the final statement of Lemma 3·1,

$$|J| \leq \sum_{vw\in D} \left(\sum_{\ell\neq i} \left(\frac{|N_\ell(v)|}{k-2}\right)^{k-2} + \sum_{\ell\neq j} \left(\frac{|N_\ell(w)|}{k-2}\right)^{k-2} \right) \leq 2|D|(k-1)\left(\frac{zn}{k-2}\right)^{k-2}.$$

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

Thus, recalling that $\delta := |D|/\binom{n}{2}$, we obtain

$$h_b \le 2\delta(k-1)\binom{n}{2}\left(\frac{zn}{k-2}\right)^{k-2}. \tag{3·7}$$

Using (3·5), (3·6), and (3·7) we obtain

$$I(R,n) \le \sum_i I(R,n_i) + \prod_\ell n_\ell \left(1 - \frac{\delta\binom{n}{2}}{\binom{n}{2} - \sum_i \binom{n_i}{2}}\right) + 2\delta(k-1)\binom{n}{2}\left(\frac{zn}{k-2}\right)^{k-2}. \tag{3·8}$$

Our final task is to upper bound the RHS.

As in Section 3.7, we see that $\delta \in I \stackrel{\text{def}}{=} \left[0, 1 - \sum_i \binom{n_i}{2}/\binom{n}{2}\right]$. Viewing (3·8) as a linear function of δ, it again suffices to check the endpoints of I.

3·8. *The extremal case*

CLAIM 3·7. *If $\delta = 0$, then* $\text{ind}\,(R) \le a$.

Proof. If $\delta = 0$, then (3·8) implies that

$$I(R,n) \le \sum_{i=1}^{k} I(R,n_i) + \prod_{i=1}^{k} n_i.$$

This is the same equation as (2·17), and we have all the same assumptions. The same argument as in Section 2.7 derives a contradiction.

3·9. *The absurd case*

Now, we consider the other endpoint of I.

CLAIM 3·8. *If $\delta = 1 - \sum_i \binom{n_i}{2}/\binom{n}{2}$, then* $\text{ind}\,(R) \le a$.

Proof. If $\delta = 1 - \sum_i \binom{n_i}{2}/\binom{n}{2}$, then (3·8) implies that

$$I(R,n) \le \sum_{i=1}^{k} I(R,n_i) + 2(k-1)\sum_{i\ne j} n_i n_j \left(\frac{z}{k-2}\right)^{k-2} n^{k-2}. \tag{3·9}$$

This is similar to (2·23) with an extra factor of approximately k^2 in the second term. We can bound the first sum using the same techniques as in Section 2.7, giving (2·23):

$$\frac{1}{\binom{n}{k}}\sum_{i=1}^{k} I(R,n_i) \le \ell\frac{\binom{c_0}{k}}{\binom{n}{k}} + (\text{ind}\,(R)+\varepsilon)\sum_{i=\ell+1}^{k} p_i^k < 0.74\,\text{ind}\,(R). \tag{3·10}$$

We now bound the second term. Dividing by $\binom{n}{k}$, we reorganize

$$\frac{2}{\binom{n}{k}}(k-1)\left(\sum_{i\ne j} n_i n_j\right)\left(\frac{z}{k-2}\right)^{k-2} n^{k-2} = \frac{2(k-1)(k^k-k)n^k}{(k-2)^{k-2}(n)_k}\left(\sum_{i\ne j} p_i p_j\right) z^{k-2} a.$$

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

We first relax $(k-1)(k^k-k) < k^{k+1}$. Observe that $k^{k-2}/(k-2)^{k-2} \le e^2$. Thus this first quotient is at most e^2k^3. For $n > M$, we have $n^k/(n)_k < 1+\varepsilon \le 1.01$. Finally, $\sum_{i\neq j} p_ip_j = (1-\sum p_i^2)/2 \le (1-1/k)/2 \le 1/2$ as $\sum p_i^2$ is minimised when $p_i = 1/k$ for all i. Thus

$$\frac{2}{\binom{n}{k}}(k-1)\left(\sum_{i\neq j} n_in_j\right)\left(\frac{z}{k-2}\right)^{k-2} n^{k-2} \le 2e^2k^3\cdot 1.01\cdot\frac{1}{2}\cdot z^{k-2}a = 1.01e^2k^3z^{k-2}a.$$

By Lemma 3·5, we know that $z^{k-2} \le (1-\eta/8)^{k-2}$, and we further see that

$$\left(1-\frac{\eta}{8}\right)^{k-2} \le \exp\left(-\frac{(k-2)\eta}{8}\right) \le \exp\left(-\frac{C}{8}\left(1-\frac{1}{k-1}\right)\log k\right) < k^{-0.9C/8}.$$

We again used $k \ge 11$ here. Thus for $C > 24/0.9 \approx 26.67$ and $k \ge 11$, we have $1.01e^2k^3z^{k-2}a < 0.25a$. Recalling that $a <$ ind (R), plugging this and (3·10) into (3·9) gives

$$\text{ind}\,(R,n) \le 0.74\,\text{ind}\,(R) + 0.25\,a < 0.99\,\text{ind}\,(R).$$

This contradiction completes the proof of the claim and the theorem.

4. *Disconnected rainbow graphs*

In this section, we show that rainbow graphs with multiple connected components are not fractalisers.

Let $R=(V,E)$ be a rainbow graph with k vertices and $\ell > 1$ connected components. Let $R = R_1 \cup \cdots \cup R_\ell$ be the connected components of size $c_1, \ldots, c_\ell$ respectively. Assume also $c_i \ge 2$ for all i (no isolated vertices). We will show that R is not a fractaliser.

We begin by upper bounding the number of copies $I(R, G_n)$ for $G_n \in \mathcal{G}_R(n)$ an iterated balanced blow-up. Then for any $i \in [\ell]$, by the same argument as for computing the inducibility of the iterated balanced blow-up (see e.g. [**14**]),

$$I(R_i, G_n) = \left(\frac{n}{k}\right)^{c_i} + k\left(\frac{n}{k^2}\right)^{c_i} + k^2\left(\frac{n}{k^3}\right)^{c_i} + \cdots = (1+o(1))\frac{n^{c_i}}{k^{c_i}-k}.$$

Any $S \subset V(G_n)$ with $G_n[S] \cong R$ has a unique partition $S = S_1 \cup \cdots \cup S_\ell$ where $G_n[S_i] \cong R_i$. Thus we can upper bound

$$I(R, G_n) \le \prod_{i=1}^{\ell} I(R_i, G_n) = (1+o(1))n^k \prod_{i=1}^{\ell}\frac{1}{k^{c_i}-k}. \tag{4·1}$$

However, consider instead the family of graphs $\mathcal{H}(n)$ consisting of separate iterated balanced blow-ups of each part. Formally, $H \in \mathcal{H}(n)$ if $|V(H)| = n$ and we have a partition $V(H) = V_1 \cup \cdots \cup V_\ell$ with the following properties:

(1) for all $i \in [\ell]$, $\left||V_i| - \frac{c_i}{k}n\right| \le 1$;
(2) for all $i \in [\ell]$, the induced subgraph $G[V_i] \in \mathcal{G}_{R_i}(|V_i|)$;
(3) for all $v \in V_i, w \in V_j$ with $i \neq j$, we have $vw \notin E(H)$.

In $\mathcal{H}(n)$, there are no edges between any copy of R_i and any copy of R_j for distinct i, j. Since R is rainbow, copies of each component R_i exist only in V_i. Then for $H_n \in \mathcal{H}(n)$, we

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

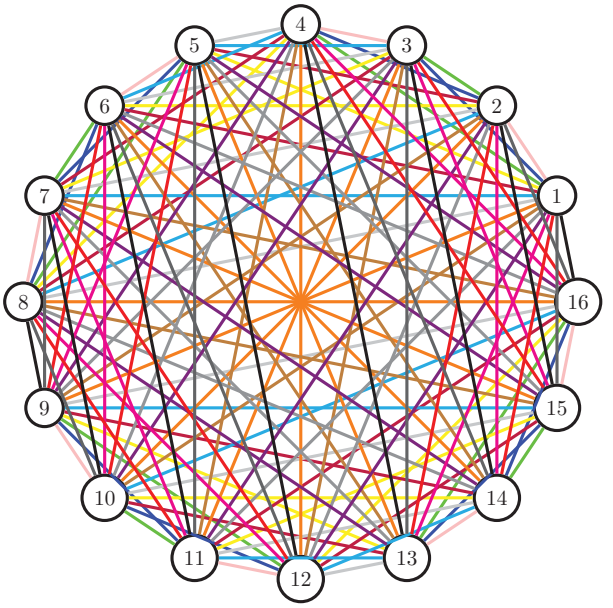


Fig. 1. A 15-edge-colouring of K_{16}.

have that $I(R, H_n) = \prod_i I(R_i, H_n[V_i])$, so

$$I(R, H_n) = \prod_{i=1}^{\ell} (1 + o(1)) \frac{c_i!}{c_i^{c_i} - c_i} \binom{\frac{c_i}{k} n}{c_i} = (1 + o(1)) n^k \prod_{i=1}^{\ell} \frac{1}{k^{c_i} - k(\frac{k}{c_i})^{c_i - 1}}. \tag{4·2}$$

Comparing (4·1) with (4·2), we subtract larger numbers in the denominator of (4·2), so the family of graphs $\mathcal{H}(n)$ induces asymptotically more copies than the family $\mathcal{G}_R(n)$. Thus R is not a fractaliser. Since R was generic, disconnected rainbow graphs without isolated vertices are not fractalisers.

5. *Concluding remark*

We now address the problem of maximising the number of rainbow cliques on k vertices (without fixing a specific rainbow colouring). In other words, for $\mathcal{G}_n$ the set of $\binom{k}{2}$-edge-coloured K_n, let

$$\gamma_k := \lim_{n \to \infty} \frac{\max_{G \in \mathcal{G}_n} |\{S \subset V(G) : |S| = k \text{ and } G[S] \text{ is rainbow}\}|}{\binom{n}{k}}. \tag{5·1}$$

We note that $\gamma_k \geq k!/(k^k - k)$.

The result of [**2**] shows that $\gamma_3 > 3!/(3^3 - 3) = 1/4$. We show that $\gamma_6 > 6!/(6^6 - 6)$. Notice that a rainbow K_6 has $\binom{6}{2} = 15$ colours. Consider the 15-edge-colouring of K_{16} displayed in Figure 1 and note that each colour class is a perfect matching. Letting the set of colours be $\{1, 2, \ldots, 15\}$, the adjacency matrix for this graph is shown below, where the i,

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

jth entry is the colour assigned to the edge ij.

$$\begin{bmatrix}
0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\
1 & 0 & 3 & 2 & 5 & 4 & 7 & 6 & 9 & 8 & 11 & 10 & 13 & 12 & 15 & 14 \\
2 & 3 & 0 & 1 & 6 & 7 & 4 & 5 & 10 & 11 & 8 & 9 & 14 & 15 & 12 & 13 \\
3 & 2 & 1 & 0 & 7 & 6 & 5 & 4 & 11 & 10 & 9 & 8 & 15 & 14 & 13 & 12 \\
4 & 5 & 6 & 7 & 0 & 1 & 2 & 3 & 12 & 13 & 14 & 15 & 8 & 9 & 10 & 11 \\
5 & 4 & 7 & 6 & 1 & 0 & 3 & 2 & 13 & 12 & 15 & 14 & 9 & 8 & 11 & 10 \\
6 & 7 & 4 & 5 & 2 & 3 & 0 & 1 & 14 & 15 & 12 & 13 & 10 & 11 & 8 & 9 \\
7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 & 15 & 14 & 13 & 12 & 11 & 10 & 9 & 8 \\
8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
9 & 8 & 11 & 10 & 13 & 12 & 15 & 14 & 1 & 0 & 3 & 2 & 5 & 4 & 7 & 6 \\
10 & 11 & 8 & 9 & 14 & 15 & 12 & 13 & 2 & 3 & 0 & 1 & 6 & 7 & 4 & 5 \\
11 & 10 & 9 & 8 & 15 & 14 & 13 & 12 & 3 & 2 & 1 & 0 & 7 & 6 & 5 & 4 \\
12 & 13 & 14 & 15 & 8 & 9 & 10 & 11 & 4 & 5 & 6 & 7 & 0 & 1 & 2 & 3 \\
13 & 12 & 15 & 14 & 9 & 8 & 11 & 10 & 5 & 4 & 7 & 6 & 1 & 0 & 3 & 2 \\
14 & 15 & 12 & 13 & 10 & 11 & 8 & 9 & 6 & 7 & 4 & 5 & 2 & 3 & 0 & 1 \\
15 & 14 & 13 & 12 & 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0
\end{bmatrix}$$

Of the 8008 subgraphs on 6 vertices, 448 of them are rainbow. Thus, the density of a rainbow K_6 in this graph is $448/8008 = 8/143 \approx 0.056$. Considering the iterated balanced blow-up of this construction, we see that

$$\begin{aligned}
\gamma_6 &\geq \lim_{n\to\infty} \frac{448\left(\frac{n}{16}\right)^6 + 16\left(448\left(\frac{n}{16^2}\right)^6 + 16\left(448\left(\frac{n}{16^3}\right)^6 + 16\,(\cdots)\right)\right)}{\binom{n}{6}} \\
&= 448 \cdot 720 \cdot \left(\frac{1}{16^6} + \frac{1}{16^{11}} + \frac{1}{16^{16}} + \cdots\right) \\
&= 322560 \cdot \frac{1}{16^6} \cdot \sum_{i=0}^{\infty} \left(\frac{1}{16^5}\right)^i \\
&= \frac{1344}{69905} \approx 0.0192.
\end{aligned}$$

This is larger than $6!/(6^6 - 6) \approx 0.0154$. It remains open to determine whether $\gamma_k = k!/(k^k - k)$ for $k = 4, 5$ and $k > 6$.

A. Appendix

Proof of (2·3). Let $q, q' \geq 0$ and $t, t' > 0$. Recall that $p(q, t)$ is the maximum of $\prod_i q_i$ where $q_1 + \cdots + q_t = q$ and each $q_i \geq 0$ is an integer. Let $q_1, \ldots, q_t$ integers such that $p(q, t) = \prod_{i=1}^{t} q_i$ and $q'_1, \ldots, q'_{t'}$ integers such that $p(q', t') = \prod_{i=1}^{t'} q'_i$. Then,

$$q_1 + \cdots + q_t + q'_1 + \cdots + q'_{t'} = q + q'.$$

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

Thus, the fact that $p(q+q', t+t')$ is a maximum gives that

$$p(q+q', t+t') \geq \prod_{i=1}^{t} q_i \prod_{i=1}^{t'} q'_i = p(q,t)p(q',t')$$

as desired.

Proof of (2·6). We will show that

$$\frac{1}{k^{k-1}-1} > 1.4\left(\frac{0.6}{k-2}\right)^{k-2}$$

for all $k \geq 11$. This is true for $k = 11$. By (2·7) and the fact that $k \geq 11$,

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

We will prove that

$$\frac{1}{e(k-1)^{k-1}} > 1.4\left(\frac{0.6}{k-2}\right)^{k-2}$$

for $k \geq 12$ by induction on k. For $k = 12$, plugging in certifies that this is true. By the inductive hypothesis, assume that

$$\frac{1}{e(k-2)^{k-2}} > 1.4\left(\frac{0.6}{k-3}\right)^{k-3}. \tag{A·1}$$

We see that

$$\frac{(k-2)^{k-2}}{(k-1)^{k-1}} > \frac{0.6(k-3)^{k-3}}{(k-2)^{k-2}}, \text{ or equivalently } f(k) := \frac{(k-2)^{2k-4}}{(k-1)^{k-1}(k-3)^{k-3}} > 0.6 \tag{A·2}$$

since

$$f(11) \approx 0.89 > 0.6$$

and

$$\frac{d}{dk}f(k) = -\frac{(k-2)^{2k-4}}{(k-1)^{k-1}(k-3)^{k-3}} \cdot \ln\left(\frac{(k-1)(k-3)}{(k-2)^2}\right) \geq 0$$

since $(k-2)^2 > (k-1)(k-3)$. Then, by (A·1) and (A·2), we see that

$$\begin{aligned}\frac{1}{e(k-1)^{k-1}} &= \frac{1}{e(k-2)^{k-2}} \cdot \frac{(k-2)^{k-2}}{(k-1)^{k-1}} \\ &> 1.4\left(\frac{0.6}{k-3}\right)^{k-3} \cdot \frac{0.6(k-3)^{k-3}}{(k-2)^{k-2}} = 1.4\left(\frac{0.6}{k-2}\right)^{k-2}.\end{aligned}$$

Proof of (2·8). We will show that

$$\frac{1}{1+\varepsilon} \cdot \frac{(k-1)^{k-1}}{k^k - k} - \frac{1}{3k} \geq z^{k-1}$$

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

for all $k \geq 11$. Recalling that $\varepsilon < \gamma/100$ and $\gamma \leq 1$, we obtain

$$\frac{1}{1+\varepsilon} \cdot \frac{(k-1)^{k-1}}{k^k - k} \geq \frac{100}{101} \cdot \frac{(k-1)^{k-1}}{k^k - k}.$$

Plugging in $k = 11$, we see that

$$\frac{100}{101} \cdot \frac{(k-1)^{k-1}}{k^k - k} - \frac{1}{3k} \geq z^{k-1}.$$

By (2·7) and the fact that $k \geq 11$,

$$\frac{100}{101} \cdot \frac{(k-1)^{k-1}}{k^k - k} \geq \frac{100}{101} \cdot \frac{1}{ek}.$$

So, it suffices to show that

$$\frac{100}{101} \cdot \frac{1}{ek} - \frac{1}{3k} \geq z^{k-1}$$

for all $k \geq 12$. We do so by induction on k. For $k = 12$, it can be verified directly. For the induction step, assume that $k \geq 13$ and

$$\frac{100}{101} \cdot \frac{1}{e(k-1)} - \frac{1}{3(k-1)} \geq z^{k-2}.$$

Using Lemma 2·4, we have $(k-1)/k \geq 12/13 > 0.5 > z$ and this yields

$$\frac{100}{101} \cdot \frac{1}{ek} - \frac{1}{3k} = \left(\frac{100}{101} \cdot \frac{1}{e(k-1)} - \frac{1}{3(k-1)}\right) \frac{k-1}{k} \geq z^{k-2} \cdot z = z^{k-1},$$

completing the proof.

Acknowledgements. The first and third author were partially supported by NSF Award DMS-1952767. The second author was partially supported by NSF Award ECCS-2217023 via the IDEAL institute. The third author was additionally partially supported by NSF Award DMS-2153576 and by a Simons Fellowship. The authors would like to thank anonymous referees for valuable comments.

REFERENCES

[1] J. Balogh, P. Hu, B. Lidický and 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 and M. Young. Rainbow triangles in three-coloured graphs. *J. Combin. Theory Ser. B* **126** (2017), 83–113.

[3] D. Conlon, J. Fox and B. Sudakov. Hypergraph Ramsey numbers. *J. Amer. Math. Soc.* **23** (2009), 247–266.

[4] T. Chao and H. Yu. Kruskal–Katona-type problems via the entropy method. *J. Combin. Theory Ser. B* **169** (2024), 480–506.

[5] P. Erdős and A. Hajnal. On Ramsey like theorems. Problems and results. Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst., Oxford, 1972). Inst. Math. Appl., Southend (1972), 123–140.

[6] J. Fox, H. Huang and C. Lee. Packing problems. Methods and Challenges in Extremal and Probabilistic Combinatorics workshop, Banff International Research Station (2015).

[7] J. Fox, L. Sauermann and F. Wei. On the inducibility problem for random Cayley graphs of abelian groups with a few deleted vertices. *Random Structures Algorithms* **59** (2021), 554–615.

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

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press

[9] H. HUANG. Personal communication (2019).
[10] D. KRÁL, S. NORIN and J. VOLEC. A bound on the inducibility of cycles. *J. Combin. Theory Ser. A* **161** (2019), 359–363.
[11] B. LIDICKÝ, C. MATTES and F. PFENDER. C_5 is almost a fractaliser. *J. Graph Theory* **104** (2023), 220–244.
[12] D. MUBAYI and A. RAZBOROV. Polynomial to exponential transition in Ramsey theory. *Proc. London Math. Soc*. **122**(1) (2021), 69–92.
[13] V. NIKIFOROV. The number of cliques in graphs of given order and size. *Trans. Amer. Math. Soc.* **363**(3) (2011), 1599–1618.
[14] N. PIPPENGER and M. GOLUMBIC. The inducibility of graphs. *J. Combin. Theory Ser. B* **19**(3) (1975), 189–203.
[15] A. A. RAZBOROV. On the minimal density of triangles in graphs. *Combin. Probab. Comput.* **17**(4) (2008), 603–618.
[16] C. REIHER. The clique density theorem. *Ann. of Math.* **184**(3) (2016), 683–707.
[17] R. YUSTER. On the exact maximum induced density of almost all graphs and their inducibility. *J. Combin. Theory Ser. B* **136** (2019), 81–109.

https://doi.org/10.1017/S0305004125101692 Published online by Cambridge University Press