

Inducibility and Subgraph Density Problems in Graphs

BY

Emily Cairncross
B.A. Oberlin College, 2021

DISSERTATION

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois Chicago, 2026

Chicago, Illinois

Defense Committee:

Dhruv Mubayi, Chair, Advisor
Vishesh Jain
Caroline Terry
György Turán
Marcus Michelen, Northwestern University

Accessibility Statement

An accessible EPUB version of this document is available at <https://ecairncr.github.io/> or can be obtained by contacting Emily Cairncross at egcairncross@gmail.com.

Dedication

To my parents, who inspire me to always be inquisitive, and to my partner, Connor O'Loughlin, who has been with me every step of the way.

Acknowledgments

I would first like to thank Dhruv Mubayi, who has been the best advisor I could have asked for. I learned so much about extremal combinatorics from him, and I appreciate his consistently high expectations of me. Thanks to Dhruv as well for travel support to attend many different conferences and the countless hours spent reading my paper and thesis drafts. I would also like to thank Clayton Mizgerd for being such a great collaborator. We filled in gaps in each others expertise, and were able to prove the results in Chapters 2 and 3 together.

I would also like to thank Vishesh Jain, Caroline Terry, György Turán, and Marcus Michelen for agreeing to be on my thesis committee and supporting such an active and positive MCS community during my time at UIC. Similarly, I would like to thank Will Perkins and Lev Reyzin for their mentorship and instruction in the first half of my program. Before I came to UIC, Elizabeth Wilmer and Susan Colley encouraged me throughout my time at Oberlin College, and I wouldn't be here without them.

Thank you to all of my peers at UIC as well as Emily Heath, Abhishek Dhawan, Nicholas Sieger, Rachel Snyder, Van Magnan, Cory Palmer, Christian Reiher, and many more for fruitful research discussions and support.

Finally, thank you to my family and friends for supporting me along the way. And thank you to Kapéj Coffee for caffeinating me through the drafting of this thesis.

Contribution of Authors

Chapters 4, 6, and 7 are joint work with Dhruv Mubayi and include results published in the *SIAM Journal on Discrete Mathematics* [8] in addition to further results and discussion. Chapter 2 comprises results published in the *Mathematical Proceedings of the Cambridge Philosophical Society* jointly with Clayton Mizgerd and Dhruv Mubayi [7]. Chapter 3 includes results proved jointly with Clayton Mizgerd, while Chapter 5 includes joint work with Dhruv Mubayi.

Contents

Accessibility Statement	ii
Dedication	iii
Acknowledgments	iv
Contribution of Authors	v
List of Figures	viii
List of Tables	ix
Summary	x
Chapter 1. Introduction	1
1. Inducibility	1
2. Subgraph Density	5
Chapter 2. Inducibility of rainbow graphs	10
1. Background and results	10
2. Rainbow cliques	12
3. Connected rainbow graphs	24
4. Disconnected rainbow graphs	37
5. Rainbow cliques without a fixed coloring	38
Chapter 3. Induced oriented cycles for low edge density	41
1. Background and results	41
2. Proof of Theorem 3.1	41
Chapter 4. The feasible region of edge-colored complete graphs	47
1. Background and results	47
2. Proof of Theorem 4.1	50
3. The feasible region	54
Chapter 5. Subgraph density of edge-weighted graphs	61
1. Background and results	61
2. Proof of Theorem 5.1	63
Chapter 6. Maximum star densities	66
1. Background and results	66
2. Maximum star densities: correcting an error	67
3. Maximum star density in a graph and its complement	74
Chapter 7. Subgraph density of ordered graphs	76

1. Background and results	76
2. Proof of Theorem 7.2	80
3. Ordered graphs on three vertices	84
Appendix A. Key inequalities for the inducibility of rainbow graphs	88
Appendix B. Supporting lemmas for maximum star densities	91
Appendix C. Copyright Statements	93
Bibliography	94
Vita	96

List of Figures

1.1 $K'_{3,4}$.	6
2.1 A 15-edge-coloring of K_{16} .	39
2.2 Adjacency matrix of the graph shown in Figure 2.1.	40
3.1 A_j when $b = v_j$.	44
3.2 A_j when $bv_j \in E(G)$.	44
3.3 A_j when $v_jb \in E(G)$.	45
4.1 $\Omega_3(K'_{2,2})$ (graphed at geogebra.org).	50
4.2 Construction maximizing $K'_{s,t}$ -density when $\sqrt{x_b} + \sqrt{x_g} \leq 1$.	52
4.3 Construction maximizing $K'_{s,s}$ -density when $\sqrt{x_b} + \sqrt{x_g} > 1$.	53
5.1 Weighted and un-weighted constructions.	61
6.1 A counterexample to [37, Claim 3.8] (shaded squares have β value 1).	74
7.1 $S_L(n, m)$.	77
7.2 $S_L(k)$ and $S_R(k)$.	77
7.3 Extremal constructions for F_{13} .	85

List of Tables

- | | | |
|---|--|----|
| 1 | Formulas for i_{ord} and I_{ord} for all ordered graphs on 3 vertices. | 80 |
|---|--|----|

Summary

In this thesis, we consider problems in extremal combinatorics concerning the number of copies of a fixed graph in another larger graph. Many of these problems can be phrased in terms of *inducibility* informally defined as follows: Given a graph F , the inducibility of F is the maximum proportion of induced copies of F in a large graph G . We also consider how to maximize or minimize the number of copies of F in G when G has a fixed edge-density. We address these problems in a variety of settings, including edge-colored graphs, edge-weighted graphs, oriented graphs, and ordered graphs. In some of these settings, we consider the number of (not necessarily induced) subgraphs of G that are isomorphic to F , while in others we only consider the induced subgraphs of G .

Our first results concern the inducibility of rainbow graphs, i.e. graphs where every edge is assigned a distinct color. This is inspired by work in [30] on the inducibility of oriented rainbow graphs. We remove the orientation of edges and 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 $k!/(k^k - k)$. We show that the same result holds if $k \geq 11$, and R is a clique. This answers a question posed by Huang [18], that is a generalization of an old problem of Erdős and Sós. We show that the connectedness assumption on R is necessary and also consider the problem when our rainbow coloring of R is not fixed.

Our next result is in the setting of oriented graphs. The fundamental conjecture in the area of inducibility, due to Pippenger and Golumbic [34], concerns the inducibility of cycles in simple graphs. Inspired by the work in [20], we prove an upper bound on the number of oriented cycles in an oriented graph that supports the Pippenger–Golumbic conjecture for graphs with lower edge density.

The remaining problems discussed in this thesis consider the same basic question of minimizing or maximizing the number of copies of a smaller graph in a larger graph, but where the edge density of the larger graph is fixed. We refer to such problems as *subgraph density* problems and consider them in the following settings:

- (1) **Edge-colored complete graphs.** For $s \geq 2$, we define a particular 3-edge-colored complete graph F_s on $2s$ vertices with colors blue, green, and red, and determine, for each (x_b, x_g) with $x_b + x_g \leq 1$ and $x_b, x_g \geq 0$, the maximum density of F in a large graph whose blue, green, and red edge sets have densities x_b, x_g and $1 - x_b - x_g$, respectively. We extend the concept of the feasible region to edge-colored complete graphs as well, with some help from the theory of graphons, and compute it entirely for F_s . These are the first nontrivial examples of colored graphs for which such complete results have been proved and are an extension of work in [26].
- (2) **Edge-weighted graphs.** Given a fixed number of edges m in a graph G and an integer $k \geq 2$, the Kruskal–Katona theorem [19, 21] implies that the maximum number of copies of the complete graph K_k in G is achieved when G is as close to a clique as possible. We address a generalization when the edges are assigned

weights. We show that in the weighted setting, where each edge-weight lies in $[0, 1]$, the weighted number of copies of K_k can exceed the bound in the unweighted setting, but only by a lower order term.

- (3) **Stars.** We consider simple graphs and correct an error in a result of Reiher and Wagner [37]. We prove that the number of k -edge stars in a graph with density $\gamma \in [0, 1]$ is asymptotically maximized by a clique and isolated vertices or its complement. Inspired by this problem, we also consider how to maximize the sum of the number of k -edge stars in a graph with density $\gamma \in [0, 1]$ and the number of copies in its complement. We show that the same constructions maximize this quantity for $k = 2$ and 3 as in the original case.
- (4) **Ordered graphs.** We consider graphs with a total order on their vertices. Among all ordered n -vertex graphs with m edges, we determine the maximum and minimum number of copies of a k -edge star whose nonleaf vertex is minimum among all vertices of the star. The constructions achieving the minimum and maximum are similar to the construction in [37] that achieves the maximum star density for simple graphs for certain edge densities. Further, we determine the maximum and minimum number of copies for almost all ordered graphs on 3 vertices.

CHAPTER 1

Introduction

We consider inducibility and subgraph density problems in many settings. In the literature, some of these settings have overlapping notation, so we will need to use nonstandard notation for some problems.

1. Inducibility

1.1. The inducibility of simple graphs. Fix a graph F on k vertices and another graph G on $n > k$ vertices. For $S \subset V$, let $G[S]$ be the *induced subgraph* of G on vertex set S . Write $N_{\text{ind}}(F, G)$ for the number of k -subsets $S \subset V(G)$ such that $G[S] \cong F$ and let

$$\varrho_{\text{ind}}(F, G) := \frac{N_{\text{ind}}(F, G)}{\binom{n}{k}}.$$

Many foundational questions in extremal graph theory deal with estimating $\varrho_{\text{ind}}(F, G)$ for various choices of F and G .

Let $N_{\text{ind}}(F, n)$ be the maximum of $N_{\text{ind}}(F, G)$ over all n vertex graphs G . A standard averaging argument implies that

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

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

$$\text{ind}(F) := \lim_{n \rightarrow \infty} \text{ind}(F, n).$$

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 $[k] := \{1, \dots, k\}$. For $n < k$, the family

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

- (1) For all $i, j \in [k]$, $||V_i| - |V_j|| \leq 1$.
- (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 iterated balanced blow-up achieves the inducibility of F , in which case we call F a *fractalizer*.

The famous Pippenger–Golumbic conjecture [34] is as follows.

CONJECTURE 1.1 (Pippenger–Golumbic [34]). *For $k \geq 5$, the cycle C_k is a fractalizer and satisfies*

$$\text{ind}(C_k) = \frac{k!}{k^k - k}.$$

Conjecture 1.1 has been resolved for $k = 5$ by Balogh, Hu, Lidický, and Pfender [2] (see also [22]), but remains open for all $k \geq 6$. Král, Norin, and Volec [20] showed that $N_{\text{ind}}(C_k, n) \leq 2n^k/k^k$. More generally, Fox, Huang, and Lee [13] and Yuster [38] independently proved that random graphs are fractalizers asymptotically almost surely. Fox, Sauermann, and Wei [14] further proved that random Cayley graphs of abelian groups with small number of vertices removed are almost surely fractalizers. In contrast, a recent result of Liu, Ma, and Zhu [24], combined with prior work in [34, 6, 5, 26, 39], shows that Turán graphs are not fractalizers. Their inducibility is achieved by larger Turán graphs, not the iterated balanced blow-up.

1.2. The inducibility of rainbow graphs. We now consider these notions on colored and directed structures. A *tournament* is an orientation of a complete graph. An edge-coloring of a graph or tournament G is a function $\chi : E(G) \rightarrow T$ where T is a set of colors; we say that G is T -colored. A colored graph or tournament G is *rainbow* if χ is injective. Two colored graphs (or tournaments) G and H are *isomorphic*, written $G \cong H$, if there exists a bijection $\varphi : V(G) \rightarrow V(H)$ such that the colors (and orientations) of all edges are preserved under φ . If F is a colored tournament or colored complete graph, then $\text{ind}(F)$ is defined

identically as in the graph case, but with these altered definitions of graph isomorphism; naturally, the underlying graph G should have the colors or orientations corresponding to F . If F is an arbitrary colored graph, then we can color all missing edges with a single new color and view F as a colored complete graph. Consequently, we can define *fractalizer* for all these structures.

There are very few results on the inducibility of colored, oriented, or directed structures. The first exact result which involved an iterated construction was due to Huang [17] who determined the inducibility of the directed star. Later, in order to solve an old conjecture of Erdős and Hajnal [11] in hypergraph Ramsey theory, Mubayi and Razborov [30] proved the following result for $k \geq 4$ (the case $k = 3$ was proven earlier by Conlon, Fox, and Sudakov [10]).

THEOREM 1.2 (Mubayi, Razborov [30]). *All rainbow tournaments R on $k \geq 4$ vertices are fractalizers.*

In this thesis, we consider the question addressed by Theorem 1.2 in the undirected setting. The first conjecture in this setting is due to Erdős and Sós from the 1970s (see [11, Equation (20)]), and implies, in particular, that a rainbow triangle is not a fractalizer. Their conjecture was proved by Balogh et. al. [3], who showed that a blow-up of a properly 3-edge-colored K_4 (instead of a rainbow K_3) achieves the inducibility of the rainbow triangle. See also [9] for similar computations, but in terms of the number of edges of each color instead of the number of vertices.

Huang [18] 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 Erdős and Sós and proved in [3] (that K_k is not a fractalizer for $k = 3$) fails to hold for larger k . Our first result for rainbow graphs addresses Huang's question and proves that rainbow K_k are fractalizers for $k \geq 11$.

THEOREM (2.2). *All rainbow cliques R on $k \geq 11$ vertices are fractalizers.*

Our proof of Theorem 2.2 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 color 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 2.2 requires a bound on the color degree of a vertex that was not needed in [30].

For large values of k , we prove the following more general result which shows that the analog of Theorem 2.2 holds for much sparser graphs. The proof requires several major new ideas.

THEOREM (2.3). *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 fractalizers.*

Note that the connectedness assumption is necessary for Theorem 2.3 to be true. In Chapter 2, we prove Theorems 2.2 and 2.3, and discuss the problem when a specific coloring is not specified.

1.3. The inducibility of oriented cycles. We now approach Conjecture 1.1 from a different direction by considering oriented cycles (with no colors). While our results are weaker here, this setting is also much closer to that of Conjecture 1.1 itself.

For an oriented graph G , we adopt the convention that, when an edge is written uv , the edge is oriented from u to v in G . We define an *oriented cycle on k vertices*, denoted \vec{C}_k , as the oriented graph with vertex set $\{v_1, v_2, v_3, v_4, \dots, v_k\}$ and edge set $\{v_1v_2, v_2v_3, \dots, v_kv_1\}$. Král, Norin, and Volec [20] proved that $N_{\text{ind}}(C_k, n) \leq 2n^k/k^k$. We augment their methods to apply to \vec{C}_k and obtain the following result.

THEOREM (3.1). *Let $k \geq 4$ be fixed. Let G be an oriented graph with n vertices and m edges. Then,*

$$N_{\text{ind}}(\vec{C}_k, G) \leq \frac{mn^{k-2}}{k^{k-1}}.$$

This implies that the bound in Conjecture 1.1 holds for oriented cycles whenever $m < (1 + o(1))(2/k) \cdot \binom{n}{2}$. In Chapter 3, we prove Theorem 3.1 and elaborate on its implications.

2. Subgraph Density

For the rest of our results, we consider subgraph density problems where we fix the number of edges in our larger graph G . We also consider subgraphs F of G that are not necessarily induced.

The *density* or *edge density* of a graph G with n vertices and m edges is $\varrho(G) := m/\binom{n}{2}$. For a graph F with $k \leq n$ vertices, let $N(F, G)$ be the number of subgraphs of G that are isomorphic to F . We are interested in the minimum and maximum values of $N(F, G)$ over graphs G with a given value of $\varrho(G)$ as n grows. We note that $N(F, G) \leq \binom{n}{k}/|\text{Aut}(F)|$, where $\text{Aut}(F)$ is the automorphism group of F and $\binom{n}{k}$ is the falling factorial $n(n-1) \cdots (n-k+1)$. Define the *(labeled) density of F in G* to be

$$\varrho(F, G) := \frac{N(F, G) \cdot |\text{Aut}(F)|}{\binom{n}{k}} \in [0, 1].$$

We note that the notation $t_{\text{inj}}(F, G)$ is often used for $\varrho(F, G)$, though it is more convenient to use $\varrho(F, G)$ in this thesis. For some settings, we add subscripts to these notations to make it clear what setting we are in.

The classical Kruskal–Katona theorem [19, 21] implies that the maximum density of K_s in a graph of density x is achieved asymptotically by graphs consisting of a clique and isolated vertices. The minimum density of K_s in a graph with density x was determined by Razborov [35] for $s = 3$, by Nikiforov [31] for $s = 4$, and by Reiher [36] for all $s \geq 4$; this is achieved by complete multipartite graphs. The exact solution for $s = 3$ was proven by H. Liu, Pikhurko, and Staden [23]. We contribute to this area of research by posing these problems in similar settings.

2.1. Edge-colored complete graphs. Our first problem is a colored generalization of the concepts introduced by Liu, Mubayi, and Reiher in [26]. In [26], the authors consider the region of possible asymptotic values of $\varrho_{\text{ind}}(F, G)$ for graphs G of fixed density. By coloring

edges in both F and G red, and coloring edges in their complements blue, it is obvious that counting induced subgraphs is the same as counting copies of a two-edge-colored clique in a (larger) two-edge-colored clique using the same two colors. This leads us to consider the maximum asymptotic density of q -edge-colored cliques in (larger) q -edge-colored cliques with given color densities.

For $2 \leq s \leq t$, let $K'_{s,t}$ be the 3-colored clique on vertex set $V = V_1 \sqcup V_2$ with $|V_1| = s$ and $|V_2| = t$ with coloring function f defined by

$$f(ij) := \begin{cases} \text{blue} & \text{if } i, j \in V_1, \\ \text{green} & \text{if } i, j \in V_2, \\ \text{red} & \text{otherwise,} \end{cases}$$

for all distinct $i, j \in [s+t]$ (see Figure 1.1).

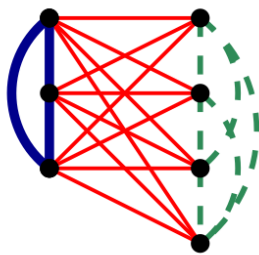


FIGURE 1.1. $K'_{3,4}$.

For any given blue and green densities, we find the maximum number of copies of $K'_{s,s}$ asymptotically and provide a construction that achieves it. We obtain partial results for $K'_{s,t}$ when $s \neq t$. This is the first nontrivial result in the edge-colored setting.

The *feasible region* of a simple graph F comprises not just the maximum subgraph density of F in a larger graph of edge density x , but all possible subgraph densities for each x (see [26]). In Chapter 4, we prove the aforementioned result on $K'_{s,t}$ and extend the theory of the feasible region to the colored setting.

2.2. Edge-weighted graphs. Instead of assigning colors to edges, we assign weights in $[0, 1]$. Let G' be an edge-weighted graph on vertex set V' and edge set $\binom{V'}{2}$ with weight function $w : \binom{V'}{2} \rightarrow [0, 1]$. Then we define

$$N_w(F, G') := \sum_{\substack{A \subset V', B \subset \binom{V'}{2} \\ (A, B) \cong F}} \prod_{e \in B} w(e)$$

the *total F -weight* of G' . Note that, if w maps instead to just $\{0, 1\}$, then $N_w(F, G')$ corresponds exactly to $N(F, G)$ when G is the simple graph on V' with all edges of weight 1 in G' . For a graph G , we often use m to denote the number of edges. For a weighted graph G' , we will use m to refer to the total weight of the edges, i.e. $m = \sum_{e \in \binom{V'}{2}} w(e)$.

The Kruskal–Katona theorem [19, 21] implies that the most copies of K_k in a graph G on m edges are achieved when G is also a clique. The following result shows that the K_k -weight cannot increase by more than $O(n^{k-2})$ when allowing edge-weights.

THEOREM (5.1). *Let $k \geq 3$. Let G' be a $[0, 1]$ -edge-weighted graph on $n \geq 3$ vertices. Set*

$$m := \left\lfloor \sum_{e \in E(G')} w(e) \right\rfloor.$$

Then there exists an n -vertex (un-weighted) graph G with m edges such that

$$N_w(K_k, G') - N(K_k, G) \leq \frac{2}{(k-4)!} \cdot n^{k-2}$$

when $k \geq 4$ and

$$N_w(K_3, G') - N(K_3, G) \leq 2n.$$

In Chapter 5, we prove Theorem 5.1 and compare the K_k -weight of both edge-weighted and non-edge-weighted constructions.

2.3. Maximum star densities. Let S_k denote the k -edge star. Ahlswede and Katona [1] determined the maximum number of S_2 's in a graph with density γ . Reiher and Wagner [37] claim to prove that the asymptotic maximum value of $N(S_k, G)$ when G has density γ is achieved by a clique and isolated vertices or its complement, called an *anticlique*.

Clearly $N(S_k, G) = \sum_{v \in V(G)} \binom{d(v)}{k}$ and $\varrho(S_k, G) = N(S_k, G) \cdot k! / (n)_{k+1}$. For $\gamma \in [0, 1]$, let

$$I(S_k, \gamma) := \sup \lim_{n \rightarrow \infty} \varrho(S_k, G_n)$$

over all sequences of graphs $(G_n)_{n=1}^{\infty}$ with $|V(G_n)| \rightarrow \infty$, $\varrho(G_n) \rightarrow \gamma$ and for which $\lim_{n \rightarrow \infty} \varrho(S_k, G_n)$ exists. For each $\gamma \in [0, 1]$, let $\eta = 1 - \sqrt{1 - \gamma}$. Then $\gamma^{(k+1)/2}$ and $\eta + (1 - \eta)\eta^k$ are the asymptotic S_k -densities in a clique with isolated vertices and the complement of a clique with isolated vertices, both with density γ . Consequently,

$$I(S_k, \gamma) \geq \max\{\gamma^{(k+1)/2}, \eta + (1 - \eta)\eta^k\}.$$

Reiher and Wagner [37] proved matching upper bounds on $I(S_k, \gamma)$. Their results are stated (and proved) using the language of graphons, which are limit object of graphs (see Lovász [27] for background on graphons).

In Chapter 6, we correct an error in their proof. We also consider the problem of maximizing the sum of the number of copies of S_k in a graph G and its complement \overline{G} . We prove that this sum is also maximized by the clique or the anticlique asymptotically for $k = 2$ and $k = 3$. This proof makes use of a result on the induced density of triangles and their complement from [26] combining results from [15, 33].

2.4. Ordered graphs. We consider the subgraph density of stars in the setting of ordered graphs and prove the first nontrivial results on S_k whose vertices have a particular order. An *ordered graph* $G = (V, E)$ is a graph with a total order on V . We usually let $V = [n] := \{1, \dots, n\}$ with the natural ordering. When we refer to an edge ij in E , it is implied that $\{i, j\} \in E$ and $i < j$. Let F be an ordered graph on $[s]$. Let $N_{\text{ord}}(F, G)$ be the number of $\{v_1, \dots, v_s\} \subseteq [n]$ with $v_1 < v_2 < \dots < v_s$ such that $v_i v_j \in E(G[v_1, \dots, v_s])$ whenever $ij \in E(F)$. We consider $N_{\text{ord}}(F, G)$ in the case that F is an appropriate ordered S_k .

Let $S_L(k)$ be the ordered star on k edges where the node is the smallest vertex. Let $S_R(k)$ be the ordered star where the node is the greatest vertex. The number of copies of these stars

is maximized or minimized by an ordered analog of the anticlique. Recall that the anticlique also maximizes the number of stars in simple graphs of certain edge density as claimed in [37] and fully proven in Chapter 6. Let $S_L(n, m)$ be the anticlique on n vertices and m edges where vertices are ordered by descending degree, and let $S_R(n, m)$ be the anticlique on n vertices and m edges where the vertices are ordered by ascending degree. (See Chapter 7 for the more technical definition.) Then, we have the following main theorem.

THEOREM (7.2). *Let G be an ordered graph with vertex set $[n]$ and m edges. Then*

$$N_{\text{ord}}(S_L(k), S_R(n, m)) \leq N_{\text{ord}}(S_L(k), G) \leq N_{\text{ord}}(S_L(k), S_L(n, m)).$$

Theorem 7.2 is proven in Chapter 7. We also discuss some ordered stars with a node that is neither smallest nor largest among the other vertices, as well as all ordered graphs on three vertices. Addressing one of these graphs, we make use of Razborov's result on the minimum triangle density in graphs [35] (see also [28, 15, 29, 32, 4, 16, 12]).

CHAPTER 2

Inducibility of rainbow graphs

This chapter is based on joint work with Clayton Mizgerd and Dhruv Mubayi in [7], published in the Mathematical Proceedings of the Cambridge Philosophical Society.

1. Background and results

Let R be a graph on k vertices. We now formally define what it means for a graph R to be a fractalizer (our definition is slightly different than that in [22]). Note that, per discussion in Subsection 1.2 of Chapter 1, this definition applies whether or not R is edge-colored.

DEFINITION 2.1. A graph R is a *fractalizer* if

$$\text{ind}(R) = \lim_{n \rightarrow \infty} \max_{G \in \mathcal{G}_R(n)} \varrho_{\text{ind}}(R, G).$$

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

Let $G \in \mathcal{G}_R(n)$ be an iterated balanced blow-up of R with vertex partition $V(G) = V_1 \cup \dots \cup V_k$. The subgraph induced by every k -set comprising exactly one vertex in each V_i is isomorphic to R . Consequently, for every $G \in \mathcal{G}_R(n)$,

$$N_{\text{ind}}(R, G) \geq \sum_{i=1}^k N_{\text{ind}}(R, G[V_i]) + \prod_{i=1}^k |V_i|.$$

Together with a standard computation (see, e.g. [30]), this yields

$$\text{ind}(R) \geq \lim_{n \rightarrow \infty} \max_{G \in \mathcal{G}_R(n)} \varrho_{\text{ind}}(R, G) \geq \frac{k!}{k^k - k}. \quad (1)$$

In most cases we consider, the fact that R is a fractalizer will imply further that $\text{ind}(R) = k!/(k^k - k)$. In particular, Theorem 1.2 implies that $\text{ind}(R) = k!/(k^k - k)$ when R is a

rainbow tournament on $k \geq 4$ vertices. However, K_k is a fractalizer with $\text{ind}(K_k) = 1$, so this implication does not always hold.

We now state our main results in full.

THEOREM 2.2. *All rainbow cliques R on $k \geq 11$ vertices are fractalizers. In particular,*

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

We make the following observations regarding Theorem 2.2.

- Theorem 2.2 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.
- Similarly, if a graph G on k vertices is known to have inducibility $k!/(k^k - k)$, then the rainbow k -clique is a fractalizer as well by the following argument. Let R be a rainbow k -clique and let $e_1, e_2, \dots, e_m \in E(R)$ such that $(V(R), \{e_1, \dots, e_m\})$ is a rainbow copy of G . Let c_1, \dots, c_m be the colors assigned to e_1, \dots, e_m , respectively. Then any construction of an edge-colored graph inducing ℓ rainbow copies of R induces at least ℓ copies of G by deleting all edges except those colored by c_1, \dots, c_m and then ignoring the edge colors. It follows that $\text{ind}(R) \leq \text{ind}(G) = k!/(k^k - k)$, so R is a fractalizer. Thus, the result of [2] that C_5 is a fractalizer with $\text{ind}(C_5) = 5!/(5^5 - 5)$ implies that the rainbow 5-clique is a fractalizer, and the result of [13] that random graphs are almost surely fractalizers implies that the rainbow k -clique is a fractalizer for large k .
- We believe our proof of Theorem 2.2 has been optimized and requires $k \geq 11$. As [2] showed that rainbow 5-cliques are fractalizers, and [3] showed that rainbow 3-cliques are not fractalizers, it remains open to determine whether rainbow k_0 -cliques are fractalizers only for $k_0 \in \{4, 6, 7, 8, 9, 10\}$.

We now lift the requirement that R be a clique and obtain the following result.

THEOREM 2.3. *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 fractalizers and satisfy*

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

We make the following observations regarding Theorem 2.3.

- Theorem 2.3 implies Theorem 2.2 for large k , since R may be viewed as a rainbow k -clique with edges deleted. Let c_1, c_2, \dots, c_m be the colors assigned to the deleted edges. Any construction of a colored complete graph inducing ℓ rainbow copies of the rainbow k -clique yields at least ℓ rainbow copies of R by deleting edges colored c_1, \dots, c_m .
- The requirement that R is connected in the statement of Theorem 2.3 is necessary, as disconnected rainbow graphs without isolated vertices are not fractalizers (see Section 4).
- We are not able to show that our requirement on minimum degree is tight, and this remains open.

Theorem 2.2 is proven in Section 2 and Theorem 2.3 is proven in Section 3. We conclude the chapter with two shorter sections on related problems. Section 4 explains why R must be connected in Theorem 2.3, while Section 5 discusses the corresponding problem for the collection of all rainbow colorings as both Theorems 2.2 and 2.3 assume a specific rainbow coloring.

2. Rainbow cliques

We give the proof of Theorem 2.2 in the following subsections.

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

Set

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

Our goal is to prove that $\text{ind}(R) \leq a$. 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 $\text{ind}(R, n) \leq \text{ind}(R) + \varepsilon$ for all $n > c_0$. Choose $M \geq \lceil 2k!c_0/\varepsilon \rceil$ such that

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

and

$$a \left(\frac{n^{k-1}}{(k-1)!} - \binom{n-1}{k-1} \right) < \gamma \left(\binom{n-1}{k-1} - \binom{n-2}{k-2} \right) \quad (3)$$

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

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

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

DEFINITION 2.4. For $q \geq 0$ and $t > 0$, let $p(q, t)$ be the maximum of $\prod_i q_i$ where $q_1 + \dots + q_t = q$ and each $q_i \geq 0$ an integer.

The AM–GM inequality yields $p(q, t) \leq (q/t)^t$ and it is straightforward to see that

$$p(q, t)p(q', t') \leq p(q + q', t + t') \quad (4)$$

for all $q, q' \geq 0$ 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 $\varphi : R \rightarrow H$ such that $\varphi(i) = x$. 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 $j \neq j'$, but all edges between $N_j(x) \cap N_{j'}(x)$ and x have the same color $\{j, j'\}$. However, $N_i(x)$ 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 . Indeed, the partition is obtained based on the color of a vertex to x . This gives

$$d(x) = \sum_{i=1}^k d_i(x) \leq \sum_{i=1}^k \prod_{j \neq i} |N_i^j(x)| \leq \sum_{i=1}^k p(|N_i(x)|, k-1). \quad (5)$$

We partition $V(H)$ into $V_1 \cup \dots \cup V_k$, 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. Set $n_i = |V_i|$ for all $i \in [k]$.

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 2.5. $d(x) \geq an^{k-1}/(k-1)!$ 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

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

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

$$d(x) = d(H) \pm \binom{n-2}{k-2}. \quad (6)$$

This follows from a standard application of Zykov symmetrization. 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 $N_{\text{ind}}(R, H) = N_{\text{ind}}(R, n)$. Hence all degrees lie in an interval of length at most $\binom{n-2}{k-2}$ and (6) 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 (3) 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. \square

2.3. Maximum color 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 colors $\{i, j\} \in T$ and $d_{\{i,j\}}(x)$ is the number of edges in H incident with x in color $\{i, j\}$. We upper bound this value.

LEMMA 2.6. $\alpha \leq 0.4$.

PROOF. Let x, i, j achieve this maximum, so that $d_{\{i,j\}}(x) = \alpha n$. Then $|N_i^j(x)| \leq \alpha n$ and $N_i(x)$ has a partition $N_i^j(x) \cup \bigcup_{\ell \neq j} N_i^\ell(x)$ 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,

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

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

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

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

color must include ℓ . Hence

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

Altogether this yields

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

Suppose for contradiction that $\alpha > 0.4$. Since $k \geq 3$, $(1+\alpha)(1-\alpha)^{k-2}$ is a decreasing function of α for $\alpha \in (0.4, 1]$, and $d(x) \geq \alpha n^{k-1}/(k-1)!$ by Lemma 2.5. Therefore

$$\frac{1}{k^{k-1}-1} = \frac{\alpha}{(k-1)!} \leq \frac{d(x)}{n^{k-1}} \leq 1.4 \left(\frac{0.6}{k-2}\right)^{k-2}. \quad (7)$$

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

2.4. 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 2.7. $z \leq 0.5$.

PROOF. Let x be such that $z = |Z(x)|/n$. Let $a_i = |N_i(x)|/n$ and assume by relabeling that $a_1 \geq a_2 = z \geq a_3 \geq \dots \geq a_k$. Since $N_j(x) \cap N_{j'}(x) \cap N_{j''}(x) = \emptyset$ for any three distinct j, j', j'' we have $\sum a_i \leq 2$. Let $a_3 + \dots + a_k = s \leq 2 - (a_1 + z)$. Write $s = qz + r$ where $q \in \mathbb{Z}^{\geq 0}$ and $0 \leq r < z$. If $x \leq y$, 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 , we obtain

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

Consequently,

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

Since $a_1 \geq z$, taking the derivative shows that for any z , this expression is increasing with a_1 . Using Lemma 2.6, we note that $a_1 + z \leq 1 + \alpha \leq 1.4$ since

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

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

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

Using (5) and Lemma 2.5 yields

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

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

$$\frac{(k-1)^{k-1}}{k^{k-1} - 1} \geq \frac{(k-1)^{k-1}}{k^{k-1}} = \left(1 - \frac{1}{k}\right)^{k-1} > \frac{1}{e} \quad (8)$$

for $k > 1$, we obtain

$$\frac{1}{e} < (\min\{1.4 - z, 1\})^{k-1} + z^{k-1} + 0.6 \cdot z^{k-2}. \quad (9)$$

As $z \leq a_1$ and $z + a_1 < 1.4$, we have $z < 0.7$. Thus, the right side of (9) 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 $z < 0.5$. \square

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

LEMMA 2.8. $|V_i| \leq (1 - 1/3k)n$ for all $i \in [k]$.

PROOF. By contradiction, Assume without loss of generality that $|V_1| > (1 - 1/3k)n$. If $x \in V_1$, then $|N_1(x)| \geq |N_i(x)|$ for all $i > 1$ so $|N_2(x)| \leq |Z(x)| \leq zn$. Using (5) we have

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

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), we get

$$\left(\frac{(k-1)^{k-1}}{k^k - k} \right) < (1 + \varepsilon) \left(z^{k-1} + \frac{1}{3k} \right). \quad (10)$$

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

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 [30], 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 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 (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

$$N_{\text{ind}}(R, H) = h_m + h_g + h_b.$$

We will bound each of these three terms separately. First, note that

$$h_m = \sum_j N_{\text{ind}}(R, H[V_j]) \leq \sum_j N_{\text{ind}}(R, n_j). \quad (11)$$

Next we turn to h_g . Let \mathcal{D} denote the number of k -sets that intersect each V_i but are not counted by h_g . So a k -set counted by \mathcal{D} 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 $V_1 \cup \dots \cup V_k$. Then

$$h_g = \prod_i n_i - \mathcal{D} \quad (12)$$

and we need to bound \mathcal{D} from below.

Note that the color of some pair in every member of \mathcal{D} 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 $\chi_H(v_i v_j) \neq \chi_R(ij) = \{i, j\}$. Let $\delta_{ij} = |D_{ij}| / \binom{n}{2}$, $D = \cup_{ij} D_{ij}$ and $\delta = |D| / \binom{n}{2}$. Let us lower bound \mathcal{D} 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 $i < j$,

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

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

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

This with along with (12) gives

$$h_g \leq \prod_{\ell=1}^k n_\ell \left(1 - \frac{\delta \binom{n}{2}}{\sum_{1 \leq i < j \leq k} n_i n_j} \right) = \prod_{\ell=1}^k n_\ell \left(1 - \frac{\delta \binom{n}{2}}{\binom{n}{2} - \sum_i \binom{n_i}{2}} \right). \quad (13)$$

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 $i' \neq i$, so the colors 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 color 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 miscolored 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 miscolored. Also, note that $2 \leq |\mathcal{C}| \leq k - 1$ since f is not contained in one color class and we have assumed f is not transversal.

If exactly $|\mathcal{C}| - 1$ vertices in \mathcal{C} are miscolored, then every edge vw where $v \in \mathcal{C}$ is miscolored and $w \in f \setminus \mathcal{C}$ is in D . Since $|f \setminus \mathcal{C}| = k - |\mathcal{C}|$, this yields at least $(|\mathcal{C}| - 1)(k - |\mathcal{C}|) \geq k - 2$ pairs from D in f . On the other hand, if all $|\mathcal{C}|$ vertices in \mathcal{C} are miscolored, 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 $|\mathcal{C}| \leq k - 2$, this yields at least $|\mathcal{C}|(k - |\mathcal{C}| - 1) \geq k - 2$ 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 color class, say the color class corresponding to color 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

$$S \geq 2(k - 2)h_b. \tag{14}$$

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

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

The crucial observation is that

$$(v, w, f) \in J_i \quad \implies \quad \{v, w\} \in D_i \quad (i = 1, 2). \quad (15)$$

To bound S from above, we use (15) 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

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

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,

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

This gives

$$S = 2|J_1| + |J_2| \leq 4|D|p(zn, k - 2) \leq 4\delta \binom{n}{2} \left(\frac{z}{k-2}\right)^{k-2} n^{k-2}. \quad (16)$$

Finally, (14) and (16) give

$$h_b \leq \frac{S}{2(k-2)} \leq \frac{2\delta \binom{n}{2}}{k-2} \left(\frac{z}{k-2}\right)^{k-2} n^{k-2}. \quad (17)$$

Using (11), (13) and (17) we have that

$$N_{\text{ind}}(R, n) \leq \sum_i N_{\text{ind}}(R, n_i) + \prod_{\ell} n_{\ell} \left(1 - \frac{\delta \binom{n}{2}}{\binom{n}{2} - \sum_i \binom{n_i}{2}} \right) + \frac{2\delta \binom{n}{2}}{k-2} \left(\frac{z}{k-2} \right)^{k-2} n^{k-2}. \quad (18)$$

Our final task is to upper bound the right side of (18).

Since $\delta \binom{n}{2} \leq \sum_{i \neq j} n_i n_j = \binom{n}{2} - \sum_i \binom{n_i}{2}$, we have $\delta \in I = [0, 1 - \sum_i \binom{n_i}{2} / \binom{n}{2}]$. Viewing (18) as a linear function of δ , it suffices to check the endpoints of I .

2.7. The extremal case.

CLAIM 2.9. *If $\delta = 0$, then $\text{ind}(R) \leq a$.*

PROOF. If $\delta = 0$, then (18) implies that

$$N_{\text{ind}}(R, n) \leq \sum_{i=1}^k N_{\text{ind}}(R, n_i) + \prod_{i=1}^k n_i. \quad (19)$$

Let $p_i := n_i/n$. Using $\max_i p_i \leq 1 - 1/3k$ by Lemma 2.8, the convexity of x^k , and $k \geq 11$ we obtain

$$\sum_{i=1}^k p_i^k \leq \left(1 - \frac{1}{3k} \right)^k + \left(\frac{1}{3k} \right)^k \leq e^{-1/3} + 33^{-11} < 0.72. \quad (20)$$

We begin by bounding the summation in (19). By relabeling if necessary, let $n_1 \leq \dots \leq n_{\ell} \leq c_0 < n_{\ell+1} \leq \dots \leq n_k$ where $\ell \geq 0$. We have that

$$\sum_{i=1}^k N_{\text{ind}}(R, n_i) \leq \ell \binom{c_0}{k} + \sum_{i=\ell+1}^k N_{\text{ind}}(R, n_i) \leq \ell \binom{c_0}{k} + (\text{ind}(R) + \varepsilon) \sum_{i=\ell+1}^k \binom{n_i}{k}. \quad (21)$$

Observe that $\binom{n_i}{k} = \binom{p_i n}{k} < p_i^k \binom{n}{k}$ since $p_i < 1$. Dividing (21) by $\binom{n}{k}$ yields

$$\frac{1}{\binom{n}{k}} \sum_{i=1}^k N_{\text{ind}}(R, n_i) \leq \ell \frac{\binom{c_0}{k}}{\binom{n}{k}} + (\text{ind}(R) + \varepsilon) \sum_{i=\ell+1}^k p_i^k. \quad (22)$$

Suppose $\ell \geq 1$. Using our bounds on ε and M along with (20), we can further bound

$$\frac{1}{\binom{n}{k}} \sum_{i=1}^k N_{\text{ind}}(R, n_i) \leq \ell \frac{\binom{c_0}{k}}{\binom{n}{k}} + (\text{ind}(R) + \varepsilon) \sum_{i=\ell+1}^k p_i^k < 0.74 \text{ind}(R) \quad (23)$$

and bound the product term

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

This yields $\text{ind}(R, n) \leq 0.74 \text{ind}(R) + \varepsilon < \text{ind}(R)$, a contradiction. Thus $\ell = 0$, so using (22) we may rewrite (19) as

$$\text{ind}(R, n) \leq (\text{ind}(R) + \varepsilon) \sum_{i=1}^k p_i^k + \frac{1}{\binom{n}{k}} \prod_{i=1}^k n_i. \quad (24)$$

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

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

Plugging this into (24) and recalling $\text{ind}(R) = a + \gamma$,

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

The first bound $\sum p_i^k + (k^k - k) \prod p_i \leq 1$ is well-known (see, e.g. [30, Equation (17)]) and the second bound comes from (20). This gives the contradiction

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

since $\varepsilon < \gamma/100$. □

2.8. The absurd case. Now, we consider the other endpoint of I .

CLAIM 2.10. *If $\delta = 1 - \sum_i \binom{n_i}{2} / \binom{n}{2}$, then $\text{ind}(R) \leq a$.*

PROOF. If $\delta = 1 - \sum_i \binom{n_i}{2} / \binom{n}{2}$, then (18) implies that

$$N_{\text{ind}}(R, n) \leq \sum_{i=1}^k N_{\text{ind}}(R, n_i) + \frac{2 \sum_{i \neq j} n_i n_j}{k-2} \left(\frac{z}{k-2} \right)^{k-2} n^{k-2}. \quad (25)$$

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

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

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

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

for $k \geq 11$ 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 (23) in (25), and $a < \text{ind}(R)$ gives

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

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

3. Connected rainbow graphs

We give the proof of Theorem 2.3 in the following subsections.

3.1. Setup. Fix k and $R = ([k], E)$ a rainbow colored 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 $k \geq 11$ 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 -colored complete graph $([k], \binom{[k]}{2})$ with coloring 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 $\text{ind}(R) \leq a$. 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.

Suppose that $n > M$ is given and H is a T -colored n -vertex graph with coloring function χ_H achieving $N_{\text{ind}}(R, n)$. This implies

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

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 color 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 (5) 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 $V_1 \cup \dots \cup V_k$, $n_i = |V_i|$, 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 and $\text{diam}(G) = \max_{v \in V(G)} \epsilon_G(v)$ the *diameter* of G . For convenience, let $\epsilon(i) := \epsilon_R(i)$ for all $i \in [k]$.

Let $B(x)$ be the set of neighbors of 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 color $\{i, j\}$.

LEMMA 2.11. *Let $i, j \in [k]$ with $\{i, j\} \in T$ and $x \in V(H)$. Then*

$$\begin{aligned} \text{(a)} \quad d_i(x) &\leq \left(\frac{|B(x)|}{k_1(i)} \right)^{k_1(i)} \left(\frac{n - |B(x)|}{k - k_1(i) - 1} \right)^{k - k_1(i) - 1} \\ \text{(b)} \quad d_i(x) &\leq \left(\frac{|N_i(x)|}{k - 1} \right)^{k-1} \\ \text{(c)} \quad d_i(x) &\leq d_{\{i,j\}}(x) \cdot \left(\frac{n - d_{\{i,j\}}(x)}{k - 2} \right)^{k-2}. \end{aligned}$$

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 $k_1(i)$ 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 color 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 $m := k_1(i) + \dots + k_{r-1}(i)$, and suppose that we have chosen $y_1, y_2, \dots, y_m \in V(H)$ to play the roles of all vertices at distance $r - 1$ or less from i in R , where $y_1, \dots, y_{k_{r-1}(i)}$ play the roles of vertices at distance exactly $r - 1$ from i . Then

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

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 $y_1, \dots, y_{k_{r-1}}$. The color 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 (4) from Section 2.1 we see that

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

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

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

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

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

When $\ell = 1$ this becomes

$$\begin{aligned} d_i(x) &\leq \sum_{\vec{y}_1 \in P_1} p(|N_i(x)| - |B_1|, k - 1 - k_1) \\ &\leq p(|B_1|, k_1) \cdot p(|N_i(x)| - |B_1|, k - k_1 - 1). \end{aligned} \tag{26}$$

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

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

so (a) holds. Alternatively, (4) also yields

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

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

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

This combined with (26) and (4) gives

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

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

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

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

This follows from an identical Zykov symmetrization argument as used in the proof of Lemma 2.5. Note that we have assumed the same inequalities for M as we did in Section 2.1.

3.4. Maximum color and non-edge degrees. The following two claims are used in the proof of Lemma 2.15 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 colors $\{i, j\} \in T$. We upper bound this value.

CLAIM 2.13. $\alpha < \eta/4$.

PROOF. Let x achieve this maximum, so that $d_{\{i,j\}}(x) = \alpha n$ for some $\{i, j\} \in T$. By Lemma 2.11(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 color $\{i, j\}$ in R . Thus by Lemma 2.11(b), we get

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

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 2.12, we get

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

Rearranging yields

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

We see that

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

so

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

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

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

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

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 color \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 2.14. $\beta < 1 - \eta/2$.

PROOF. Assume for contradiction that $\beta \geq 1 - \eta/2$. 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 2.11(a) gives

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

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

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

$$d_i(x) \leq \left(\frac{n}{k - 1} \right)^{k-1} \left(\frac{(1 - \beta)^q \beta^{1-q}}{q^q (1 - q)^{1-q}} \right)^{k-1}. \quad (27)$$

We will first bound the term

$$\frac{(1 - \beta)^q \beta^{1-q}}{q^q (1 - q)^{1-q}}. \quad (28)$$

Regarding (28) 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 (28), i.e.

$$\frac{(1-\beta)^q \beta^{1-q}}{q^q (1-q)^{1-q}} \leq \frac{(\eta/2)^q (1-\eta/2)^{1-q}}{q^q (1-q)^{1-q}}. \quad (29)$$

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 (29) 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 (27) 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 2.12, 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 (8) 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 . \square

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 2.15. $z < 1 - \eta/8$.

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 $|N_1(x) \cap Z(x)| \leq d_{\{i,j\}}(x) + d_\emptyset(x)$. It follows that

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

where $d_{\{i,j\}}(x) \leq \alpha n$ and $d_\emptyset(x) \leq \beta n$. Thus, $|N_i(x)| + |Z(x)| < (2 - \eta/4)n$ by Claims 2.13 and 2.14. Since $|Z(x)| \leq |N_i(x)|$, this gives $z < 1 - \eta/8$. \square

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

LEMMA 2.16. $|V_i| \leq (1 - 1/3k)n$ for all $i \in [k]$.

PROOF. By contradiction, assume without loss of generality that $|V_1| > (1 - 1/3k)n$. If $x \in V_1$, then $|N_1(x)| \geq |N_i(x)|$ for all $i > 1$ so $|N_2(x)| \leq |Z(x)| \leq zn$. Applying Lemma 2.11(b) to $d_2(x)$ gives

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

and we further see that

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

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

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

Using the standard inequality (8) and then Lemma 2.15 gives

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

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

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 [30]. While we do not have as much information in this case, without a focus on optimizing for small k , we allow ourselves to be less strict with the counting arguments.

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 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 (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 $\varphi : [k] \rightarrow f$ with $\varphi(i) \in V_i$ for all i is not a graph isomorphism from $R \rightarrow H[f]$. Let $h_m = |H_m|$, $h_g = |H_g|$ and $h_b = |H_b|$ so that

$$N_{\text{ind}}(R, H) = h_m + h_g + h_b.$$

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

$$h_m = \sum_{j=1}^k N_{\text{ind}}(R, H[V_j]) \leq \sum_{j=1}^k N_{\text{ind}}(R, n_j) \tag{30}$$

and

$$h_g \leq \prod_{\ell=1}^k n_\ell \left(1 - \frac{\delta \binom{n}{2}}{\sum_{1 \leq i < j \leq k} n_i n_j} \right) = \prod_{\ell=1}^k n_\ell \left(1 - \frac{\delta \binom{n}{2}}{\binom{n}{2} - \sum_i \binom{n_i}{2}} \right). \quad (31)$$

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 colored 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 $V_1 \cup \dots \cup V_k = V(H)$. By the final statement of Lemma 2.11,

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

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

$$h_b \leq 2\delta(k-1) \binom{n}{2} \left(\frac{zn}{k-2} \right)^{k-2}. \quad (32)$$

Using (30), (31), and (32) we obtain

$$N_{\text{ind}}(R, n) \leq \sum_i N_{\text{ind}}(R, n_i) + \prod_{\ell} n_{\ell} \left(1 - \frac{\delta \binom{n}{2}}{\binom{n}{2} - \sum_i \binom{n_i}{2}} \right) + 2\delta(k-1) \binom{n}{2} \left(\frac{zn}{k-2} \right)^{k-2}. \quad (33)$$

Our final task is to upper bound the right side of (33).

As in Section 3.7, we see that $\delta \in I = [0, 1 - \sum_i \binom{n_i}{2} / \binom{n}{2}]$. Viewing (33) as a linear function of δ , it again suffices to check the endpoints of I .

3.8. The extremal case.

CLAIM 2.17. *If $\delta = 0$, then $\text{ind}(R) \leq a$.*

PROOF. If $\delta = 0$, then (33) implies that

$$N_{\text{ind}}(R, n) \leq \sum_{i=1}^k N_{\text{ind}}(R, n_i) + \prod_{i=1}^k n_i.$$

This is the same equation as (19), and we have all the same assumptions. The same argument as in Section 2.7 derives a contradiction. \square

3.9. The absurd case.

Now, we consider the other endpoint of I .

CLAIM 2.18. *If $\delta = 1 - \sum_i \binom{n_i}{2} / \binom{n}{2}$, then $\text{ind}(R) \leq a$.*

PROOF. If $\delta = 1 - \sum_i \binom{n_i}{2} / \binom{n}{2}$, then (33) implies that

$$N_{\text{ind}}(R, n) \leq \sum_{i=1}^k N_{\text{ind}}(R, n_i) + 2(k-1) \sum_{i \neq j} n_i n_j \left(\frac{z}{k-2} \right)^{k-2} n^{k-2}. \quad (34)$$

This is similar to (25) 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 (23):

$$\frac{1}{\binom{n}{k}} \sum_{i=1}^k N_{\text{ind}}(R, n_i) \leq \ell \frac{\binom{c_0}{k}}{\binom{n}{k}} + (\text{ind}(R) + \varepsilon) \sum_{i=\ell+1}^k p_i^k < 0.74 \text{ind}(R). \quad (35)$$

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

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

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

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

By Lemma 2.15, we know that $z^{k-2} \leq (1 - \eta/8)^{k-2}$, and we further see that

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

We again used $k \geq 11$ here. Thus for $C > 24/0.9 \approx 26.67$ and $k \geq 11$, we have $1.01e^2 k^3 z^{k-2} a < 0.25a$. Recalling that $a < \text{ind}(R)$, plugging this and (35) into (34) gives

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

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

4. Disconnected rainbow graphs

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

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

We begin by upper bounding the number of copies $N_{\text{ind}}(R, G_n)$ 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. [34]),

$$N_{\text{ind}}(R_i, G_n) = \binom{n}{k}^{c_i} + k \binom{n}{k^2}^{c_i} + k^2 \binom{n}{k^3}^{c_i} + \cdots = (1 + o(1)) \frac{n^{c_i}}{k^{c_i} - k}.$$

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

$$N_{\text{ind}}(R, G_n) \leq \prod_{i=1}^{\ell} N_{\text{ind}}(R_i, G_n) = (1 + o(1)) n^k \prod_{i=1}^{\ell} \frac{1}{k^{c_i} - k}. \quad (36)$$

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 $i \in [\ell]$, $||V_i| - \frac{c_i}{k}n| \leq 1$.
- (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 have that $N_{\text{ind}}(R, H_n) = \prod_i N_{\text{ind}}(R_i, H_n[V_i])$, so

$$N_{\text{ind}}(R, H_n) = \prod_{i=1}^{\ell} (1 + o(1)) \frac{c_i!}{c_i^{c_i} - c_i} \binom{\frac{c_i}{k}n}{c_i} = (1 + o(1)) n^k \prod_{i=1}^{\ell} \frac{1}{k^{c_i} - k \binom{k}{c_i}^{c_i-1}}. \quad (37)$$

Comparing (36) with (37), we subtract larger numbers in the denominator of (37), so the family of graphs $\mathcal{H}(n)$ induces asymptotically more copies than the family $\mathcal{G}_R(n)$. Thus R is not a fractalizer. Since R was generic, disconnected rainbow graphs without isolated vertices are not fractalizers.

5. Rainbow cliques without a fixed coloring

We now address the problem of maximizing the number of rainbow cliques on k vertices without fixing a specific rainbow coloring. In other words, for \mathcal{G}_n the set of $\binom{k}{2}$ -edge-colored

K_n , let

$$\gamma_k := \lim_{n \rightarrow \infty} \frac{\max_{G \in \mathcal{G}_n} |\{S \subset V(G) : |S| = k \text{ and } G[S] \text{ is rainbow}\}|}{\binom{n}{k}}. \quad (38)$$

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

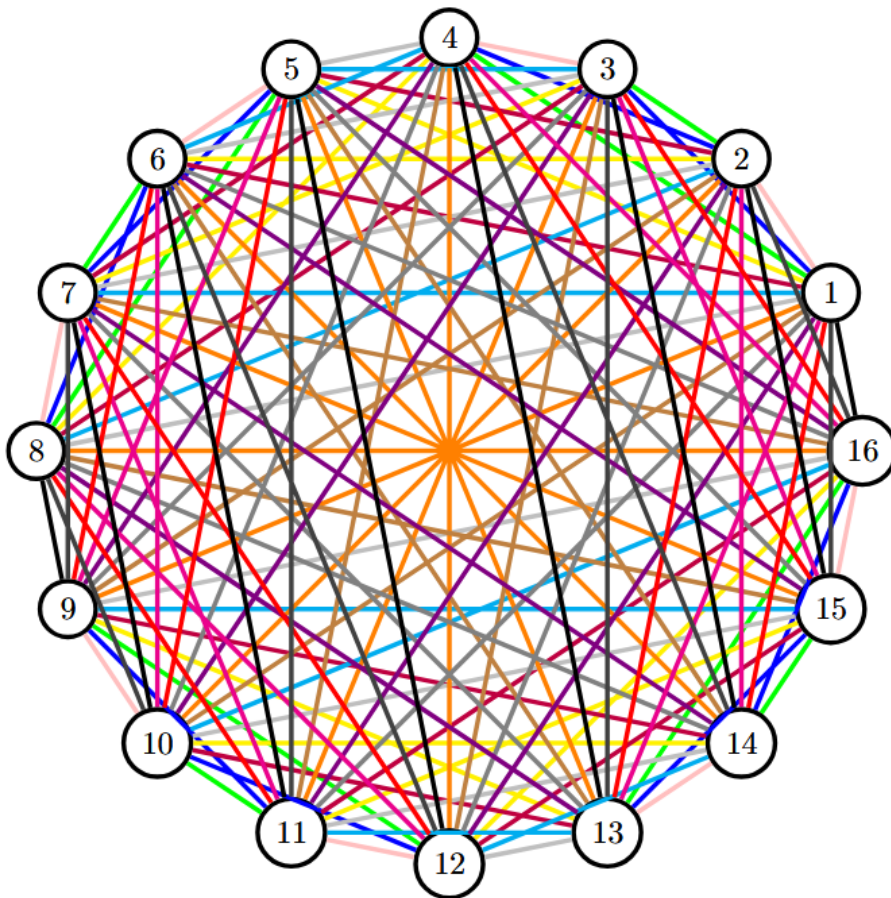


FIGURE 2.1. A 15-edge-coloring of K_{16} .

The result of [3] 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 $\binom{6}{2} = 15$ colors. Consider the 15-edge-coloring of K_{16} displayed in Figure 2.1 and note that each color class is a perfect matching. Letting the set of colors be $\{1, 2, \dots, 15\}$, the adjacency matrix for this graph is shown in Figure 2.2, where the i, j -th entry is the color assigned to the edge ij .

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

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

FIGURE 2.2. Adjacency matrix of the graph shown in Figure 2.1.

of this construction, we see that

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

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 $k > 6$.

CHAPTER 3

Induced oriented cycles for low edge density

This chapter is based on joint work with Clayton Mizgerd.

1. Background and results

Recall that Conjecture 1.1 states that $\text{ind}(C_k) = k!/(k^k - k)$ for $k \geq 5$. In this chapter, we consider the oriented cycle \vec{C}_k instead and prove the following.

THEOREM 3.1. *Let $k \geq 4$ be fixed. Let G be an oriented graph with n vertices and m edges. Then,*

$$N_{\text{ind}}(\vec{C}_k, G) \leq \frac{mn^{k-2}}{k^{k-1}}.$$

The bound in Conjecture 1.1 corresponds to

$$N_{\text{ind}}(\vec{C}_k, G) \leq \frac{n^k}{k^k}. \tag{39}$$

Since $m = O(n^2)$, the main difference is in the k^{k-1} in the denominator of our bound. In fact, Theorem 3.1 proves that (39) holds whenever $m \leq n^2 k^{k-1}/(k^k - k)$. Asymptotically, this is when m is about $(2/k) \cdot \binom{n}{2}$, which is the same number of edges as in the iterated balanced blow-up of \vec{C}_k on n vertices. Thus, up until the edge density at which inducibility is conjectured to be achieved, the desired bound holds for oriented cycles.

2. Proof of Theorem 3.1

Given k and G , we will count the number of k -tuples of vertices $(v_1, v_2, v_3, v_4, \dots, v_k)$ such that $v_1 v_2 \dots v_k$ is an induced oriented cycle of length k in G (so G contains edges $v_1 v_2, v_2 v_3$, etc., up to $v_k v_1$). As in [20], we call such k -tuples *good*. For $A = (v_1, \dots, v_k)$, define

$$w(A) = \prod_{i=2}^k \frac{1}{n_i},$$

where

- n_2 is m ,
- n_3 is the number of neighbors b of v_2 such that v_1v_2b is an induced oriented path of length 3,
- n_i for $i = 4, \dots, k$ is the number of neighbors b of v_{i-1} such that $v_1v_2 \dots v_{i-1}b$ is an induced oriented path of length i ,
- n_k is the number of vertices b such that $v_1v_2 \dots v_{k-1}b$ is an induced oriented cycle of length k .

Note that Král, Norin, and Volec [20] started by choosing a vertex, then chose additional vertices one at a time. The fact that we are first choosing an edge is the main difference in our work, which alters multiple steps and bounds throughout the proof of Theorem 3.1.

The following result shows that we can in some sense think of $w(A)$ as the probability of choosing A if we pick an edge at random, then add vertices one at a time to form a good k -tuple. We will call any connected subgraph of an oriented cycle an *oriented path*.

LEMMA 3.2. *Let $k \geq 4$ and G an oriented graph. Summing over all good k -tuples A ,*

$$\sum_A w(A) \leq 1.$$

PROOF. We consider the following random process for generating a k -tuple of vertices. We choose an edge $v_1v_2 \in E(G)$ uniformly at random. We then select a random vertex v_3 uniformly among those vertices $u \in V$ such that v_1v_2u induces an oriented path of length 2. We continue to iterate this process until $v_1 \dots v_{k-1}$ is an oriented path of length $k - 2$. We then choose v_k uniformly among those vertices $v \in V$ such that $\{v_1, \dots, v_{k-1}, v_k\}$ induces \vec{C}_k . If at any point there are no valid vertices and we are attempting to choose from an empty set of vertices, then we abort this process and fill in the remaining slots with any vertices.

The probability of drawing any “good” k -tuple A is exactly $w(A)$. Thus by the law of total probability, summing over all good k -tuples A ,

$$\sum_A w(A) \leq 1$$

as claimed. \square

Now that we have a way to weight the good k -tuples, we are prepared to prove Theorem 3.1. We follow a similar outline to the proof of [20, Theorem 1].

PROOF OF THEOREM 3.1. Let k be fixed. Let G be an oriented graph with n vertices and m edges. We count the number of good k -tuples in G .

Fix an induced cycle $v_1 v_2 \dots v_k$ and let A_j be the good k -tuple $(v_j, v_{j+1}, \dots, v_{j-1})$ modulo k for $j \in [k]$. Our main goal is to find a lower bound on $w(A_1) + \dots + w(A_k)$. The AM-GM inequality gives us that

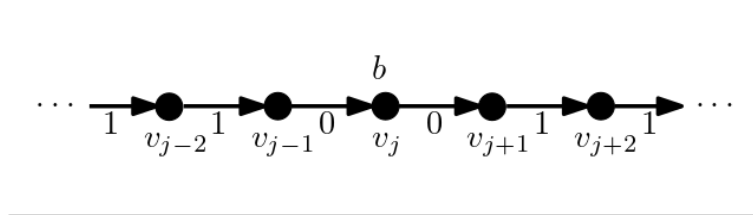
$$\left(\prod_{j=1}^k w(A_j) \right)^{\frac{1}{k}} \leq \frac{w(A_1) + \dots + w(A_k)}{k}. \quad (40)$$

Let $n_{j,\ell}$ be n_ℓ as in the calculation of $w(A_j)$. The definition of $w(A)$ and another application of the AM-GM inequality further implies that

$$\left(\prod_{j=1}^k \frac{1}{w(A_j)} \right)^{\frac{1}{k(k-2)}} = \left(\prod_{j=1}^k m n_{j,3} n_{j,4} \dots n_{j,k} \right)^{\frac{1}{k(k-2)}} \leq \frac{m^{\frac{1}{k-2}}}{k(k-2)} \cdot \sum_{j=1}^k \sum_{\ell=3}^k n_{j,\ell}. \quad (41)$$

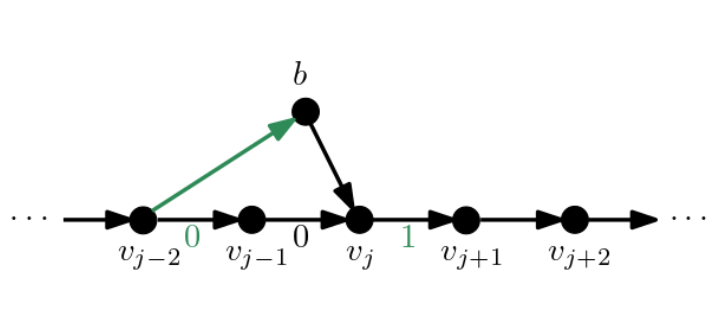
Let $b \in V(G)$ be any vertex. If b contributes 1 to $n_{j,\ell}$ for some $j \in [k]$ and $\ell \in \{3, 4, \dots, k\}$, then $v_{j+\ell-2} b \in E(G)$ and $v_i b \notin E(G)$ for all $j \leq i < j + \ell - 2$ modulo k . Therefore, b can contribute at most 1 to $\sum_{\ell=3}^k n_{j,\ell}$ for any $j \in [k]$. We show that each vertex b contributes at most $k-2$ to the double sum in (41) by showing that there exist two values $j_1, j_2 \in [k]$ such that b contributes 0 to both $\sum_{\ell=3}^k n_{j_1,\ell}$ and $\sum_{\ell=3}^k n_{j_2,\ell}$. We do so by cases:

Case 1: $b = v_j$ for some $j \in [k]$. See Figure 3.1 for an illustration. In this and following figures in this proof, each edge $v_i v_{i+1}$ is labeled by the contribution of b to $\sum_{\ell=3}^k n_{i,\ell}$. We see that b contributes 0 to $n_{j-1,\ell}$ and $n_{j,\ell}$ for all $\ell \in \{3, 4, \dots, k\}$, so we set $j_1 = j-1$ and $j_2 = j$.

FIGURE 3.1. A_j when $b = v_j$.

Case 2: $b \neq v_j$ for all $j \in [k]$ and b has no edges to or from v_j for any $j \in [k]$. We see that $n_{j,\ell} = 0$ for all $j \in [k]$ and $\ell \in \{3, 4, \dots, k\}$, so we may take any distinct $j_1, j_2 \in [k]$.

Case 3: $b \neq v_j$ for all $j \in [k]$ and $bv_j \in E(G)$ for some $j \in [k]$. We see that b contributes 0 to $n_{j-1,\ell}$ for all $\ell \in \{3, 4, \dots, k\}$. If b contributes 0 to $n_{j,\ell}$ for all $\ell \in \{3, 4, \dots, k\}$ as well, then we may set $j_1 = j - 1$ and $j_2 = j$. Otherwise, b must contribute 1 to $n_{j,k}$ since the edge v_jv_{j+1} is picked first, and a vertex with an edge to v_j will only be added if it is the last vertex, completing the induced oriented cycle. For this to occur, we must have that $v_{j-2}b \in E(G)$. See Figure 3.2 for an illustration, where the edge $v_{j-2}b$ and its implications are depicted in green. Thus, b contributes 0 to $n_{j-2,\ell}$ for all $\ell \in \{3, 4, \dots, k\}$, so we set $j_1 = j - 2$ and $j_2 = j - 1$.

FIGURE 3.2. A_j when $bv_j \in E(G)$.

Case 4: $b \neq v_j$ for all $j \in [k]$ and $v_jb \in E(G)$ for some $j \in [k]$. We see that b contributes 0 to $n_{j,\ell}$ for all $\ell \in \{3, 4, \dots, k\}$. If b contributes 0 to $n_{j+1,\ell}$ for all $\ell \in \{3, 4, \dots, k\}$ as well,

then it contributes 0 to both $\sum_{\ell=3}^k n_{j,\ell}$ and $\sum_{\ell=3}^k n_{j+1,\ell}$. Otherwise, there must be some $j_0 \in [k] \setminus \{j, j+1\}$ such that $v_{j_0}b \in E(G)$. See Figure 3.3 for an illustration, where the edge $v_{j_0}b$ and its implications are depicted in green. Thus, b contributes 0 to $n_{j_0,\ell}$ for all $\ell \in \{3, 4, \dots, k\}$ and we are done.

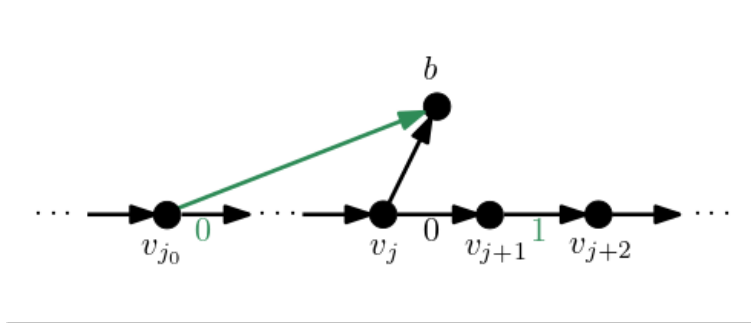


FIGURE 3.3. A_j when $v_j b \in E(G)$.

Thus we have that

$$\sum_{j=1}^k \sum_{\ell=3}^k n_{j,\ell} \leq n(k-2).$$

This, together with (41), shows that

$$\left(\prod_{j=1}^k \frac{1}{w(A_j)} \right)^{\frac{1}{k(k-2)}} \leq \frac{nm^{\frac{1}{k-2}}}{k},$$

which further implies that

$$\frac{k^{k-2}}{mn^{k-2}} \leq \left(\prod_{j=1}^k w(A_j) \right)^{\frac{1}{k}}. \quad (42)$$

Equations (40) and (42) together give us that

$$\frac{k^{k-1}}{mn^{k-2}} \leq w(A_1) + \dots + w(A_k). \quad (43)$$

Summing over all good k -tuples and applying Lemma 3.2, we see that

$$\frac{k^{k-1}}{mn^{k-2}} \cdot N_{\text{ind}}(\vec{C}_k, G) \leq 1.$$

Therefore,

$$N_{\text{ind}}(\vec{C}_k, G) \leq \frac{mn^{k-2}}{k^{k-1}}.$$

as desired. \square

The bound in Theorem 3.1 is not tight, but we believe that the only place where slack appears in the proof is in (42).

As mentioned earlier, all changes in our proof from that in [20] come from the fact that we choose an edge first, and then vertices, when finding an induced oriented cycle in G , rather than just choosing vertices. This means that our bound depends on m , but it also means that b contributes at most $k - 2$ to the double sum in (41), rather than at most $k - 1$ in [20]. The orientation of the edges also implies that b can contribute at most 1 to $\sum_{\ell=3}^k n_{j,\ell}$, while without orientation it can contribute at most 2. Thus, while our bound varies with the edge density of G , it is much closer to the conjectured bound for lower edge densities than the bound in [20].

CHAPTER 4

The feasible region of edge-colored complete graphs

This chapter is based on joint work with Dhruv Mubayi in [8], published in the SIAM Journal on Discrete Mathematics.

1. Background and results

A q -colored graph is a graph $G = (V, E)$ together with a coloring function $f : E \rightarrow C$, where $|C| = q$. Fix $q \in \mathbb{Z}^+$ and let $G = (V, E)$ be a q -colored complete graph with coloring function f . For $1 \leq i \leq q$, let $e_i(G) = |\{e \in E : f(e) = i\}|$ and let

$$\varrho_i(G) := \frac{e_i(G)}{\binom{|V|}{2}}$$

be the *density of color i* . Given a q -colored complete graph F with $|V(F)| = s$ and coloring function g , a subset $X \subset V$ with $|X| = s$ is a *copy of F in G* if there is a bijection $\varphi : V(F) \rightarrow X$ such that $g(uv) = g(\varphi(u)\varphi(v))$ for all distinct $u, v \in V(F)$. Let $N_q(F, G)$ be the number of copies of F in G and let

$$\varrho_q(F, G) := \frac{N_q(F, G)}{\binom{|V|}{s}}$$

be the *density of F in G* .

Let $(G_n)_{n=1}^\infty$ be a sequence of q -colored complete graphs with $|V(G_n)| \rightarrow \infty$. The sequence $(G_n)_{n=1}^\infty$ is *F -good* if $x_i = \lim_{n \rightarrow \infty} \varrho_i(G_n)$ exists for all $i \in [q]$ and $y = \lim_{n \rightarrow \infty} \varrho_q(F, G_n)$ exists. In this case, $(G_n)_{n=1}^\infty$ *realizes* (x_1, \dots, x_{q-1}, y) . Note that we only list x_1, \dots, x_{q-1} since $x_q = 1 - (x_1 + \dots + x_{q-1})$. Define

$$I_q(F, (x_1, \dots, x_{q-1})) := \sup \{y : (x_1, \dots, x_{q-1}, y) \in [0, 1]^q \text{ realized by some } F\text{-good } (G_n)\}.$$

In fact, this supremum is a maximum (see Proposition 4.8).

For $2 \leq s \leq t$, recall that $K'_{s,t}$ is the 3-colored clique on vertex set $V = V_1 \sqcup V_2$ with $|V_1| = s$ and $|V_2| = t$ with coloring function f defined by

$$f(ij) := \begin{cases} \text{blue} & \text{if } i, j \in V_1, \\ \text{green} & \text{if } i, j \in V_2, \\ \text{red} & \text{otherwise,} \end{cases}$$

for all distinct $i, j \in [s+t]$ (see Figure 1.1 from Chapter 1). Our main theorem is as follows.

THEOREM 4.1. *Let $2 \leq s \leq t$ and $x_b, x_g, x_r \in [0, 1]$ such that $x_b + x_g + x_r = 1$. Then*

$$I_3(K'_{s,s}, (x_b, x_g)) = \begin{cases} x_b^{s/2} x_g^{s/2} \binom{2s}{s} & \text{if } \sqrt{x_b} + \sqrt{x_g} \leq 1, \\ \left(\frac{x_r}{2}\right)^s \binom{2s}{s} & \text{if } \sqrt{x_b} + \sqrt{x_g} \geq 1. \end{cases}$$

Furthermore, if $\sqrt{x_b} + \sqrt{x_g} \leq 1$, then

$$I_3(K'_{s,t}, (x_b, x_g)) = x_b^{s/2} x_g^{t/2} \binom{s+t}{s}.$$

Theorem 4.1 determines $I_3(K'_{s,s}, (x_b, x_g))$ for all vectors (x_b, x_g) in the region $x_b, x_g \geq 0$ and $x_b + x_g \leq 1$. A simple optimization shows that

$$\text{ind}(K'_{s,s}) = I_3(K'_{s,s}, (1/4, 1/4)) = \left(\frac{1}{4}\right)^s \binom{2s}{s}.$$

Note that $I_3(K'_{s,t}, (x_b, x_g))$ is not known when $\sqrt{x_b} + \sqrt{x_g} > 1$ and $s \neq t$.

We will now consider all possible subgraph densities of $K'_{s,t}$, or any other edge-colored complete graph. This work follows the structure of Liu, Mubayi, and Reiher in [26]. Let $q \in \mathbb{Z}^+$ be the number of colors, F a q -colored complete graph on s vertices, G a q -colored complete graph, and $(G_n)_{n=1}^\infty$ a sequence of q -colored complete graphs. Then, a *q-colored quantum graph* Q is a formal linear combination of finitely many q -colored complete graphs, i.e.

$$Q = \sum_{i=1}^m \lambda_i F_i$$

where $m \in \mathbb{Z}_{\geq 0}$, $\lambda_i \in \mathbb{R}$ for all i , and F_i are all q -colored complete graphs. All definitions on q -colored complete graphs extend to q -colored quantum graphs by linearity. For convenience, we set $\vec{x} := (x_1, \dots, x_{q-1}) = \lim_{n \rightarrow \infty} (\varrho_1(G_n), \dots, \varrho_{q-1}(G_n))$. The *feasible region* of Q is

$$\Omega_q(Q) = \{(\vec{x}, y) \in [0, 1]^q : \text{realized by } Q\text{-good } (G_n)_{n=1}^\infty\}.$$

Similarly to our definition of I_q , we define

$$i_q(Q, \vec{x}) := \inf\{y : (\vec{x}, y) \in \Omega_q(Q)\},$$

which we also later show is in fact a minimum (see Proposition 4.8).

Set $\Delta'_{q-1} := \{\vec{x} \in [0, 1]^{q-1} : x_1 + \dots + x_{q-1} \leq 1\}$ a projection of the $(q-1)$ -dimensional standard simplex. For most graphs F , the minimum subgraph density $i_q(F, \vec{x})$ is simply zero.

PROPOSITION 4.2. *Given a q -colored non-monochromatic complete graph F , $i_q(F, \vec{x}) = 0$ for all $\vec{x} \in \Delta'_{q-1}$.*

When F is monochromatic, the minimum clique density bounds on simple graphs from [35, 31, 36] apply directly. See also Remark 4.4.

The following result shows that $\Omega_q(F)$ is completely determined by $I_q(F, \vec{x})$ and $i_q(F, \vec{x})$.

THEOREM 4.3. *For every q -colored quantum graph Q ,*

$$\Omega_q(Q) = \{(\vec{x}, y) \in [0, 1]^{q-1} \times \mathbb{R} : i_q(Q, \vec{x}) \leq y \leq I_q(Q, \vec{x})\}.$$

Moreover, the boundary functions $i_q(Q, \vec{x})$ and $I_q(Q, \vec{x})$ are continuous and their partial derivatives exist almost everywhere.

For $i \in [q]$, let ϕ_i be the function that assigns to every q -colored graph F a simple graph $G = (V(F), E)$ where E is the set of edges colored i in F .

REMARK 4.4. *For any q -colored complete graph F , we have that $I_q(F, \vec{x}) \leq I(\phi_i(F), x_i)$ for all $i \in [q]$ where $x_q = 1 - (x_1 + \dots + x_{q-1})$. This is because, given any F -good sequence realizing a specific F -density, if we just delete all edges except those of color i in both the*

sequence and F , the new sequence is still $\phi_i(F)$ -good and realizes the same, or higher, $\phi_i(F)$ -density. Written differently,

$$I_q(F, \vec{x}) \leq \min_{i \in [q]} \{I(\phi_i(F), x_i)\}.$$

The second case of the proof of Theorem 4.1 is an example showing that this inequality can be tight.

By Proposition 4.2, we know that $i_3(K'_{s,t}, (x_b, x_g)) = 0$ for all $x_b, x_g, x_r \in [0, 1]$ such that $x_b + x_g + x_r = 1$ since $K'_{s,t}$ is not monochromatic. Both $i_3(K'_{s,t}, \vec{x})$ and $I_3(K'_{s,t}, \vec{x})$ are clearly continuous and almost everywhere differentiable since $I_3(K'_{s,t}, \vec{x})$ is only not differentiable on the line $\sqrt{x_b} + \sqrt{x_g} = 1$, which has measure 0. A visualization of the feasible region of $K'_{2,2}$ is shown in Figure 4.1. The horizontal axes are x_b and x_g , while the vertical axis is y . The feasible region itself looks somewhat like a sand dune.

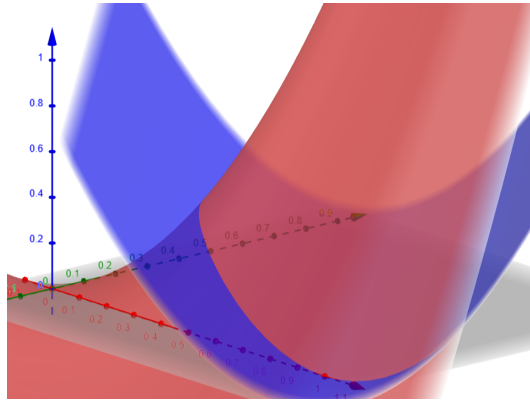


FIGURE 4.1. $\Omega_3(K'_{2,2})$ (graphed at geogebra.org).

In Section 2, we prove Theorem 4.1. In Section 3, we prove Proposition 4.2 and Theorem 4.3 along with other supporting results.

2. Proof of Theorem 4.1

We start with some necessary definitions for simple graphs. Let $(G_n) := (G_n)_{n=1}^\infty$ be a sequence of graphs with $\lim_{n \rightarrow \infty} |V(G_n)| = \infty$. The sequence (G_n) is F -good if both

$x = \lim_{n \rightarrow \infty} \varrho(G_n)$ and $y = \lim_{n \rightarrow \infty} \varrho_{\text{ind}}(F, G_n)$ exist. In this case, we say that (G_n) realizes (x, y) . Define

$$I_{\text{ind}}(F, x) := \max\{y : (x, y) \in [0, 1]^2 \text{ is realized by some } F\text{-good } (G_n)\},$$

$$i_{\text{ind}}(F, x) := \min\{y : (x, y) \in [0, 1]^2 \text{ is realized by some } F\text{-good } (G_n)\}$$

as in [26]. The following two results will be needed to give upper bounds.

THEOREM 4.5 (Kruskal–Katona [19, 21]). *Let $r \geq 2$ be an integer. Then for every $x \in [0, 1]$ we have $I_{\text{ind}}(K_r, x) \leq x^{r/2}$.*

THEOREM 4.6 (Liu, Mubayi, Reiher [26, Theorem 1.16]). *Let $s \geq 2$ be an integer. Then for every $x \in [0, 1]$ we have $I_{\text{ind}}(K_{s,s}, x) \leq \binom{2s}{s} x^s / 2^s$.*

PROOF OF THEOREM 4.1. We address the two cases separately.

Case 1: $\sqrt{x_b} + \sqrt{x_g} \leq 1$.

We first prove the upper bound. Suppose that $(G_n)_{n=1}^\infty$ is a $K'_{s,t}$ -good sequence that realizes $((x_b, x_g), I_3(K'_{s,t}, (x_b, x_g)))$ and let f_n be the coloring function for G_n . Let $K'_s = ([s], \binom{s}{2})$ with coloring function $f_b \equiv \text{blue}$. Similarly, let $K'_t = ([t], \binom{t}{2})$ have coloring function $f_g \equiv \text{green}$. Since $K'_{s,t}$ contains exactly one copy of K'_s and one copy of K'_t ,

$$N_3(K'_{s,t}, G_n) \leq N_3(K'_s, G_n) N_3(K'_t, G_n). \quad (44)$$

By deleting edges that are not colored blue and applying Theorem 4.5, we obtain

$$\limsup_{n \rightarrow \infty} \frac{N_3(K'_s, G_n)}{\binom{|V(G_n)|}{s}} \leq I_3(K'_s, (x_b, x_g)) \leq I_{\text{ind}}(K_s, x_b) \leq x_b^{s/2}.$$

Similarly, by deleting edges not colored green, we obtain

$$\limsup_{n \rightarrow \infty} \frac{N_3(K'_t, G_n)}{\binom{|V(G_n)|}{t}} \leq I_3(K'_t, (x_b, x_g)) \leq I_{\text{ind}}(K_t, x_g) \leq x_g^{t/2}.$$

Together with (44), we get

$$I_3(K'_{s,t}, (x_b, x_g)) \leq \lim_{n \rightarrow \infty} \frac{x_b^{s/2} \binom{|V(G_n)|}{s} \cdot x_g^{t/2} \binom{|V(G_n)|}{t}}{\binom{|V(G_n)|}{s+t}} = x_b^{s/2} x_g^{t/2} \binom{s+t}{s}.$$

We now prove the lower bound. As $\lfloor n\sqrt{x_b} \rfloor + \lfloor n\sqrt{x_g} \rfloor \leq n\sqrt{x_b} + n\sqrt{x_g} \leq n$, there exist disjoint subsets A_n and B_n of $[n]$ with $|A_n| = \lfloor n\sqrt{x_b} \rfloor$ and $|B_n| = \lfloor n\sqrt{x_g} \rfloor$ for all n . For each $n \geq 1$, let $G_n = ([n], \binom{[n]}{2})$ have coloring function f_n defined by

$$f_n(ij) = \begin{cases} \text{blue} & \text{if } i, j \in A_n, \\ \text{green} & \text{if } i, j \in B_n, \\ \text{red} & \text{otherwise.} \end{cases}$$

See Figure 4.2 for an illustration. Clearly, $\lim_{n \rightarrow \infty} \varrho_{\text{blue}}(G_n) = x_b$ and $\lim_{n \rightarrow \infty} \varrho_{\text{green}}(G_n) =$

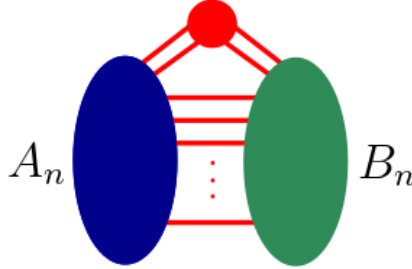


FIGURE 4.2. Construction maximizing $K'_{s,t}$ -density when $\sqrt{x_b} + \sqrt{x_g} \leq 1$.

x_g . If A is a copy of K'_s in G_n and B is a copy of K'_t in G_n , then $A \cup B$ is a copy of $K'_{s,t}$ in G_n , as $f_n(ij) = \text{red}$ for all $i \in A_n, j \in B_n$. Thus, $(G_n)_{n=1}^\infty$ realizes (x_b, x_g, y) , where

$$y = \lim_{n \rightarrow \infty} \varrho_3(K'_{s,t}, G_n) = \lim_{n \rightarrow \infty} \frac{\binom{\lfloor n\sqrt{x_b} \rfloor}{s} \binom{\lfloor n\sqrt{x_g} \rfloor}{t}}{\binom{n}{s+t}} = \lim_{n \rightarrow \infty} \frac{(s+t)! n^s x_b^{s/2} n^t x_g^{t/2}}{n^{s+t} s! t!} = x_b^{s/2} x_g^{t/2} \binom{s+t}{s}.$$

Consequently, $I_3(K'_{s,t}, (x_b, x_g)) \geq x_b^{s/2} x_g^{t/2} \binom{s+t}{s}$.

Case 2: $\sqrt{x_b} + \sqrt{x_g} \geq 1$.

We first prove the upper bound. Set $x_r := 1 - x_b - x_g$. Let $(G_n)_{n=1}^\infty$ be a $K'_{s,s}$ -good sequence that realizes $((x_b, x_g), I_3(K'_{s,s}, (x_b, x_g)))$. Every copy of $K'_{s,s}$ contains a red copy of

$K_{s,s}$. Consequently, Theorem 4.6 implies that

$$I_3(K'_{s,s}, (x_b, x_g)) = \lim_{n \rightarrow \infty} \frac{N_3(K'_{s,s}, G_n)}{\binom{|V(G_n)|}{2s}} \leq I_{\text{ind}}(K_{s,s}, x_r) \leq \left(\frac{x_r}{2}\right)^s \binom{2s}{s}.$$

We now prove the lower bound. Note that

$$\lfloor n\sqrt{x_b} \rfloor + \left\lfloor n \frac{x_r}{2\sqrt{x_b}} \right\rfloor \leq \frac{2x_b + x_r}{2\sqrt{x_b}} n = \frac{1 + x_b - x_g}{2\sqrt{x_b}} n \leq \frac{2\sqrt{x_b}}{2\sqrt{x_b}} n = n$$

since $\sqrt{x_b} + \sqrt{x_g} \geq 1$ implies $x_g \geq 1 - 2\sqrt{x_b} + x_b$. For each n , let $A_n \sqcup B_n \subseteq [n]$, where $|A_n| = \lfloor n\sqrt{x_b} \rfloor$ and $|B_n| = \lfloor nx_r/(2\sqrt{x_b}) \rfloor$ and let $G_n = ([n], \binom{[n]}{2})$ have coloring function f_n defined by

$$f_n(ij) = \begin{cases} \text{blue} & \text{if } i, j \in A_n, \\ \text{red} & \text{if } i \in A_n, j \in B_n \text{ or } j \in A_n, i \in B_n, \\ \text{green} & \text{otherwise.} \end{cases}$$

See Figure 4.3 for an illustration. Now $\lim_{n \rightarrow \infty} \varrho_{\text{blue}}(G_n) = x_b$, and

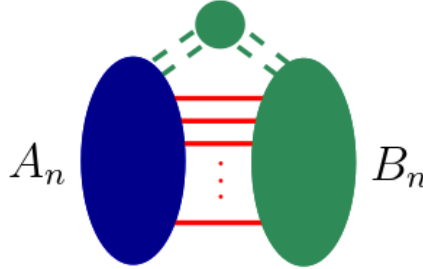


FIGURE 4.3. Construction maximizing $K'_{s,s}$ -density when $\sqrt{x_b} + \sqrt{x_g} > 1$.

$$\lim_{n \rightarrow \infty} \varrho_{\text{red}}(G_n) = \lim_{n \rightarrow \infty} \frac{\lfloor n\sqrt{x_b} \rfloor \cdot \lfloor nx_r/2\sqrt{x_b} \rfloor}{\binom{n}{2}} = \frac{2x_r\sqrt{x_b}}{2\sqrt{x_b}} = x_r.$$

We conclude that $\lim_{n \rightarrow \infty} \varrho_{\text{green}}(G_n) = 1 - x_b - x_r = x_g$. Every pair of s vertices in A_n and s vertices in B_n induces a copy of $K'_{s,s}$, so $(G_n)_{n=1}^\infty$ realizes (x_b, x_g, y) , where

$$y = \lim_{n \rightarrow \infty} \frac{\binom{\lfloor n\sqrt{x_b} \rfloor}{s} \cdot \binom{\lfloor nx_r/2\sqrt{x_b} \rfloor}{s}}{\binom{n}{2s}} = \lim_{n \rightarrow \infty} \frac{n^s x_b^{s/2} n^s x_r^s x_b^{-s/2}}{n^{2s} 2^s} \cdot \binom{2s}{s} = \left(\frac{x_r}{2}\right)^s \binom{2s}{s}.$$

Consequently, $I_3(F, (x_b, x_g)) \geq \left(\frac{x_r}{2}\right)^s \binom{2s}{s}$. □

3. The feasible region

The results in this section follow a similar structure to those in [26], and we make use of the following proposition.

PROPOSITION 4.7 (Liu, Mubayi, Reiher [26, Proposition 1.4]). *If F denotes a graph which is neither complete nor empty, then $i_{\text{ind}}(F, x) = 0$ for all $x \in [0, 1]$.*

This allows us to prove Proposition 4.2.

PROOF OF PROPOSITION 4.2. Let $\vec{x} = (x_1, x_2, \dots, x_{q-1}) \in \Delta'_{q-1}$ and pick $i \in [q-1]$ such that $\varrho_i(F) \neq 0$. Note that $\varrho_i(F) \neq 1$ by assumption. Then, consider the uncolored simple graph F_i obtained from F by only keeping edges with color i . As F_i is neither complete nor empty, we can apply Proposition 4.7. Then $i_{\text{ind}}(F_i, x_i) = 0$, so there is a sequence of graphs $(G_n)_{n=1}^{\infty}$ that realizes $(x_i, 0)$. Now, we construct a new sequence $(G'_n)_{n=1}^{\infty}$ where G'_n is a complete graph on $V(G_n)$. Given $e \in E(G_n)$, we color $e \in E(G'_n)$ color i . Then, we arbitrarily color $x_j \binom{n}{2} + o(1)$ edges color j for all $j \in [q] \setminus \{i\}$ so that all color densities are correct asymptotically. Then, $(G'_n)_{n=1}^{\infty}$ must realize $(\vec{x}, 0)$ since

$$\varrho_q(F, G'_n) \leq \varrho_{\text{ind}}(F_i, G_n)$$

for all n , and $\lim_{n \rightarrow \infty} \varrho_{\text{ind}}(F_i, G_n) = 0$. □

Not only is the lower bound on the feasible region zero for most q -colored graphs F , but the feasible region itself is also closed for all F (and all q -colored quantum graphs Q).

PROPOSITION 4.8. *For every q -colored quantum graph Q , the set $\Omega_q(Q)$ is closed.*

PROOF. We follow the same proof structure of [25, Proposition 1.3].

Let (\vec{x}, y) be any limit point of $\Omega_q(Q)$. We will build a sequence of q -colored complete graphs $(G_n)_{n=1}^{\infty}$ that will be Q -good and realize (\vec{x}, y) . This will imply by definition that $(\vec{x}, y) \in \Omega_q(Q)$, and thus that $\Omega_q(Q)$ is closed.

Let $d(\cdot, \cdot)$ represent standard Euclidean distance. If each G_n in the sequence satisfies $v(G_n) \geq n$, $d((\varrho_1(G_n), \dots, \varrho_{q-1}(G_n)), \vec{x}) \leq \frac{1}{n}$, and $|\varrho_q(Q, G_n) - y| \leq \frac{1}{n}$, then $(G_n)_{n=1}^{\infty}$

will be Q -good and realize (\vec{x}, y) . To construct such a sequence, for a given n , we take $(\vec{x}', y') \in \Omega_q(Q)$ such that $d(\vec{x}, \vec{x}') \leq \frac{1}{2n}$ and $|y - y'| \leq \frac{1}{2n}$. This is possible since (\vec{x}, y) is a limit point. Since $(\vec{x}', y') \in \Omega_q(Q)$, there is a Q -good sequence of q -colored complete graphs realizing (\vec{x}', y') . Therefore, there is some q -colored complete graph G_n with at least n vertices such that $d((\varrho_1(G_n), \dots, \varrho_{q-1}(G_n)), \vec{x}') \leq \frac{1}{2n}$ and $|\varrho_q(Q, G_n) - y'| \leq \frac{1}{2n}$. Thus, for this choice of G_n , we find that $d((\varrho_1(G_n), \dots, \varrho_{q-1}(G_n)), \vec{x})$ is at most

$$d((\varrho_1(G_n), \dots, \varrho_{q-1}(G_n)), \vec{x}') + d(\vec{x}', \vec{x}) \leq \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n}$$

and

$$|\varrho_q(Q, G_n) - y| \leq |\varrho_q(Q, G_n) - y'| + |y' - y| \leq \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n},$$

so $(G_n)_{n=1}^\infty$ realizes (\vec{x}, y) as desired. \square

Thus, “sup” and “inf” can be replaced with “max” and “min” respectively in the definitions of I_q and i_q . The next result shows that all subgraph densities between the minimum and the maximum can be achieved.

PROPOSITION 4.9. *Let Q be a q -colored quantum graph, $\vec{x} \in \Delta'_{q-1}$, and $y_1 < y_2$. If $(\vec{x}, y_1), (\vec{x}, y_2) \in \Omega_q(Q)$, then $(\vec{x}, y) \in \Omega_q(Q)$ for all $y \in [y_1, y_2]$.*

Note that, if Q has no monochromatic constituents, we can take $y_1 = 0$ by Proposition 4.2.

PROOF. Fix $y \in [y_1, y_2]$. Let $(G'_n)_{n=1}^\infty$ be a Q -good sequence of q -colored complete graphs realizing (\vec{x}, y_1) and similarly for $(G''_n)_{n=1}^\infty$ realizing (\vec{x}, y_2) . By diagonalization, we may assume that $(G'_n)_{n=1}^\infty$ and $(G''_n)_{n=1}^\infty$ are F -good for all q -colored complete graphs F as there are countably many such graphs for fixed q . Thus, by [27, Theorem 17.8], both sequences converge to a limit object, called a $[q]$ -graphon.

While this proof is possible without the use of graphons, it is much more pleasant to read with them. Since we only use them in this proof in this chapter, we only mention necessary definitions and results. For more details and context, see the book by Lovász [27]. A *graphon* W is a symmetric, measurable function $W : [0, 1]^2 \rightarrow [0, 1]$ and is the limit

object of a convergent sequence of dense graphs. A $[q]$ -*graphon* ω consists of q graphons W_i for $i \in [q]$, where W_i corresponds to edges of color i , and is the limit object of q -colored complete graphs. Let ω' be the limit of $(G'_n)_{n=1}^\infty$ and ω'' the limit of $(G''_n)_{n=1}^\infty$. Let ω' be represented by W'_1, \dots, W'_q and ω'' by W''_1, \dots, W''_q .

By the definition of convergence,

$$\int_{[0,1]^2} W'_i(x, y) \, dx dy = \int_{[0,1]^2} W''_i(x, y) \, dx dy = x_i$$

for all $i \in [q]$. In this sense, ω' and ω'' have the same color densities as $(G'_n)_{n=1}^\infty$ and $(G''_n)_{n=1}^\infty$. Similarly, for F a q -colored complete graph on vertex set $[k]$ with coloring function χ , the *homomorphism density of F in ω'* is

$$t(F, \omega') := \int_{[0,1]^k} \prod_{i \in [q]} \prod_{\substack{a, b \in [k] \\ \chi(ab)=i}} W'_i(x_a, x_b) \, d_{x_1} \cdots d_{x_k} = y_1$$

and the homomorphism density of F in ω'' is

$$t(F, \omega'') := \int_{[0,1]^k} \prod_{i \in [q]} \prod_{\substack{a, b \in [k] \\ \chi(ab)=i}} W''_i(x_a, x_b) \, d_{x_1} \cdots d_{x_k} = y_2.$$

This extends by linearity to Q .

Take any $\lambda \in [0, 1]$ and consider the $[q]$ -graphon defined by $\lambda\omega' + (1 - \lambda)\omega''$, i.e. the $[q]$ -graphon represented by $\lambda W'_i + (1 - \lambda)W''_i$ for all $i \in [q]$. The density of color i is still x_i for all $i \in [q]$. For a graph F , it is straightforward to check that $t(F, \lambda\omega' + (1 - \lambda)\omega'')$ is a polynomial of λ with bounded degree and coefficients as F and q are fixed. This extends by linearity to any fixed Q and implies that the homomorphism density of Q is continuous in λ . By the Intermediate Value Theorem, there exists some $\lambda \in [0, 1]$ such that $\omega := \lambda\omega' + (1 - \lambda)\omega''$ satisfies $t(Q, \omega) = y$.

As with simple graphs and graphons (see [27, Proposition 11.32]), we may use ω to generate a random sequence of q -colored complete graphs that converge to ω with probability 1. Therefore, there exists a convergent sequence $(G_n)_{n=1}^\infty$ of q -colored complete graphs with limit ω . By definition, $(G_n)_{n=1}^\infty$ realizes (\vec{x}, y) . Therefore, $(\vec{x}, y) \in \Omega_q(Q)$. \square

We will need the following technical lemma. Its proof is similar to that of [26, Lemma 2.3], so we only include the parts of the proof that are different.

LEMMA 4.10. *For every q -colored quantum graph Q and fixed $\vec{x}_0 := (x_2, \dots, x_{q-1}) \in \Delta'_{q-2}$, there exist constants $\ell \geq 1$ and $C \geq 0$ such that for all x, x' with $0 < x \leq x' \leq 1 - (x_2 + \dots + x_{q-1})$ we have*

$$\frac{I_q(Q, (x', \vec{x}_0))}{(x')^\ell} \leq \frac{I_q(Q, (x, \vec{x}_0))}{x^\ell} + C \left(\left(\frac{1}{x} \right)^\ell - \left(\frac{1}{x'} \right)^\ell \right).$$

PROOF. Fix $0 < x \leq x' \leq 1$, set $\alpha = \sqrt{x'/x} - 1$, and consider a Q -good sequence $(G'_n)_{n=1}^\infty$ that realizes $((x', \vec{x}_0), I_q(Q, (x', \vec{x}_0)))$. Assume without loss of generality that $v(G'_n) = n$ for every $n \geq 1$. Let G_n be the q -colored complete graph which is the union of G'_n and a set of $\lfloor \alpha n \rfloor$ other vertices, where no new edges have color 1. We see that

$$\lim_{n \rightarrow \infty} \varrho_1(G_n) = \lim_{n \rightarrow \infty} \frac{\varrho_1(G'_n) \binom{n}{2}}{\binom{n + \lfloor \alpha n \rfloor}{2}} = \frac{x'}{(1 + \alpha)^2} = x.$$

We then color the edges between G'_n and these new vertices, and these new vertices and themselves, colors $2, 3, \dots, q$ so that $\lim_{n \rightarrow \infty} \varrho_i(G_n) = x_i$ for all $i \in \{2, \dots, q-1\}$ and $\lim_{n \rightarrow \infty} \varrho_q(G_n) = x'_q + (x' - x)$. This is possible since, before coloring the new edges, all color densities are less than or equal to $x, x_2, \dots, x'_q + (x' - x)$, and these densities add to one. Thus,

$$I_q(Q, (x, \vec{x}_0)) \geq \limsup_{n \rightarrow \infty} \varrho_q(Q, G_n). \quad (45)$$

Let $Q = \sum_{i \in P} \lambda_i F_i + \sum_{j \in N} \lambda_j F_j$ with $\lambda_i > 0$ for $i \in P$ and $\lambda_j < 0$ for $j \in N$. Set $\ell_i := v(F_i)$ for all $i \in P \cup N$ and

$$\ell := \max \left\{ \frac{\ell_i}{2} : i \in P \cup N \right\}.$$

The exact reasoning as in the proof of [26, Lemma 2.3] then gives that

$$\varrho_q(Q, G_n) \geq \frac{\varrho_q(Q, G'_n)}{(x'/x)^\ell} - C \left(1 - \left(\frac{x}{x'} \right)^\ell \right) - o(1)$$

for $C := \sum_{j \in N} (-\lambda_j)(\ell_j + 1) \geq 0$. This together with 45 gives that

$$I_q(Q, (x, \vec{x}_0)) \geq \frac{I_q(Q, (x', \vec{x}_0))}{(x'/x)^\ell} - C \left(1 - \left(\frac{x}{x'} \right)^\ell \right).$$

Dividing by x^ℓ gives

$$\frac{I_q(Q, (x', \vec{x}_0))}{(x')^\ell} \leq \frac{I_q(Q, (x, \vec{x}_0))}{x^\ell} + C \left(\left(\frac{1}{x} \right)^\ell - \left(\frac{1}{x'} \right)^\ell \right),$$

as desired. □

We are almost ready to prove Theorem 4.3, but we need one more fact.

FACT 4.11. Let Q be a q -colored quantum graph.

(a) The feasible regions of Q and $-Q$ are symmetric to each other about the \vec{x} -axes.

Hence, $I_q(-Q, \vec{x}) = -i_q(Q, \vec{x})$ and $i_q(-Q, \vec{x}) = -I_q(Q, \vec{x})$ for all $x \in \Delta'_{q-1}$.

(b) As we can always permute the colors, given any $\sigma \in S_q$,

$$I_q(Q, \vec{x}) = I_q(\sigma Q, f(\sigma(\vec{x}, 1 - (x_1 + \cdots + x_{q-1})))) ,$$

$$i_q(Q, \vec{x}) = i_q(\sigma Q, f(\sigma(\vec{x}, 1 - (x_1 + \cdots + x_{q-1}))))$$

for all $x \in [0, 1]$ where $f(a_1, \dots, a_q) = (a_1, \dots, a_{q-1})$ and S_q acts on Q by correspondingly permuting the edge-colors of all its constituents.

We prove Theorem 4.3, our main result on the structure of the feasible region of q -colored complete graphs.

PROOF OF THEOREM 4.3. Proposition 4.9 directly implies that

$$\Omega_q(Q) = \{(\vec{x}, y) \in [0, 1]^{q-1} \times \mathbb{R} : i_q(Q, \vec{x}) \leq y \leq I_q(Q, \vec{x})\},$$

so it just remains to show that $i_q(Q, \vec{x})$ and $I_q(Q, \vec{x})$ are continuous and their partial derivatives exist almost everywhere.

First, we just show that $I_q(F, \vec{x})$ is continuous for any q -colored complete graph F on s vertices. Let $\epsilon > 0$ small and $\vec{x} \in \Delta'_{q-1}$ be fixed. Then, take $\alpha > 0$ small enough that

$$(1 + \alpha)^s < \frac{I_q(F, \vec{x})}{I_q(F, \vec{x}) - \epsilon}.$$

This is possible since $\lim_{\alpha \rightarrow 0} (1 + \alpha)^s = 1$ and $I_q(F, \vec{x}) / (I_q(F, \vec{x}) - \epsilon) > 1$. Now, we need to take $\delta > 0$ such that

$$\alpha_i = \sqrt{\frac{x_i}{x_i - \delta}} - 1 \leq \alpha$$

for all $i \in [q]$. Since q is finite, $\lim_{\delta \rightarrow 0} \sqrt{x_i/(x_i - \delta)} - 1 = 0$, and $\alpha > 0$, this is possible. Set $\alpha' := \max_{i \in [q]} \{\alpha_i\}$ and take any $\vec{x}' \in \Delta'_{q-1}$ such that $d(\vec{x}, \vec{x}') < \delta$. Thus, $|x_i - x'_i| < \delta$ for all $i \in [q]$. Let $(G_n)_{n=1}^\infty$ be a sequence of q -colored complete graphs that realizes $(\vec{x}, I_q(F, \vec{x}))$. Now, we create a new sequence of q -colored complete graphs $(G'_n)_{n=1}^\infty$ where we replace G_n with G'_n by adding $\alpha'n$ vertices to G_n and coloring the new edges so that G'_n has as close to \vec{x}' color proportions as possible. Since we added $\alpha'n$ more vertices, the proportion of edges of color i before coloring the new edges is now

$$\frac{(x_i + o_n(1)) \binom{n}{2}}{\binom{n+\alpha'n}{2}} = \frac{x_i}{(1 + \alpha_i)^2} + o_n(1).$$

Since $\alpha' \geq \alpha_i$ for all $i \in [q]$,

$$\frac{x_i}{(1 + \alpha_i)^2} + o_n(1) \leq x_i - \delta + o_n(1) \leq x'_i + o_n(1)$$

for all $i \in [q]$, so we can indeed color the additional edges so that G'_n has color density $x'_i + o_n(1)$ for each color $1 \leq i \leq q$. This tells us that $(G'_n)_{n=1}^\infty$ realizes the color densities \vec{x}' , so

$$I_q(F, \vec{x}') \geq \lim_{n \rightarrow \infty} \varrho_q(F, G'_n) \geq \lim_{n \rightarrow \infty} \frac{I_q(F, \vec{x}) \binom{n}{s}}{\binom{n+\alpha'n}{s}} = \frac{I_q(F, \vec{x})}{(1 + \alpha')^s} \geq \frac{I_q(F, \vec{x})}{(1 + \alpha)^s} > I_q(F, \vec{x}) - \epsilon$$

due to our choice of α .

We also need to show that we can take δ small enough so that $I_q(F, \vec{x}') < I_q(F, \vec{x}) + \epsilon$ for all $\vec{x}' \in \Delta'_{q-1}$ such that $d(\vec{x}, \vec{x}') < \delta$ in order to conclude that $|I_q(F, \vec{x}) - I_q(F, \vec{x}')| < \epsilon$. By contradiction, suppose that, for all choices of $\delta > 0$, we have $I_q(F, \vec{x}') \geq I_q(F, \vec{x}) + \epsilon$ for some \vec{x}' with $d(\vec{x}, \vec{x}') < \delta$. Taking in particular $\delta_n := 1/n$, let \vec{x}_n be such a vector for δ_n . Thus, we see that $\lim_{n \rightarrow \infty} \vec{x}_n = \vec{x}$. For a given n , take G_n an element of a sequence realizing $(\vec{x}_n, I_q(F, \vec{x}_n))$ such that $v(G_n) \geq n$ and $\varrho_q(F, G_n) \geq I_q(F, \vec{x}) + \epsilon - 1/n$. Since $\{\varrho_q(F, G_n)\}_{n=1}^\infty$ is an infinite sequence bounded below by 0 and above by 1, the Bolzano–Weierstrass theorem implies that there is a convergent subsequence $\{\varrho_q(F, G'_n)\}_{n=1}^\infty$. We see that $(G'_n)_{n=1}^\infty$ realizes $(\vec{x}, y) \in \Omega_q(F)$ where

$$y = \lim_{n \rightarrow \infty} \varrho_q(F, G_n) \geq \lim_{n \rightarrow \infty} I_q(F, \vec{x}) + \epsilon - \frac{1}{n} = I_q(F, \vec{x}) + \epsilon.$$

However, this contradicts the definition of $I_q(F, \vec{x})$, so we have reached a contradiction and $I_q(F, \vec{x})$ is continuous.

Let $Q = \sum_{i=1}^m \lambda_i F_i$ be some q -colored quantum graph. For each F_i , let $\delta_i > 0$ be such that $|I_q(F_i, \vec{x}), I_q(F_i, \vec{x}')| < \epsilon/(m\lambda_i)$ for all $\vec{x}' \in \Delta'_{q-1}$ such that $d(\vec{x}, \vec{x}') < \delta_i$. Then, take $\delta := \min_{i \in [q]} \{\delta_i\}$, so that

$$|I_q(Q, \vec{x}), I_q(Q, \vec{x}')| \leq \sum_{i=1}^m \lambda_i |I_q(F_i, \vec{x}), I_q(F_i, \vec{x}')| < \sum_{i=1}^m \frac{\epsilon}{m} = \epsilon.$$

Therefore, $I_q(Q, \vec{x})$ is continuous.

Further, we know by Fact 4.11(a) that $i_q(Q, \vec{x}) = -I_q(-Q, \vec{x})$, so $i_q(Q, \vec{x})$ is also continuous.

It remains to prove that the partial derivatives of $i_q(Q, \vec{x})$ and $I_q(Q, \vec{x})$ exist almost everywhere. By Fact 4.11(a) and (b), it suffices to just show that the partial derivatives of $I_q(Q, \vec{x})$ with respect to x_1 exist almost everywhere. Let $\ell \geq 1$, $C \geq 0$ be the constants from Lemma 4.10. Then, the function $g : (0, 1] \rightarrow \mathbb{R}$ defined by $g(x) = I_q(Q, (x, \vec{x}_0)) + C/x^\ell$ is decreasing. Therefore, g is almost everywhere differentiable, and thus $I_q(Q, (x, \vec{x}_0))$ is as well for any fixed $\vec{x}_0 \in \Delta'_{q-2}$. \square

Subgraph density of edge-weighted graphs

This chapter is based on joint work with Dhruv Mubayi.

1. Background and results

The Kruskal–Katona theorem [19, 21] implies that, given a fixed number of vertices n and number of edges $m = \binom{a}{2} + b$ in a simple graph G for $b \in [0, a)$,

$$N(K_k, G) \leq \binom{a}{k} + \binom{b}{k-1}.$$

This is achieved by the construction on the left in Figure 5.1.

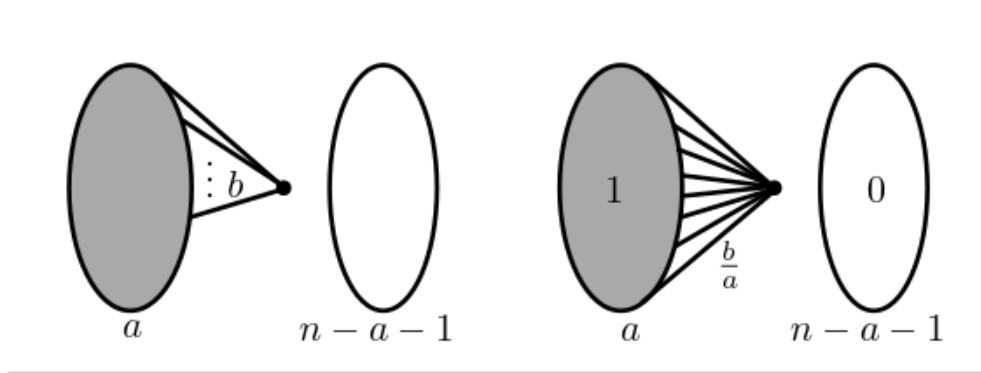


FIGURE 5.1. Weighted and un-weighted constructions.

In the edge-weighted setting, the construction on the right in Figure 5.1 has higher K_k -weight whenever $b \neq 0$. Nonetheless, we prove that the total K_k -weight of an edge-weighted graph G' on n vertices cannot be much higher than the maximum number of copies of K_k in a simple graph G .

THEOREM 5.1. *Let $k \geq 3$. Let G' be a $[0, 1]$ -edge-weighted graph on $n \geq 3$ vertices. Set*

$$m := \left\lfloor \sum_{e \in E(G')} w(e) \right\rfloor$$

Then there exists an n -vertex (un-weighted) graph G with m edges such that

$$N_w(K_k, G') - N(K_k, G) \leq \frac{2}{(k-4)!} \cdot n^{k-2}$$

when $k \geq 4$ and

$$N_w(K_3, G') - N(K_3, G) \leq 2n.$$

Theorem 5.1 says that we lose at most a constant factor of n^{k-2} cliques of order k by restricting to the un-weighted regime. Note that the number of copies of K_k in a graph typically has order n^k , so this difference is asymptotically negligible.

Further, the bound in Theorem 5.1 is tight. Let $a = \Theta(n)$ and $b = a/2$. The difference between the K_k -weight of the right edge-weighted graph construction of Figure 5.1 and the number of copies of K_k in the left graph construction of Figure 5.1 is

$$\binom{a}{k-1} \left(\frac{b}{a}\right)^{k-1} - \binom{b}{k-1}.$$

This simplifies to

$$\frac{1}{(k-1)!} \left(\frac{(a)_{k-1}}{a^{k-1}} \cdot b^{k-1} - (b)_{k-1} \right) = \frac{1}{(k-1)!} \left(\frac{(a)_{k-1}}{2^{k-1}} - \left(\frac{a}{2}\right)_{k-1} \right).$$

Multiplying out, we see that this can be written as $1/(2^{k-1}(k-1)!)$ times

$$\left(a^{k-1} - \frac{(k-1)(k-2)}{2} a^{k-2} + O(n^{k-3}) \right) - \left(a^{k-1} - \frac{(k-1)(2k-4)}{2} a^{k-2} + O(n^{k-3}) \right),$$

which in turn can be simplified to

$$\frac{(k-1)(k-2)}{2} a^{k-2} + O(n^{k-3}).$$

Therefore, $\binom{a}{k-1} (b/a)^{k-1} - \binom{b}{k-1}$ is equal to

$$\frac{1}{2^{k-1}(k-1)!} \left(\frac{(k-1)(k-2)}{2} a^{k-2} + O(n^{k-3}) \right) = \frac{1}{2^k(k-3)!} a^{k-2} + O(n^{k-3}),$$

which is positive and $\Theta(n^{k-2})$ since $a = \Theta(n)$. Therefore, the K_k -weight of the edge-weighted construction is $\Theta(n^{k-2})$ more than the number of copies of K_k in the simple graph construction, so the order of the bound in Theorem 5.1 is tight.

2. Proof of Theorem 5.1

If an edge of a $[0, 1]$ -edge-weighted graph is assigned a weight in $(0, 1)$, we call it a *fractional* edge. We prove Theorem 5.1 by changing the edge weights of a $[0, 1]$ -edge-weighted graph G' repeatedly until no fractional edges are left. Then we are left with a $\{0, 1\}$ -edge-weighted graph that corresponds to a simple graph G .

PROOF OF THEOREM 5.1. Let $k \geq 3$, let G' a $[0, 1]$ -edge-weighted graph on $n \geq 3$ vertices, and set

$$m := \left\lfloor \sum_{e \in E(G')} w(e) \right\rfloor.$$

We will need to be able to track the weight of all cliques sharing a common edge without including the weight of the edge itself or a second specified edge. For $e, f \in E(G')$, let

$$N_k(e, f) = \sum_{\substack{S \in \binom{[n]}{k} \\ e \subseteq S, f \notin S}} \prod_{e' \in \binom{S}{2} \setminus e} w(e').$$

Then, the sum of weights of all k -cliques that contain e , but not f , is $w(e)N_k(e, f)$.

Suppose there are two disjoint, fractional edges $e, f \in E(G')$ with weight $x := w(e)$ and $y := w(f)$. Assume without loss of generality that $N_k(e, f) \leq N_k(f, e)$. If we decrease $w(e)$ by $\ell \in [0, 1]$ and increase $w(f)$ by ℓ , then the sum of edge-weights is still m . Further, the weight of cliques containing e , but not f , decreases by $\ell N_k(e, f)$, and the weight of cliques including f , but not e , increases by $\ell N_k(f, e) \geq \ell N_k(e, f)$. It remains to track the change in weight of those cliques containing both e and f . If $k = 3$, no triangles contain both

disjoint edges. If $k \geq 4$, then there are at most $\binom{n-4}{k-4}$ many k -cliques containing both e and f , so the weight of these cliques decreases by at most $\binom{n-4}{k-4}$ since all k -clique weights are in $[0, 1]$. Therefore, overall, the change in edge-weights can cause a total decrease in k -clique weight of at most $\binom{n-4}{k-4}$. Choosing the maximum ℓ such that $w(e) - \ell, w(f) + \ell \in [0, 1]$, we repeat this process until no pairs of disjoint, fractional edges remain. Call the resulting edge-weighted graph G'' . Since there is at least one less fractional edge after each operation, and there are $\binom{n}{2}$ edges in the graph, we perform this operation at most $\binom{n}{2}$ times. Therefore, $N_w(K_k, G') - N_w(K_k, G'')$ is at most

$$\binom{n}{2} \binom{n-4}{k-4} \leq \frac{n^{k-2}}{2 \cdot (k-4)!}$$

for $k \geq 4$ and at most 0 for $k = 3$.

Note that all fractional edges in G'' form either a triangle or a star, so there are at most n remaining fractional edges. Suppose there are two fractional edges $e, f \in e(G'')$ with weight $x := w(e)$ and $y := w(f)$. Since all fractional edges are incident, e and f share a vertex. Assume that $N_k(e, f) \leq N_k(f, e)$, decrease $w(e)$ by $\ell \in [0, 1]$, and increase $w(f)$ by ℓ . Again, the weight of cliques containing e , but not f , decreases by $\ell N_k(e, f)$, and the weight of cliques including f , but not e , increases by $\ell N_k(f, e) \geq \ell N_k(e, f)$. There are at most $\binom{n-3}{k-3}$ many k -cliques containing both e and f , so the weight of these cliques decreases by at most $\binom{n-3}{k-3}$. Therefore, overall, the change in edge-weights can cause a total decrease in k -clique weight of at most $\binom{n-3}{k-3}$. Choosing the maximum ℓ such that $w(e) - \ell, w(f) + \ell \in [0, 1]$, we repeat this process until no pairs of disjoint, fractional edges remain. Call the resulting edge-weighted graph G''' . Since there is at least one less fractional edge after each operation, and there are at most n fractional edges in G'' , we perform this operation at most n times. Therefore, $N_w(K_k, G'') - N_w(K_k, G''')$ is at most

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

If $\sum_{e \in E(G')} w(e) \notin \mathbb{Z}$, then there is one remaining fractional edge in G''' . Set its weight to zero and call the resulting (un-weighted) graph G . Note that $N(K_k, G)$ differs from

$N_w(K_k, G''')$ by at most

$$\binom{n-2}{k-2} \leq \frac{n^{k-2}}{(k-2)!}.$$

Putting everything together, we see that

$$N_w(K_k, G') - N(K_k, G) \leq \left(\frac{1}{2} + \frac{1}{k-3} + \frac{1}{(k-2)(k-3)} \right) \cdot \frac{n^{k-2}}{(k-4)!} \leq \frac{2n^{k-2}}{(k-4)!}$$

for $k \geq 4$ and

$$N_w(K_3, G') - N(K_3, G) \leq \left(\frac{1}{0!} + \frac{1}{1!} \right) n = 2n,$$

as desired. □

CHAPTER 6

Maximum star densities

This chapter is based on joint work with Dhruv Mubayi in [8], published in the SIAM Journal on Discrete Mathematics.

1. Background and results

Reiher and Wagner [37] make great use of the theory of graphons for their work on star densities. While we considered an edge-coloring variant of the standard graphon in the proof of Proposition 4.9 in Chapter 4, we go into more detail here. Formally, a *graphon* W is a symmetric, measurable function $W : [0, 1]^2 \rightarrow [0, 1]$ (see Lovász [27]). For a graphon W , let $d_W(x) = \int_0^1 W(x, y) dy$ be the *degree* of x in W and let

$$t(\cdot, W) = \int_{[0,1]^2} W(x, y) dx dy$$

be the *density* of W . We observe that

$$\sum_{v \in V(G)} d(v)^k = N(S_k, G) + O(n^k)$$

in graphs. Correspondingly, let

$$t(S_k, W) = \int_0^1 d_W^k(x) dx$$

be the *homomorphism density* of S_k in W .

THEOREM 6.1 (Reiher–Wagner [37]). *Let W be a graphon and let k be a positive integer. Set $\gamma = t(\cdot, W)$ and $\eta = 1 - \sqrt{1 - \gamma}$. Then*

$$t(S_k, W) \leq \max\{\gamma^{(k+1)/2}, \eta + (1 - \eta)\eta^k\}.$$

By the general theory of graphons, Theorem 6.1 gives the same upper bound for $I(S_k, \gamma)$. There appears to be an error in the proof of [37, Proposition 3.7], which is necessary for the proof of Theorem 6.1. We correct the proof.

We also consider the problem of maximizing the number of copies of S_k in graph G and its complement \overline{G} simultaneously. We expect the number of copies of S_k in both G and \overline{G} to be $\Theta(n^{k+1})$, so any terms that are $O(n^k)$ or $o(n^{k+1})$ can be viewed as error terms. We prove the following result for $k = 2$.

THEOREM 6.2. *For all graphs G on n vertices with edge density γ , we have that*

$$N(S_2, G) + N(S_2, \overline{G}) \leq 3 \binom{n}{3} \max \{2\gamma^{3/2} - 2\gamma + 1, 2(1 - \gamma)^{3/2} - 2(1 - \gamma) + 1\} + o(n^3).$$

This bound is tight. When $2\gamma^{3/2} - 2\gamma + 1$ is the maximum, the optimal construction asymptotically is the clique. Otherwise, the optimal construction is the anticlique. Note that these are the same constructions as for Theorem 6.1.

For $k = 3$, we observe that

$$N(S_3, G) + N(S_3, \overline{G}) = \sum_{v \in V(G)} (d(v)^3 + (n - d(v))^3) + O(n^3).$$

This motivates our next result.

THEOREM 6.3. *For all graphs G on n vertices with edge density γ , we have that*

$$\sum_{v \in V(G)} (d(v)^3 + (n - d(v))^3) \leq \frac{n^4}{6} \max \{3\gamma^{3/2} - 3\gamma + 1, 3(1 - \gamma)^{3/2} - 3(1 - \gamma) + 1\} + o(n^4).$$

Theorem 6.3 is also tight, and the constructions are again the clique and anticlique.

We prove Theorem 6.1 in Section 2. We prove Theorems 6.2 and 6.3 in Section 3.

2. Maximum star densities: correcting an error

In this section, we will correct the proof of Theorem 6.1 from [37]. This involves defining a new parameter (called $T(W)$) on graphons that we will optimize. Nevertheless, many parts

of the argument are identical to those in [37], and we will indicate when this is the case in various lemmas, claims and propositions.

Given a measurable function $F : [0, 1] \rightarrow \mathbb{R}$, a graphon W , and $\gamma \in [0, 1]$, let

$$D(F, W) := \int_0^1 F(d_W(x)) dx,$$

$$\text{MAX}(\gamma, F) := \max\{(1 - \sqrt{\gamma})F(0) + \sqrt{\gamma}F(\sqrt{\gamma}), (1 - \eta)F(\eta) + \eta F(1)\},$$

where $\eta = 1 - \sqrt{1 - \gamma}$. A measurable function $F : [0, 1] \rightarrow \mathbb{R}$ is *good* for W if

$$D(F, W) \leq \text{MAX}(\gamma, F),$$

where $\gamma = t(\cdot | \cdot, W)$ and $\eta = 1 - \sqrt{1 - \gamma}$; F is *bad* for W if it is not good for W . A collection of measurable functions is good (bad) for W if all its members are good (bad) for W . In [37], a set \mathcal{C} of twice differentiable convex functions $F : [0, 1] \rightarrow \mathbb{R}$ satisfying certain conditions is defined. We do not need the details of the conditions here, so we do not state them, but we note that \mathcal{C} contains the function $F(x) = x^k$.

The error in [37] appears in the proof of the following proposition.

PROPOSITION 6.4 (corresponds to [37, Proposition 3.7]). *\mathcal{C} is good for all step graphons.*

To prove Proposition 6.4, we introduce the following more refined notion of “good.” For any $\delta > 0$, say that $F \in \mathcal{C}$ is δ -good for a graphon W if

$$D(F, W) < \text{MAX}(\gamma, F) + \delta$$

for $\gamma = t(\cdot | \cdot, W)$ and $\eta = 1 - \sqrt{1 - \gamma}$. Say that F is δ -bad for W if it is not δ -good for W .

We next show that several lemmas from [37] still apply when we replace “good” with “ δ -good.” The proofs are almost exact copies of those in [37], and are given in the Appendix.

LEMMA 6.5 (corresponds to [37, Lemma 3.2]). *If all functions in \mathcal{C} are δ -good for a graphon W , then the same is true for the graphon $1 - W$.*

Given a graphon W and a real number $\lambda \in [0, 1]$, let $[\lambda, W]$ be the graphon satisfying

$$[\lambda, W](x, y) = \begin{cases} 0 & \text{if } 0 \leq x < \lambda \text{ or } 0 \leq y < \lambda, \\ W\left(\frac{x-\lambda}{1-\lambda}, \frac{y-\lambda}{1-\lambda}\right) & \text{otherwise.} \end{cases}$$

LEMMA 6.6 (corresponds to [37, Lemma 3.5]). *If $\lambda \in [0, 1]$ and the graphon W has the property that all functions in \mathcal{C} are δ -good for it, then the same applies to $[\lambda, W]$.*

Similarly, let $[W, \lambda]$ be the graphon satisfying

$$[W, \lambda](x, y) = \begin{cases} W\left(\frac{x}{1-\lambda}, \frac{y}{1-\lambda}\right) & \text{if } 0 \leq x \leq 1 - \lambda \text{ and } 0 \leq y \leq 1 - \lambda, \\ 1 & \text{otherwise.} \end{cases}$$

The next lemma follows from the previous two lemmas and the observation that $[W, \lambda]$ is isomorphic to $1 - [\lambda, 1 - W]$.

LEMMA 6.7 (corresponds to [37, Lemma 3.6]). *If all functions in \mathcal{C} are δ -good for the graphon W and $\lambda \in [0, 1]$, then all functions in \mathcal{C} are good for $[W, \lambda]$ as well.*

We now prove Proposition 6.4. Let \mathcal{S} be the collection of all step graphons and let \mathcal{S}_i be the collection of all step graphons with i parts.

PROOF OF PROPOSITION 6.4. Suppose for a contradiction that there exists $W' \in \mathcal{S}$ and $F' \in \mathcal{C}$ such that F' is bad for W' . This means that $D(F', W') > \text{MAX}(\gamma, W')$. Then there exists $\delta > 0$ such that $D(F', W') \geq \text{MAX}(\gamma, W') + \delta$. In other words, F' is δ -bad for W' . Let

$$\mathcal{A} = \mathcal{A}(\delta) := \{(F, W) \in \mathcal{C} \times \mathcal{G} : F \text{ is } \delta\text{-bad for } W\}.$$

Note that $(F', W') \in \mathcal{A}$, so $\mathcal{A} \neq \emptyset$. Partition \mathcal{A} into $\bigcup_{i=1}^{\infty} \mathcal{A}_i$, where

$$\mathcal{A}_i := \{(F, W) \in \mathcal{C} \times \mathcal{G}_i : F \text{ is } \delta\text{-bad for } W\}.$$

Let k be the smallest integer such that $\mathcal{A}_k \neq \emptyset$. As all convex functions are good for all constant graphons (see [37, Observation 2.1]), we have $k \geq 2$. Pick $F \in \mathcal{C}$ such that

$(F, W) \in \mathcal{A}_k$ for some W . Let

$$\mathcal{W} := \{W : (F, W) \in \mathcal{A}_k\}.$$

Let μ be the Lebesgue measure on \mathbb{R} . Each $W \in \mathcal{W}$ is a step function with respect to a partition $\mathcal{P} = \{P_1, \dots, P_k\}$ of the unit interval with $\alpha_i = \mu(P_i)$ for all $i \in [k]$ and β_{ij} the value attained by W on $P_i \times P_j$ for $i, j \in [k]$. By the choice of k , we deduce $\alpha_1, \dots, \alpha_k > 0$ for all $W \in \mathcal{W}$. Define

$$T(W) := \sum_{i=1}^k \alpha_i (d_W(i))^2,$$

where

$$d_W(i) = \int_0^1 W(i, y) dy = \sum_{j=1}^k \alpha_j \beta_{ij}$$

for all $i \in [k]$.

CLAIM 6.8. $T := \sup_{W \in \mathcal{W}} T(W) = \max_{W \in \mathcal{W}} T(W)$.

PROOF. First we note that each \mathcal{S}_ℓ has a natural compact topology corresponding to convergence of all parameters α_i and β_{ij} .

In particular, the space is homeomorphic to $\Delta^{\ell-1} \times [0, 1]^{\binom{\ell+1}{2}}$, where $\Delta^{\ell-1}$ is the standard $(\ell-1)$ -simplex since $\sum_{i=1}^{\ell} \alpha_i = 1$. In this topology, it is straightforward to check that $T(W)$, $D(F, W)$ and $\text{MAX}(\gamma_W, F)$ are continuous functions of W provided that F is continuous and that the property of being δ -bad is preserved under taking the limit. (This is why we needed to define the property of being δ -bad, as this would not be true if we replaced “ δ -bad” with just “bad.”) In particular, this implies that the set \mathcal{W} is closed, hence compact. Thus, $T(\mathcal{W})$ must attain its maximum in the compact set \mathcal{W} . \square

Fix $W \in \mathcal{W}$ with partition $\mathcal{P} = \{P_1, \dots, P_k\}$ and parameters α_i and $\beta_{ij} \in [0, 1]$ for all $i, j \in [k]$ such that $T(W) = T$. Recall that by the minimality of k , we know that $\alpha_i \in (0, 1)$. By the definition of \mathcal{W} , F is δ -bad for W . Set $d_i := d_W(i)$. To obtain the necessary contradiction to complete the proof of the proposition, we will show that F is δ -good for W ,

i.e. that

$$D(F, W) = \int_0^1 F(d_W(x)) dx = \sum_{i=1}^k \alpha_i F(d_i) < \text{MAX}(\gamma, F) + \delta,$$

for $\gamma = t(\cdot, W) = \sum_{i=1}^k \alpha_i d_i$. Without loss of generality, we may assume $d_1 \leq d_2 \leq \dots \leq d_k$.

CLAIM 6.9 (corresponds to [37, Claim 3.8]). *If $1 \leq r < s \leq k$ and $\beta_{ir} > 0$, then $\beta_{is} = 1$.*

PROOF OF CLAIM 6.9. Suppose, for contradiction, that $\beta_{ir} > 0$ and $\beta_{is} < 1$. Define $Q \in \mathcal{S}_k$ with the same partition \mathcal{P} as follows: let δ_{ij} denote the Kronecker delta, and set, for $x \in P_m$ and $y \in P_n$,

$$Q(x, y) = \begin{cases} -(1 + \delta_{ir})\alpha_s & \text{if } \{m, n\} = \{i, r\}, \\ (1 + \delta_{is})\alpha_r & \text{if } \{m, n\} = \{i, s\}, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\varepsilon \geq 0$ be maximal such that $W_\varepsilon = W + \varepsilon Q$ still satisfies $W_\varepsilon(x, y) \in [0, 1]$ for all $x, y \in [0, 1]$. By our assumptions on β_{ir} and β_{is} , we know that $\varepsilon > 0$.

For all $j \in [k]$, let d'_j denote the value attained by $d_{W_\varepsilon}(x)$ for all $x \in P_j$. We have

$$d'_r - d_r = -(1 + \delta_{ir})\alpha_i\alpha_s\varepsilon + \delta_{ir}(1 + \delta_{is})\alpha_i\alpha_s\varepsilon = -\alpha_i\alpha_s\varepsilon,$$

and similarly $d'_s - d_s = \alpha_i\alpha_r\varepsilon$. Further, if $i \notin \{r, s\}$, then

$$d'_i - d_i = \alpha_r(\beta_{ir} - \alpha_s) + \alpha_s(\beta_{is} + \alpha_r) - (\alpha_r\beta_{ir} + \alpha_s\beta_{is}) = 0.$$

Therefore $d'_j = d_j$ for all $j \notin \{r, s\}$. Consequently,

$$\begin{aligned} T(W_\varepsilon) - T(W) &= \sum_{j=1}^k \alpha_j (d'_j)^2 - \sum_{j=1}^k \alpha_j d_j^2 = \alpha_r (d_r - \alpha_i\alpha_s\varepsilon)^2 - \alpha_r d_r^2 + \alpha_s (d_s + \alpha_i\alpha_r\varepsilon)^2 - \alpha_s d_s^2 \\ &= 2\varepsilon\alpha_i\alpha_r\alpha_s(d_s - d_r) + \varepsilon^2\alpha_i^2\alpha_r\alpha_s(\alpha_r + \alpha_s) \\ &\geq \varepsilon^2\alpha_i^2\alpha_r\alpha_s(\alpha_r + \alpha_s) > 0 \end{aligned}$$

since $\alpha_i, \alpha_r, \alpha_s, \varepsilon > 0$ and $d_s \geq d_r$ by assumption. Thus $T(W_\varepsilon) > T(W)$, and clearly $W_\varepsilon \in \mathcal{S}_k$, so we conclude by the choice of W that $W_\varepsilon \notin \mathcal{W}$. To complete the proof, we will now obtain the contradiction $W_\varepsilon \in \mathcal{W}$ by showing that F is δ -bad for W_ε . First, observe that

$$\int_{[0,1]^2} Q(x, y) \, dx dy = \alpha_i \alpha_r \alpha_s ((1 + \delta_{is})(2 - \delta_{is}) - (1 + \delta_{ir})(2 - \delta_{ir})) = 0.$$

This implies that $t(\cdot, W_\varepsilon) = t(\cdot, W) = \gamma$. Next,

$$\begin{aligned} D(F, W_\varepsilon) - D(F, W) &= \int_0^1 F(d_{W_\varepsilon}(x)) \, dx - \int_0^1 F(d_W(x)) \, dx \\ &= \sum_{j=1}^k \alpha_j (F(d'_j) - F(d_j)) \\ &= \alpha_s (F(d_s + \alpha_i \alpha_r \varepsilon) - F(d_s)) + \alpha_r (F(d_r - \alpha_i \alpha_s \varepsilon) - F(d_r)) \\ &\geq \alpha_i \alpha_r \alpha_s \varepsilon (F'(d_s) - F'(d_r)) \geq 0 \end{aligned}$$

by convexity of F and $d_s \geq d_r$. Since F is δ -bad for W , we obtain

$$D(F, W_\varepsilon) \geq D(F, W) \geq \text{MAX}(\gamma, F) + \delta.$$

This shows that F is δ -bad for W_ε and completes the proof of the claim. \square

CLAIM 6.10 (corresponds to [37, Claim 3.9]). $\beta_{1k} > 0$.

PROOF. If this does not hold, then $\beta_{1k} = 0$ and Claim 6.9 imply that $\beta_{1i} = 0$ for all $i \in [k-1]$. It follows that there exists $W' \in \mathcal{S}_{k-1}$ such that W is isomorphic to $[\alpha_1, W']$. Due to the minimality of k all functions in \mathcal{C} are δ -good for W' and by Lemma 6.6 the same applies to the graphon W , contrary to its choice. \square

CLAIM 6.11 (corresponds to [37, Claim 3.10]). $\beta_{1k} < 1$.

PROOF. Suppose for contradiction that $\beta_{1k} = 1$. Then $\beta_{k1} = \beta_{1k} > 0$, and Claim 6.9 implies $\beta_{ki} = 1$ for all $1 \leq i \leq k$. So some $W' \in \mathcal{S}_{k-1}$ has the property that $[W', \alpha_k]$ is isomorphic to W . This implies that all functions in \mathcal{C} are δ -good for W' and then Lemma 6.7 implies that all functions in \mathcal{C} are δ -good for W , contradiction. \square

By Claims 6.10 and 6.11, we have $0 < \beta_{1k} < 1$. Moreover, by Claim 6.9, $\beta_{1i} = 0$ for all $i \in [k-1]$ and $\beta_{jk} = 1$ for all j with $2 \leq j \leq k$. Divide P_k into two measurable subsets Q_k and Q_{k+1} satisfying $\mu(Q_k) = (1 - \beta_{1k})\alpha_k$ and, consequently, $\mu(Q_{k+1}) = \beta_{1k}\alpha_k$. Set $Q_i = P_i$ for $i \in [k-1]$ and $\mathcal{Q} = \{Q_1, \dots, Q_{k+1}\}$. Let W' be the step graphon with respect to \mathcal{Q} defined as follows: for $x \in Q_i$ and $y \in Q_j$,

$$W'(x, y) = \begin{cases} \beta_{ij} & \text{if } 2 \leq i \leq k \text{ and } 2 \leq j \leq k, \\ 0 & \text{if } i = 1 \text{ and } j \in [k] \text{ or vice versa,} \\ 1 & \text{if } i = k+1 \text{ or } j = k+1. \end{cases}$$

By the last two clauses W' is isomorphic to a graphon of the form $[[\alpha_1/(1 - \beta_{1k}\alpha_k), W''], \beta_{1k}\alpha_k]$ for some graphon W'' , and by the first clause $W'' \in \mathcal{G}_{k-1}$. So Lemmas 6.6 and 6.7 show that F is δ -good for W' . Set $d' := d_k - \alpha_1\beta_{1k}$ and $d'' := d_k + \alpha_1(1 - \beta_{1k})$. Since $t(\cdot, W') = t(\cdot, W) = \gamma$ and $(1 - \beta_{1k})d' + \beta_{1k}d'' = d_k$, Jensen's Inequality implies that

$$D(F, W) = \sum_{i=1}^k \alpha_i F(d_i) \leq \sum_{i=1}^{k-1} \alpha_i F(d_i) + \alpha_k((1 - \beta_{1k})F(d') + \beta_{1k}F(d'')) = D(F, W').$$

Therefore, $D(F, W) \leq D(F, W') < \text{MAX}(\gamma, F) + \delta$ so F is δ -good for W , a contradiction. \square

The proof of Theorem 6.1 follows from Proposition 6.4 exactly as in [37].

The error we correct is in the proof of [37, Claim 3.8] (which corresponds to Claim 6.9 in this thesis). In [37], T is defined as the number of pairs $(i, j) \in [k]^2$ for which $\beta_{ij} \in \{0, 1\}$ and $W \in \mathcal{S}_k$ is chosen to maximize T . Then [37, Claim 3.8] states that $\beta_{ir} > 0$ and $\beta_{is} < 1$ together imply that W_ε is 0 or 1 on at least $T+1$ of the sets $P_i \times P_j$ by construction. However, this is not true if $\beta_{ir} = 1$ and $\beta_{is} = 0$. For example, consider the graphon W defined on the partition $P_1 = [0, 2/5)$, $P_2 = [2/5, 4/5)$, $P_3 = [4/5, 1]$ by

$$W(x, y) = \begin{cases} 1 & \text{if } (x, y) \in (P_1 \times P_3) \cup (P_2 \times P_2) \cup (P_2 \times P_3) \\ 0 & \text{if } (x, y) \in (P_1 \times P_1) \cup (P_1 \times P_2) \cup (P_3 \times P_3) \end{cases}$$

(see Figure 6.1). We see that $d_1 = 1/5$, $d_2 = 3/5$, and $d_3 = 4/5$, so $d_1 \leq d_2 \leq d_3$ is satisfied. Setting $i = 3$, $r = 2$, and $s = 3$, we have that $\beta_{ir} = 1 > 0$ and $\beta_{is} = 0 < 1$. However, all entries are already 0 or 1 in W , so W_ε cannot equal 0 or 1 on more sets than W . Note that any step graphon with parts ordered by degree such that $\beta_{ir} = 1$ and $\beta_{is} = 0$ for some i and $r < s$ is also a counterexample; the other entries need not equal 0 or 1 as in this example.

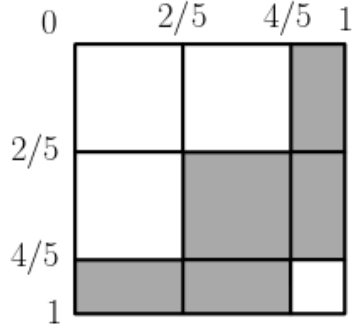


FIGURE 6.1. A counterexample to [37, Claim 3.8] (shaded squares have β value 1).

3. Maximum star density in a graph and its complement

We first prove Theorem 6.2.

PROOF OF THEOREM 6.2. Counting the number of times each induced subgraph of G on 3 vertices contributes to $N(S_2, G) + N(S_2, \overline{G})$, we see that

$$N(S_2, G) + N(S_2, \overline{G}) = 3N_{\text{ind}}(K_3, G) + N_{\text{ind}}(S_2, G) + N_{\text{ind}}(\overline{S_2}, G) + 3N_{\text{ind}}(\overline{K_3}, G).$$

Since every 3-vertex subgraph of G is one of K_3 , S_2 , $\overline{S_2}$, or $\overline{K_3}$, it is also true that

$$N_{\text{ind}}(K_3, G) + N_{\text{ind}}(S_2, G) + N_{\text{ind}}(\overline{S_2}, G) + N_{\text{ind}}(\overline{K_3}, G) = \binom{n}{3}.$$

Thus,

$$N(S_2, G) + N(S_2, \overline{G}) = \binom{n}{3} + 2(N_{\text{ind}}(K_3, G) + N_{\text{ind}}(\overline{K_3}, G)).$$

Liu, Mubayi, and Reiher [26, Theorem 1.7] combined results of Goodman [15] and Olpp [33] to prove that

$$N_{\text{ind}}(K_3, G) + N_{\text{ind}}(\overline{K_3}, G) \leq \binom{n}{3} (1 - \min\{\gamma - \gamma^{3/2}, (1 - \gamma) - (1 - \gamma)^{3/2}\}) + o(n^3).$$

Applying this bound and simplifying, we see that

$$N(S_2, G) + N(S_2, \overline{G}) \leq 3 \binom{n}{3} \max\{2\gamma^{3/2} - 2\gamma + 1, 2(1 - \gamma)^{3/2} - 2(1 - \gamma) + 1\} + o(n^3),$$

as desired. \square

To prove Theorem 6.3, we make use of graphons.

PROOF OF THEOREM 6.3. Recall that $x^k \in \mathcal{C}$. We see further that

$$x^3 + (1 - x)^3 = 3x^2 - 3x + 1 \in \mathcal{C}$$

by [37, Lemma 3.1], [37, Proposition 4.4], and the fact that $F \in \mathcal{C}$ implies that $cF \in \mathcal{C}$ for any positive $c \in \mathcal{C}$ by the definition of \mathcal{C} . Proposition 6.4 gives that $x^3 + (1 - x)^3$ is good for all step graphons. As step graphons are dense in the space of all graphons (see the proof of [37, Theorem 1.3]), this implies for all graphons W that

$$\begin{aligned} \int_0^1 d_W(x)^3 + (1 - d_W(x))^3 dx &\leq \max\{(1 - \sqrt{\gamma})F(0) + \sqrt{\gamma}F(\sqrt{\gamma}), (1 - \eta)F(\eta) + \eta F(1)\} \\ &= \max\left\{1 - \sqrt{\gamma} + \sqrt{\gamma}(\gamma^{3/2} + (1 - \sqrt{\gamma})^3), \right. \\ &\quad \left. \sqrt{1 - \gamma} \left((1 - \sqrt{1 - \gamma})^3 + (1 - \gamma)^{3/2} \right) + 1 - \sqrt{1 - \gamma} \right\} \\ &= \max\{3\gamma^{3/2} - 3\gamma + 1, 3(1 - \gamma)^{3/2} - 3(1 - \gamma) + 1\} \end{aligned}$$

where $\eta = 1 - \sqrt{1 - \gamma}$. By the general theory of graphons, this implies that

$$\sum_{v \in V(G)} (d(v)^3 + (n - d(v))^3) \leq \frac{4n^4}{4!} \max\{3\gamma^{3/2} - 3\gamma + 1, 3(1 - \gamma)^{3/2} - 3(1 - \gamma) + 1\} + o(n^4)$$

for all graphs G on n vertices with edge density γ . \square

CHAPTER 7

Subgraph density of ordered graphs

This chapter is based on joint work with Dhruv Mubayi in [8], published in the SIAM Journal on Discrete Mathematics.

1. Background and results

Recall that an ordered graph $G = (V, E)$ is a graph with a total order on V . We will let $V = [n] := \{1, \dots, n\}$ with the natural ordering. For F an ordered graph on $[s]$, further recall that $N_{\text{ord}}(F, G)$ is the number of subsets $\{v_1, \dots, v_s\} \subseteq [n]$ with $v_1 < v_2 < \dots < v_s$ such that $v_i v_j \in E(G[v_1, \dots, v_s])$ whenever $ij \in E(F)$. We consider $N_{\text{ord}}(F, G)$ when F is an ordered star. We will show that the following constructions achieve the maximum and minimum number of copies of certain ordered stars, in particular when the node is either the smallest or greatest vertex.

CONSTRUCTION 7.1. For positive integers n and $m \leq \binom{n}{2}$, let a be the largest integer such that $f(n, a) := \binom{a}{2} + a(n - a) \leq m$. As $f(n, n) = \binom{n}{2}$, we have $0 \leq a \leq n$. Set $b = m - f(n, a)$. Since $f(n, a + 1) - f(n, a) = n - a - 1$, we conclude that $0 \leq b < n - a - 1$. Let $S_L(n, m)$ be the ordered graph with vertex set $[n]$ and edge set

$$\{vw : v \in [a], w \in [n]\} \cup \{\{a + 1, j\} : a + 2 \leq j \leq a + b + 1\}.$$

In words, $S_L(n, m)$ comprises a complete graph on $[a]$, and in addition has all edges between $[a]$ and $[n] \setminus [a]$ and b edges between $a + 1$ and the b smallest vertices in $[n] \setminus [a + 1]$ (see Figure 7.1). Let $S_R(n, m)$ be defined as $S_L(n, m)$, but where the total order on the vertices is reversed.

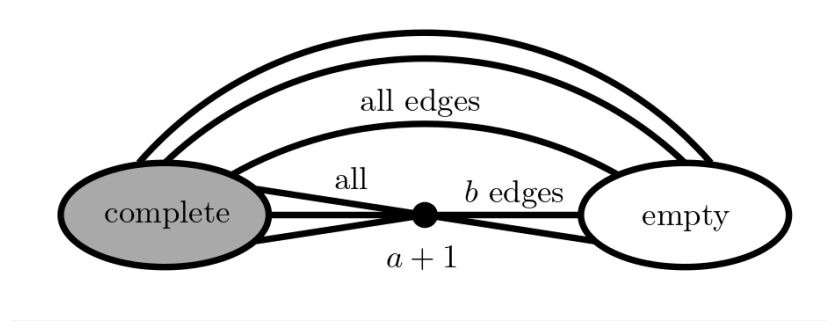


FIGURE 7.1. $S_L(n, m)$.

Formally, $S_L(k) := S_L(k + 1, k)$ is the ordered *left star* and $S_R(k) := S_R(k + 1, k)$ is the ordered *right star*. Note that $S_L(k)$ has $a = 1$ and $b = 0$ (see Figure 7.2).

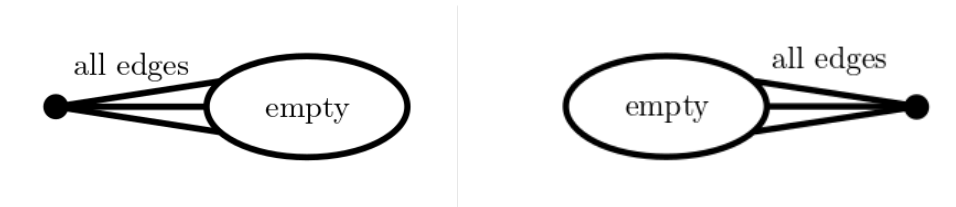


FIGURE 7.2. $S_L(k)$ and $S_R(k)$.

Our main result for ordered graphs is the following theorem.

THEOREM 7.2. *Let G be an ordered graph with vertex set $[n]$ and m edges. Then*

$$N_{\text{ord}}(S_L(k), S_R(n, m)) \leq N_{\text{ord}}(S_L(k), G) \leq N_{\text{ord}}(S_L(k), S_L(n, m)).$$

Theorem 7.2 implies similar results for $S_R(k)$. There are, up to obvious symmetries, $\lceil (k + 1)/2 \rceil$ different ordered stars with k edges and, apart from $S_L(k)$, it remains open to prove analogous results to Theorem 7.2 for them. Indeed, obtaining sharp bounds for these stars seems nontrivial. We address the first open case when $k = 2$.

Let M be the ordered graph with vertex set $[3]$ and edge set $\{12, 23\}$. It seems very difficult to obtain exact results for $N_{\text{ord}}(M, G)$ so we consider asymptotic growth rates. Let

$$\varrho_{\text{ord}}(F, G) := \frac{N_{\text{ord}}(F, G)}{\binom{n}{s}}$$

be the *density of F in G* . Since the vertices are ordered, each s -tuple of vertices can contribute at most one copy of F so $\varrho_{\text{ord}}(F, G) \in [0, 1]$. Let $(G_n) := (G_n)_{n=1}^{\infty}$ be a sequence of ordered graphs with $\lim_{n \rightarrow \infty} |V(G_n)| = \infty$. The sequence (G_n) is *F -good* if both $\lim_{n \rightarrow \infty} \varrho(G_n)$ and $\lim_{n \rightarrow \infty} \varrho_{\text{ord}}(F, G_n)$ exist. In this case, we set $x := \lim_{n \rightarrow \infty} \varrho(G_n)$ and $y := \lim_{n \rightarrow \infty} \varrho_{\text{ord}}(F, G_n)$ and say that (G_n) *realizes (x, y)* . Define

$$I_{\text{ord}}(F, x) := \sup\{y : (x, y) \in [0, 1] \text{ is realized by some } F\text{-good } (G_n)\},$$

$$i_{\text{ord}}(F, x) := \inf\{y : (x, y) \in [0, 1] \text{ is realized by some } F\text{-good } (G_n)\}.$$

We first give a construction that achieves $i_{\text{ord}}(, x)$. As in [26], for any integers $n \geq k \geq 2$ and real $x \in (\frac{k-2}{k-1}, \frac{k-1}{k}]$, let $H^*(n, x)$ be the complete k -partite graph on n vertices with parts V_1, \dots, V_k of sizes $|V_1| = \dots = |V_{k-1}| = \lfloor \alpha_k n \rfloor$ and $|V_k| = n - (k-1)\lfloor \alpha_k n \rfloor$, where

$$\alpha_k = \frac{1}{k} \left(1 + \sqrt{1 - \frac{k}{k-1}x} \right).$$

It is simple to check that $\lim_{n \rightarrow \infty} \varrho(H^*(n, x)) = x$. We define

$$g_3(x) := \lim_{n \rightarrow \infty} \frac{N(K_3, H^*(n, x))}{\binom{n}{3}}.$$

Let $H'(n, x)$ be an ordered graph obtained from $H^*(n, x)$ with any vertex ordering for which $u < v$ whenever $u \in V_i, v \in V_j$ and $i < j$. Then $N_{\text{ord}}(M, H'(n, x)) = N_{\text{ord}}(K_3, H^*(n, x))$ and this implies that

$$i_{\text{ord}}(M, x) \leq g_3(x).$$

Recall the definition of $i_{\text{ind}}(F, x)$ from Section 2 of Chapter 4. Lovász and Simonovits [28] conjectured that $g_3(x)$ is the minimal density of K_3 in simple graphs, and this was proven by Razborov.

THEOREM 7.3 (Razborov [35, Theorem 3.1]). *For all $x \in [0, 1]$,*

$$i_{\text{ind}}(K_3, x) = g_3(x).$$

As $i_{\text{ind}}(K_3, x) = i_{\text{ord}}(K_3, x)$ due to the structure of K_3 , we have that $i_{\text{ord}}(K_3, x) = g_3(x)$. Note that M is a subgraph of the ordered K_3 , so $i_{\text{ord}}(K_3, x) \leq i_{\text{ord}}(M, x)$. Therefore,

$$g_3(x) = i_{\text{ord}}(K_3, x) \leq i_{\text{ord}}(M, x)$$

and we conclude that

$$i_{\text{ord}}(M, x) = i_{\text{ord}}(K_3, x) = g_3(x).$$

Determining $I_{\text{ord}}(M, x)$ appears to be more difficult.

CONSTRUCTION 7.4. We construct a sequence of graphs $(P(n, x))_{n=1}^{\infty}$ for any $x \in [0, 1]$. For each n , define $P(n, x)$ to be the ordered graph on $[n] = A \sqcup B \sqcup C$, where $|B| = \lfloor n(1 - \sqrt{1-x}) \rfloor$, $\|A\| - \|C\| \leq 1$, and $a < b < c$ for all $a \in A, b \in B, c \in C$ with edge set

$$\{ab : a \in A, b \in B\} \cup \{b_1b_2 : b_1, b_2 \in B\} \cup \{bc : b \in B, c \in C\}.$$

A short calculation shows that $\lim_{n \rightarrow \infty} \varrho(P(n, x)) = x$ and

$$\lim_{n \rightarrow \infty} \varrho_{\text{ord}}(M, P(n, x)) = \frac{\eta(3 - \eta^2)}{2}$$

for $\eta = 1 - \sqrt{1-x}$.

CONSTRUCTION 7.5. We construct a sequence of graphs $(Q(n, x))_{n=1}^{\infty}$ for any $x \in [0, 1]$. For each n , define $Q(n, x)$ to be the ordered graph on $[n]$ with edge set

$$\{ij : j - i \leq \lfloor (1 - \sqrt{1-x})n \rfloor\}.$$

A short calculation shows that $\lim_{n \rightarrow \infty} \varrho(Q(n, x)) = x$ and that for $\eta = 1 - \sqrt{1-x}$,

$$\lim_{n \rightarrow \infty} \varrho_{\text{ord}}(M, Q(n, x)) = \begin{cases} 6\eta^3 + 6(1 - 2\eta)\eta^2 & \text{if } \eta \leq 1/2, \\ 2\eta^3 - 6\eta^2 + 6\eta - 1 & \text{if } \eta \geq 1/2. \end{cases}$$

Constructions 7.4 and 7.5 show that for $\eta = 1 - \sqrt{1-x}$,

$$I_{\text{ord}}(M, x) \geq \max \left\{ \lim_{n \rightarrow \infty} \varrho_{\text{ord}}(M, P(n, x)), \lim_{n \rightarrow \infty} \varrho_{\text{ord}}(M, Q(n, x)) \right\}.$$

It is an interesting open problem to determine if the inequality is sharp.

PROBLEM 7.6. *Is*

$$I_{\text{ord}}(M, x) = \max \left\{ \lim_{n \rightarrow \infty} \varrho_{\text{ord}}(M, P(n, x)), \lim_{n \rightarrow \infty} \varrho_{\text{ord}}(M, Q(n, x)) \right\}$$

for any (possibly all) $x \in [0, 1]$?

A short calculation shows that there exists $x_0 \in [0, 1]$ such that $\lim_{n \rightarrow \infty} \varrho_{\text{ord}}(M, Q(n, x)) \leq \lim_{n \rightarrow \infty} \varrho_{\text{ord}}(M, P(n, x))$ iff $x < x_0$.

We end this section by summarizing the current knowledge on ordered graphs with three vertices in Table 1. Theorem 7.2 is proven in Section 2, and justification for the formulas in Table 1 are given in Section 3.

F	$i_{\text{ord}}(F, x)$	$I_{\text{ord}}(F, x)$
$n = 3, m = 1$	$(1 - \sqrt{1 - x})^2 (2\sqrt{1 - x} + 1)$	$3x - 2x^{3/2}$
$S_L(2), S_R(2)$	$(1 - \sqrt{1 - x})^2 (2\sqrt{1 - x} + 1)$	$1 - (1 - x)^{3/2}$
M	$g_3(x)$	unknown
$n = 3, m = 3$	$g_3(x)$	$x^{3/2}$

TABLE 1. Formulas for i_{ord} and I_{ord} for all ordered graphs on 3 vertices.

2. Proof of Theorem 7.2

Recall that we are assuming $G = ([n], E)$. For a vertex i in G , the *right-degree* of i is $d_G^+(i) := |\{j > i : ij \in E\}|$. Recall that when we write uv for an edge, we implicitly mean $u < v$. The *left vertex* of $vw \in E$ is v . A vertex v has *full right-degree* if $E \supset \{vw : w > v\}$. Note that $N_{\text{ord}}(S_L(k), G) = \sum_{i \in V(G)} \binom{d_G^+(i)}{k}$.

PROOF OF THEOREM 7.2. If $m = 0$ or $m = \binom{n}{2}$, the theorem is trivial, so we assume that $0 < m < \binom{n}{2}$.

Let $G = ([n], E)$ with $|E| = m > 0$ such that G maximizes the number of copies of $S_L(k)$ among all ordered n vertex m edge graphs. Suppose that there exists $i \in [n - 1]$ such that $d_G^+(i) < d_G^+(i + 1)$. Let G_i be the ordered graph obtained from G by interchanging the

positions of i and $i + 1$ in the ordering of $V(G)$. We say that G_i is obtained from G by *swapping* i and $i + 1$.

Since i and $i + 1$ are consecutive in the ordering of G , we have that $d_G^+(j) = d_{G_i}^+(j)$ for all $j \in [n] \setminus \{i, i + 1\}$. If there is no edge between i and $i + 1$, then $d_G^+(i) = d_{G_i}^+(i)$ and $d_G^+(i + 1) = d_{G_i}^+(i + 1)$ as well, so $N(S_L(k), G_i) = N(S_L(k), G)$. If there is an edge between i and $i + 1$, then $d_{G_i}^+(i) = d_G^+(i) - 1$ and $d_{G_i}^+(i + 1) = d_G^+(i + 1) + 1$. By convexity of the binomial coefficient and the fact that $d_G^+(i) < d_G^+(i + 1)$,

$$\begin{aligned} N_{\text{ord}}(S_L(k), G_i) - N_{\text{ord}}(S_L(k), G) &= \sum_{j \in V(G_i)} \binom{d_{G_i}^+(j)}{k} - \sum_{j \in V(G)} \binom{d_G^+(j)}{k} \\ &= \left(\binom{d_{G_i}^+(i)}{k} + \binom{d_{G_i}^+(i + 1)}{k} \right) - \left(\binom{d_G^+(i)}{k} + \binom{d_G^+(i + 1)}{k} \right) \\ &= \left(\binom{d_G^+(i) - 1}{k} + \binom{d_G^+(i + 1) + 1}{k} \right) - \left(\binom{d_G^+(i)}{k} + \binom{d_G^+(i + 1)}{k} \right) \geq 0. \end{aligned}$$

We continue swapping adjacent vertices in this way until there is no $i \in [n - 1]$ such that $d_G^+(i) < d_G^+(i + 1)$. Call the resulting graph G' . Then $d_{G'}^+(v) \geq d_{G'}^+(w)$ for all $v < w$ and $N_{\text{ord}}(S_L(k), G') = N_{\text{ord}}(S_L(k), G)$ by our assumption on G .

Let $v_{G'} := \min\{i \in [n] : d_{G'}^+(i) < n - i\}$, which exists since $m < \binom{n}{2}$, and choose $w \in [n]$ such that $w > v_{G'}$ and $v_{G'}w \notin E(G')$. Let $x_{G'} := \max\{i \in [n] : d_{G'}^+(i) > 0\}$, which exists since $m > 0$. For brevity, set $v' := v_{G'}$ and $x' := x_{G'}$. Note that if $v' = 1$, then $x' \geq 1 = v'$; if $v' \geq 2$, then the definition of v' implies that $d_{G'}^+(v' - 1) = n - v' + 1 > 0$, so $x' \geq v' - 1$. Thus $v' \leq x' + 1$. In the next paragraph, we show that there is a graph G'' with the same number of copies of $S_L(k)$ as G' satisfying $x_{G''} \leq v_{G''} \leq x_{G''} + 1$.

Suppose that $v' < x'$ and fix $y \in [n]$ such that $x'y \in E(G')$ (recall that this notation implies $x' < y$). Let $H = ([n], (E(G') \cup \{v'w\}) \setminus \{x'y\})$. Then $d_H^+(v') = d_{G'}^+(v') + 1$ and $d_H^+(x') = d_{G'}^+(x') - 1$. By convexity of the binomial coefficient and the fact that $d_{G'}^+(v') \geq d_{G'}^+(x')$,

$$N_{\text{ord}}(S_L(k), H) \geq N_{\text{ord}}(S_L(k), G') = N_{\text{ord}}(S_L(k), G).$$

We continue adding and deleting edges in this way until we reach a new graph, call it G'' , satisfying $v_{G''} \geq x_{G''}$. Notice that $x_{G''} \leq v_{G''} \leq x_{G''} + 1$ and $N_{\text{ord}}(S_L(k), G'') = N_{\text{ord}}(S_L(k), G)$.

Setting $v := v_{G''}$, we further see that

$$E(G'') = \{ij : i \in [v-1], j \in [n]\} \cup A$$

for some $A \subset \{vj : j \in \{v+1, v+2, \dots, n\}\}$ with $|A| = b$ for some $0 \leq b < n - v$. Since $d_{G''}^+(v) = b$ and $d_{G''}^+(w) = 0$ for $w > v$,

$$N_{\text{ord}}(S_L(k), G) = N_{\text{ord}}(S_L(k), G'') = N_{\text{ord}}(S_L(k), S_L(n, m))$$

as required.

Let $G = ([n], E)$ with $|E| = m$ such that G minimizes the number of copies of $S_L(k)$ among all ordered n vertex m edge graphs. Suppose further that there are some $i, j \in V(G)$ with $i < j$ such that $d_G^+(j) < n - j$ and $d_G^+(i) > d_G^+(j)$. Then choose $v \in [n]$ such that $j < v$ and $jk \notin E$. As $d_G^+(i) > d_G^+(j) \geq 0$, we also choose $w \in [n]$ such that $i < w$ and $iw \in E$. Let $H = ([n], (E \cup \{jv\}) \setminus \{iw\})$. Then $d_H^+(i) = d_G^+(i) - 1$ and $d_H^+(j) = d_G^+(j) + 1$. By convexity of the binomial coefficient,

$$N_{\text{ord}}(S_L(k), H) \leq N_{\text{ord}}(S_L(k), G).$$

By our assumption on G , we must have $N_{\text{ord}}(S_L(k), H) = N_{\text{ord}}(S_L(k), G)$. We repeatedly remove an edge iw and add an edge jv for $i < j$ satisfying $d^+(j) < n - j$ and $d^+(i) > d^+(j)$ until no such i and j exist. Call the resulting graph G' and note that G' has m edges. Then $N_{\text{ord}}(S_L(k), G') = N_{\text{ord}}(S_L(k), G)$ and

$$\text{for all } i < j \text{ in } G', \text{ either } d_{G'}^+(j) = n - j \text{ or } d_{G'}^+(i) \leq d_{G'}^+(j). \quad (46)$$

We note here that the right degree sequence of $S_R(n, m)$ is given by (not necessarily in the vertex order):

$$0, 1, \dots, a-1, \underbrace{a, \dots, a}_{n-a-b}, \underbrace{a+1, \dots, a+1}_b. \quad (47)$$

Our plan is to alter G' until its right degree sequence is the same as in (47). This will allow us to conclude that $N_{\text{ord}}(S_L(k), G') = N_{\text{ord}}(S_L(k), S_R(n, m))$.

Set $v_{G'} := \min\{j \in [n] : d_{G'}^+(j) = n - j\}$. For brevity, set $v' := v_{G'}$. Then $d_{G'}^+(i) \leq d_{G'}^+(j)$ for all $i < j \leq v' - 1$ by (46). Furthermore, if there exists $i \geq v'$ such that $d_{G'}^+(i) < n - i$, then we can find consecutive vertices x, y such that $v' \leq x < y$, $d_{G'}^+(y) < n - y$ and $d_{G'}^+(x) = n - x$. But this is impossible as (46) implies that $d_{G'}^+(y) \geq d_{G'}^+(x) = n - x > n - y$. Consequently, $d_{G'}^+(i) = n - i$ for all $i \in \{v', v' + 1, \dots, n\}$.

If $v' = n$, then $d_{G'}^+(n-1) = 0$ and $d_{G'}^+(i) \leq d_{G'}^+(n-1) = 0$ for all $i \in [n-1]$, so $|E(G')| = 0$. This contradicts $m > 0$, so we must have $v' < n$. If $d_{G'}^+(i) \geq n - v' + 1$ for some $i \in [v' - 1]$, then since $i \leq v' - 1$, we have $d_{G'}^+(v' - 1) \geq d_{G'}^+(i) \geq n - v' + 1 = n - (v' - 1)$. This implies that $d_{G'}^+(v' - 1) = n - (v' - 1)$, contradicting the definition of v' . Therefore $v' < n$ and $d_{G'}^+(i) \leq n - v'$ for all $i \in [v' - 1]$.

Next, suppose that $d_{G'}^+(1) < n - v' - 1$. Note that $d_{G'}^+(v') = n - v' > n - v' - 1 > d_{G'}^+(1)$ implies that $d_{G'}^+(v') \geq d_{G'}^+(1) + 2$. Choose $x \in [n] \setminus \{1\}$ such that $1x \notin E(G')$ and $y \in \{v' + 1, \dots, n\}$ such that $v'y \in E(G')$. Let $H' = ([n], (E(G') \cup \{1x\}) \setminus \{v'y\})$. Then $d_{H'}^+(1) = d_{G'}^+(1) + 1$ and $d_{H'}^+(v') = d_{G'}^+(v') - 1$. By convexity of binomial coefficients and the fact that $d_{G'}^+(v') \geq d_{G'}^+(1) + 2$,

$$N_{\text{ord}}(S_L(k), H') \leq N_{\text{ord}}(S_L(k), G') = N_{\text{ord}}(S_L(k), G).$$

We continue adding and deleting edges in this way until we reach a graph, call it G'' , that satisfies $d_{G''}^+(1) \geq n - v_{G''} - 1$. Note that $n - v_{G''} - 1 \leq d_{G''}^+(1) \leq n - v_{G''}$.

Set $v := v_{G''}$. Since the right-degrees of vertices x are nondecreasing for $1 \leq x \leq v$, there is some $0 < b_0 \leq v$ such that $d_{G''}^+(i) = n - v - 1$ for all $i \in [v - b_0]$, $d_{G''}^+(i) = n - v$ for the remaining b_0 vertices $v - b_0 + 1 \leq i \leq v$; we have already observed that $d_{G''}^+(i) = n - i$ for all $i > v$. Thus the right degree sequence of G'' is given by

$$0, 1, \dots, n - v - 2, \underbrace{n - v - 1, \dots, n - v - 1}_{v+1-b_0}, \underbrace{n - v, \dots, n - v}_{b_0}. \quad (48)$$

If $b_0 = v$, set $b = 0$ and $a = n - v$. Otherwise, set $b = b_0$ and $a = n - v - 1$. In either case, the degree sequence of $S_R(n, m)$ in (47) is the same as that of G'' in (48) and therefore

$$N_{\text{ord}}(S_L(k), G) = N_{\text{ord}}(S_L(k), G'') = N_{\text{ord}}(S_L(k), S_R(n, m))$$

as required. \square

3. Ordered graphs on three vertices

First, we address the ordered graphs on three vertices and only one edge. Let F_{ij} have vertex set $\{1, 2, 3\}$ and only one edge ij for $i, j \in \{1, 2, 3\}$ distinct. We obtain the following result.

PROPOSITION 7.7. *For any $x \in [0, 1]$, we have that $I_{\text{ord}}(F_{12}, x) = 3x - 2x^{3/2}$ and $i_{\text{ord}}(F_{12}, x) = (1 - \sqrt{1-x})^2 (2\sqrt{1-x} + 1)$.*

PROOF. Let $G = ([n], E)$ be an ordered graph with $\lfloor x \binom{n}{2} \rfloor$ edges and note that every edge $ij \in E(G)$ (for $i < j$) contributes exactly $n - j$ copies of F_{12} .

Let G be chosen to maximize the number of copies of F_{12} . Suppose that $i_1 j_i \notin E$ and $i_2 j_2 \in E$ for some $i_1, i_2, j_1, j_2 \in [n]$ such that $i_1 < j_1$ and $j_1 < j_2$. Now delete $i_2 j_2$ and add $i_1 j_1$. By doing this, we have gained

$$n - j_1 - (n - j_2) = j_2 - j_1 > 0$$

copies of F_{12} . This contradicts our choice of G , so no such pair of edges exists. Thus, G contains all of the edges with the least right vertex, so G must be a clique on the smallest $\sqrt{x} \cdot n$ vertices asymptotically. Taking the limit, we see that

$$I_{\text{ord}}(F_{12}, x) = \lim_{n \rightarrow \infty} \frac{\binom{\lfloor n\sqrt{x} \rfloor}{3} + (n - \lfloor n\sqrt{x} \rfloor) \binom{\lfloor n\sqrt{x} \rfloor}{2}}{\binom{n}{3}} = x^{3/2} + 3(1 - \sqrt{x})x = 3x - 2x^{3/2}.$$

Now let G instead be chosen to minimize the number of copies of F_{12} . Suppose that $i_1 j_i \in E$ and $i_2 j_2 \notin E$ for some $i_1, i_2, j_1, j_2 \in [n]$ such that $i_2 < j_2$ and $j_1 < j_2$. Now delete

i_1j_1 and add i_2j_2 . By doing this, we have lost

$$n - j_1 - (n - j_2) = j_2 - j_1 > 0$$

copies of F_{12} . This contradicts our choice of G , so no such pair of edges exists. Thus, G contains all of the edges with the greatest right vertex, so G must be $S_R(n, x\binom{n}{2})$ asymptotically.

Taking the limit, we see that

$$i_{\text{ord}}(F_{12}, x) = \lim_{n \rightarrow \infty} \frac{\binom{n - \lfloor n\sqrt{1-x} \rfloor}{3} + \lfloor n\sqrt{1-x} \rfloor \cdot \binom{n - \lfloor n\sqrt{1-x} \rfloor}{2}}{\binom{n}{3}},$$

which is equal to

$$(1 - \sqrt{1-x})^3 + 3\sqrt{1-x}(1 - \sqrt{1-x})^2 = (1 - \sqrt{1-x})^2(2\sqrt{1-x} + 1)$$

as claimed. □

By symmetry, the same is true for F_{23} . The proof is a bit different for F_{13} .

PROPOSITION 7.8. *For any $x \in [0, 1]$, we have that $I_{\text{ord}}(F_{13}, x) = 3x - 2x^{3/2}$ and $i_{\text{ord}}(F_{13}, x) = (1 - \sqrt{1-x})^2(2\sqrt{1-x} + 1)$.*

PROOF. Let $G = ([n], E)$ be an ordered graph with $\lfloor x\binom{n}{2} \rfloor$ edges and note that every edge $ij \in E(G)$ contributes exactly $j - i - 1$ copies of F_{13} . For an edge $ij \in E$, let $j - i$ be its *length*.

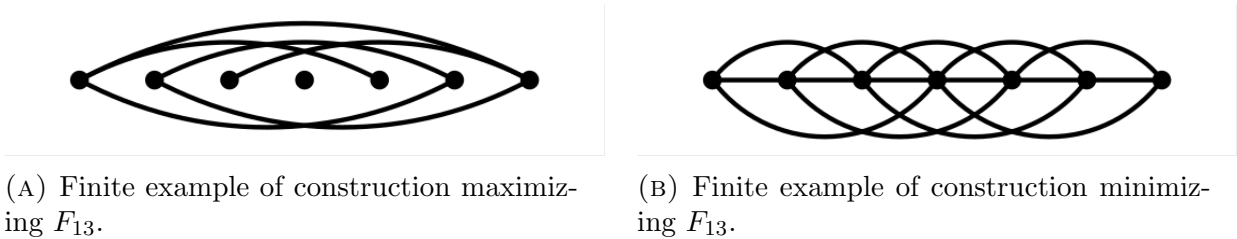


FIGURE 7.3. Extremal constructions for F_{13} .

Let G be chosen to maximize the number of copies of F_{13} . Suppose that $i_1j_i \notin E$ and $i_2j_2 \in E$ for some $i_1, i_2, j_1, j_2 \in [n]$ such that $i_1 < j_1$ and $j_1 - i_1 > j_2 - i_2$. Now delete i_2j_2 and add i_1j_1 . By doing this, we have gained

$$j_1 - i_1 - 1 - (j_2 - i_2 - 1) = (j_1 - i_1) - (j_2 - i_2) > 0$$

copies of F_{13} . This contradicts our choice of G , so no such pair of edges exists. Thus, G contains all of the edges with greatest length. For G to have edge density x , it must contain all edges of length at least $(1 - \sqrt{x})n$ asymptotically. A finite example of this construction is visualized in Figure 7.3a. Taking the limit, we see that

$$I_{\text{ord}}(F_{13}, x) = \lim_{n \rightarrow \infty} \frac{\sum_{\ell=(1-\sqrt{x})n}^n \binom{n-\ell}{3}}{\binom{n}{3}} = -x(2\sqrt{x} - 3) = 3x - 2x^{3/2}.$$

Now let G instead be chosen to minimize the number of copies of F_{13} . Suppose that $i_1j_i \in E$ and $i_2j_2 \notin E$ for some $i_1, i_2, j_1, j_2 \in [n]$ such that $i_1 < j_1$ and $j_1 - i_1 > j_2 - i_2$. Now delete i_1j_1 and add i_2j_2 . By doing this, we have lost

$$j_1 - i_1 - 1 - (j_2 - i_2 - 1) = (j_1 - i_1) - (j_2 - i_2) > 0$$

copies of F_{13} . This contradicts our choice of G , so no such pair of edges exists. Thus, G contains all of the edges with least length. For G to have edge density x , it must contain all edges of length at most $(1 - \sqrt{1-x})n$ asymptotically. A finite example of this construction is visualized in Figure 7.3b. Taking the limit, we see that

$$i_{\text{ord}}(F_{13}, x) = \lim_{n \rightarrow \infty} \frac{\sum_{\ell=1}^{(1-\sqrt{1-x})n} \binom{n-\ell}{3}}{\binom{n}{3}} = (1 - \sqrt{1-x})^2 (2\sqrt{1-x} + 1),$$

as claimed. □

Now, we consider $S_L(k)$ and $S_R(k)$. Due to symmetry, we just compute $I_{\text{ord}}(S_L(k), x)$ and $i_{\text{ord}}(S_L(k), x)$. By Theorem 7.2, we know that the construction maximizing the number of copies of $S_L(k)$ is $S_L(n, x \binom{n}{2})$, while the construction minimizing the number of copies is $S_R(n, x \binom{n}{2})$. Notice that, as we are only considering the number of copies asymptotically,

we may assume that $b = 0$. In $S_L(n, x \binom{n}{2})$, every triple of vertices except those entirely contained in the empty section include $S_L(k)$ as a subgraph, so we see that

$$I_{\text{ord}}(S_L(k), x) = \lim_{n \rightarrow \infty} \frac{\binom{n}{3} - \binom{\sqrt{1-x}n}{3}}{\binom{n}{3}} = 1 - (1-x)^{3/2}.$$

In $S_R(n, x \binom{n}{2})$, it is exactly the triples of vertices with at least two vertices in the complete section that include $S_L(k)$ as a subgraph, so we see that

$$i_{\text{ord}}(S_L(k), x) = \lim_{n \rightarrow \infty} \frac{\sqrt{1-x}n \binom{(1-\sqrt{1-x})n}{2} + \binom{(1-\sqrt{1-x})n}{3}}{\binom{n}{3}},$$

which is equal to

$$3\sqrt{1-x}(1-\sqrt{1-x})^2 + (1-\sqrt{1-x})^3 = (1-\sqrt{1-x})^2(2\sqrt{1-x}+1)$$

as claimed.

It remains to justify the last row of Table 1. When $F = K_3$, the ordered graph case is the same as the simple graph case. Thus we are able to apply Theorem 7.3 to see that $i_{\text{ord}}(K_3, x) = g_3(x)$ and the Kruskal–Katona theorem [19, 21] to see that $I_{\text{ord}}(K_3, x) = x^{3/2}$.

APPENDIX A

Key inequalities for the inducibility of rainbow graphs

We prove key inequalities from Chapter 2.

PROOF OF (4). Let $q, q' \geq 0$ and $t, t' > 0$. Recall that $p(q, t)$ is the maximum of $\prod_i q_i$ where $q_1 + \dots + q_t = q$ and each $q_i \geq 0$ is an integer. Let q_1, \dots, q_t integers such that $p(q, t) = \prod_{i=1}^t q_i$ and $q'_1, \dots, q'_{t'}$ integers such that $p(q', t') = \prod_{i=1}^{t'} q'_i$. Then,

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

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 (7). 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 (8) 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

$$\frac{1}{e(k-2)^{k-2}} > 1.4 \left(\frac{0.6}{k-3} \right)^{k-3}. \tag{49}$$

We see that

$$\frac{(k-2)^{k-2}}{(k-1)^{k-1}} > \frac{0.6(k-3)^{k-3}}{(k-2)^{k-2}}, \text{ or equivalently } f(k) := \frac{(k-2)^{2k-4}}{(k-1)^{k-1}(k-3)^{k-3}} > 0.6 \quad (50)$$

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 (49) and (50), 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 (10). We will show that

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

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 (8) 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 2.7, 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. □

APPENDIX B

Supporting lemmas for maximum star densities

We prove supporting lemmas from Chapter 6.

PROOF OF LEMMA 6.5. Let $F \in \mathcal{C}$ be the function that we want to prove δ -good for $1-W$. Using the fact that $G : [0, 1] \rightarrow \mathbb{R}$ given by $G(x) := F(1-x)$ from [37, Lemma 3.1(ii)] is δ -good for W we find that

$$\gamma = t(\cdot, 1-W) = 1 - t(\cdot, W) \text{ and } \eta = 1 - \sqrt{1-\gamma}$$

satisfy

$$\begin{aligned} \int_0^1 F(d_{1-W}(x)) &= \int_0^1 G(d_W(x)) \\ &< \max \left\{ \left(1 - \sqrt{1-\gamma}\right) G(0) + \sqrt{1-\gamma} G\left(\sqrt{1-\gamma}\right), \sqrt{\gamma} G(1-\sqrt{\gamma}) + (1-\sqrt{\gamma}) G(1) \right\} + \delta \\ &= \max \{ (1-\eta)F(\eta) + \eta F(1), (1-\sqrt{\gamma})F(0) + \sqrt{\gamma}F(\sqrt{\gamma}) \} + \delta \end{aligned}$$

as desired. □

PROOF OF LEMMA 6.6. Let $F \in \mathcal{C}$ be any function that we want to prove δ -good for $[\lambda, W]$. By [37, Lemma 3.1(iii)], the function $H : [0, 1] \rightarrow \mathbb{R}$ given by $H(x) = F((1-\lambda)x)$ for all $x \in [0, 1]$ is in \mathcal{C} . Thus it is δ -good for W , which tells us that

$$\int_0^1 H(d_W(x)) \, dx < \max \{ (1-\sqrt{\gamma})H(0) + \sqrt{\gamma}H(\sqrt{\gamma}), (1-\eta)H(\eta) + \eta H(1) \} + \delta,$$

where $\gamma = t(\cdot, W)$ and $\eta = 1 - \sqrt{1-\gamma}$. Since

$$\begin{aligned} \int_0^1 F(d_{[\lambda, W]}(x)) \, dx &= \lambda F(0) + \int_\lambda^1 F\left((1-\lambda)d_W\left(\frac{x-\lambda}{1-\lambda}\right)\right) \, dx \\ &= \lambda F(0) + (1-\lambda) \int_0^1 H(d_W(x)) \, dx, \end{aligned}$$

it follows that either

$$\begin{aligned} \int_0^1 F(d_{[\lambda, W]}(x)) \, dx &< \lambda F(0) + (1 - \lambda)((1 - \sqrt{\gamma})F(0) + \sqrt{\gamma}F((1 - \lambda)\sqrt{\gamma}) + \delta) \\ &= \lambda F(0) + (1 - \lambda)(1 - \sqrt{\gamma})F(0) + (1 - \lambda)\sqrt{\gamma}F((1 - \lambda)\sqrt{\gamma}) + (1 - \lambda)\delta \\ &\leq \lambda F(0) + (1 - \lambda)(1 - \sqrt{\gamma})F(0) + (1 - \lambda)\sqrt{\gamma}F((1 - \lambda)\sqrt{\gamma}) + \delta, \end{aligned}$$

or

$$\begin{aligned} \int_0^1 F(d_{[\lambda, W]}(x)) \, dx &< \lambda F(0) + (1 - \lambda)((1 - \eta)F((1 - \lambda)\eta) + \eta F(1 - \lambda) + \delta) \\ &= \lambda F(0) + (1 - \lambda)(1 - \eta)F((1 - \lambda)\eta) + (1 - \lambda)\eta F(1 - \lambda) + (1 - \lambda)\delta \\ &\leq \lambda F(0) + (1 - \lambda)(1 - \eta)F((1 - \lambda)\eta) + (1 - \lambda)\eta F(1 - \lambda) + \delta. \end{aligned}$$

In the former case the right side simplifies to

$$\left(1 - \sqrt{\gamma'}\right) F(0) + \sqrt{\gamma'} F\left(\sqrt{\gamma'}\right) + \delta,$$

where $\gamma' = (1 - \lambda)^2\gamma = t(|, [\lambda, W])$, meaning that F is, in particular, δ -good for $[\lambda, W]$.

So we may assume that the second alternative occurs. Setting $x = \lambda$, $y = (1 - \lambda)\eta$, and $z = (1 - \lambda)(1 - \eta)$ we thus get

$$\int_0^1 F(d_{[\lambda, W]}(x)) \, dx < xF(0) + zF(y) + yF(y + z) + \delta.$$

Since $y^2 + 2yz = (1 - \lambda)^2(2\eta - \eta^2) = (1 - \lambda)^2\gamma = \gamma'$, it follows in view of [37, Lemma 3.3] that

$$\int_0^1 F(d_{[\lambda, W]}(x)) \, dx < \max \left\{ \left(1 - \sqrt{\gamma'}\right) F(0) + \sqrt{\gamma'} F\left(\sqrt{\gamma'}\right), (1 - \eta')F(\eta') + \eta'F(1) \right\} + \delta,$$

where $\eta' = 1 - \sqrt{1 - \gamma'}$. This tells us that F is indeed δ -good for $[\lambda, W]$. \square

APPENDIX C

Copyright Statements

Much of the material in Chapter 2 was previously published by Cambridge University Press [7]. On their website, it states:

“This is an open access article distributed under the terms of the Creative Commons CC BY license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.”

Chapters 4, 6, and 7 include material that was first published in the SIAM Journal on Discrete Mathematics in 2025 [8], published by the Society for Industrial and Applied Mathematics (SIAM). Copyright © by SIAM. Unauthorized reproduction of work from this article is prohibited. The policy as stated in the Consent to Publish form is:

“The Author may reproduce and distribute the Work (including derivative works) in connection with the Author’s teaching, technical collaborations, conference presentations, lectures, or other scholarly works and professional activities as well as to the extent the fair use provisions of the U.S. Copyright Act permit. If the copyright is granted to the Publisher, then the proper notice of the Publisher’s copyright should be provided. . . . In any authorized reproduction or duplication of the work, in whole or in part, by the Author or an entity as permitted herein, the original publication must be properly credited in the following manner: “First Published in [Publication] in [volume and number, or year], published by the Society for Industrial and Applied Mathematics (SIAM)” and the copyright notice as stated in the article itself (e.g., “Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.”) must be placed on all copies.”

Bibliography

- [1] R. AHLWEDE AND G. O. H. KATONA. Graphs with maximal number of adjacent pairs of edges. *Acta Math. Acad. Sci. Hungar.* **32** (1978) 97–120.
- [2] 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.
- [3] J. BALOGH, P. HU, B. LIDICKÝ, F. PFENDER, J. VOLEC AND M. YOUNG. Rainbow triangles in three-colored graphs. *J. Combin. Theory Ser. B* **126** (2017) 83–113.
- [4] B. BOLLOBÁS. Relations between sets of complete subgraphs. In Proceedings of the Fifth British Combinatorial Conference, *Utilitas Math.*, Winnipeg (1976) 161–170.
- [5] B. BOLLOBÁS, Y. EGAWA, A. HARRIS, AND G. JIN. The maximal number of induced r -partite subgraphs. *Graphs Combin.*, **11** (1995) 1–19.
- [6] J. BROWN AND A. SIDORENKO. The inducibility of complete bipartite graphs. *J. Graph Theory*, **18** (1994) 629–645.
- [7] E. CAIRNCROSS, C. MIZGERD AND D. MUBAYI. Inducibility of rainbow graphs. *Math. Proc. Cambridge Philos. Soc.* **180** (2026) 191–218.
- [8] E. CAIRNCROSS AND D. MUBAYI. Ordered and colored subgraph density problems. *SIAM J. Discrete Math.* **39** (2025) 75–91.
- [9] T. CHAO AND H. YU. Kruskal–Katona-type problems via the entropy method. *J. Combin. Theory Ser. B* **169** (2024) 480–506.
- [10] D. CONLON, J. FOX AND B. SUDAKOV. Hypergraph Ramsey numbers. *J. Amer. Math. Soc.* **23** (2009) 247–266.
- [11] 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.
- [12] D. C. FISHER. Lower bounds on the number of triangles in a graph. *J. Graph Theory* **13** (1989) 505–512.
- [13] J. FOX, H. HUANG AND C. LEE. Packing problems. Methods and Challenges in Extremal and Probabilistic Combinatorics workshop, Banff International Research Station (2015).
- [14] J. FOX, L. SAUERMAN 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.
- [15] A. W. GOODMAN. On sets of acquaintances and strangers at any party. *Amer. Math. Monthly* **66** (1959) 778–783.
- [16] H. G. HADŽIIVANOV, V. S. NIKIFOROV. The Nordhaus-Stewart-Moon-Moser inequality. *Serdica* **4** (1978) 344–350.
- [17] H. HUANG. On the maximum induced density of directed stars and related problems. *SIAM J. Discrete Math.* **28** (2014) 92–98.
- [18] H. HUANG. Personal communication (2019).
- [19] G. KATONA. A theorem of finite sets. In *Theory of Graphs*, Academic Press, New York; Budapest (1968) 187–207.
- [20] D. KRÁL, S. NORIN AND J. VOLEC. A bound on the inducibility of cycles. *J. Combin. Theory Ser. A* **161** (2019) 359–363.
- [21] J. B. KRUSKAL. The number of simplices in a complex. In *Mathematical Optimization Techniques*, Univ. of California Press, Berkeley, California (1963) 251–278.
- [22] B. LIDICKÝ, C. MATTES AND F. PFENDER. C_5 is almost a fractalizer. *J. Graph Theory* **104** (2023) 220–244.
- [23] H. LIU, O. PIKHURKO, AND K. STADEN. The exact minimum number of triangles in graphs with given order and size. *Forum of Math Pi* **8** (2020) 1–144.

- [24] X. LIU, J. MA, AND T. ZHU. The inducibility of Turán graphs. Preprint, available at arXiv:2601.10548 (2026).
- [25] X. LIU, AND D. MUBAYI. The feasible region of hypergraphs. *J. Combin. Theory Ser. B* **148** (2021) 23–59.
- [26] X. LIU, D. MUBAYI, AND C. REIHER. The feasible region of induced graphs. *J. Combin. Theory Ser. B* **158** (2023) 105–135.
- [27] L. LOVÁSZ. Large Networks and Graph Limits. *Colloquium Presentations 60*, American Mathematical Society, Providence, Rhode Island (2012).
- [28] L. LOVÁSZ, M. SIMONOVITS. On the number of complete subgraphs of a graph II. In *Studies in Pure Mathematics*, P. Erdős, L. Alpár, G. Halász, and A. Sárközy, ed., Birkhäuser, Basel (1983) 459–495.
- [29] J. W. MOON, L. MOSER. On a problem of Turán. *Magyar Tud. Akad. Mat. Kutató Int. Közl* **7** (1962) 283–286.
- [30] D. MUBAYI AND A. RAZBOROV. Polynomial to exponential transition in Ramsey theory. *Proc. Lond. Math. Soc.* **122** (2021) no. 1, 69–92.
- [31] V. S. NIKIFOROV. The number of cliques in graphs of given order and size. *Trans. Amer. Math. Soc.* **363** (2011) 1599–1618.
- [32] E. A. NORDHAUS, B. M. STEWART. Triangles in an ordinary graph. *Canadian J. Math.* **15** (1963) 33–41.
- [33] D. OLPP. A conjecture of Goodman and the multiplicities of graphs. *Australas. J. Combin.* **14** (1996) 267–282.
- [34] N. PIPPENGER AND M. GOLUBIC. The inducibility of graphs. *J. Combin. Theory Ser. B* **19** (1975) no. 3, 189–203.
- [35] A. A. RAZBOROV. On the minimal density of triangles in graphs. *Combin. Probab. Comput.* **17** (2008) no. 4, 603–618.
- [36] C. REIHER. The clique density theorem. *Ann. of Math.* **184** (2016) 683–707.
- [37] C. REIHER AND S. WAGNER. Maximum star densities. *Studia Sci. Math. Hungar.* **55** (2017) 238–259.
- [38] R. YUSTER. On the exact maximum induced density of almost all graphs and their inducibility. *J. Combin. Theory Ser. B* **136** (2019) 81–109.
- [39] R. YUSTER. Inducibility in H -free graphs and inducibility of Turán graphs. *J. Combin. Theory Ser. B* **178** (2026) 1–26.

Vita

Education

2021-2026 Ph.D. in Mathematics, University of Illinois Chicago.

2017-2021 B.A. in Mathematics, Oberlin College.

Publications

E. CAIRNCROSS, C. MIZGERD AND D. MUBAYI. Inducibility of rainbow graphs. *Math. Proc. Cambridge Philos. Soc.* **180** (2026) 191–218.

E. CAIRNCROSS AND D. MUBAYI. Ordered and colored subgraph density problems. *SIAM J. Discrete Math.* **39** (2025) 75–91.

E. CAIRNCROSS, J. OGLAND-HAND, B. ADAMS, AND R. MIDDLETON. Nationwide cost and capacity estimates for sedimentary basin geothermal power and implications for geologic CO₂ storage. *Frontiers in Energy Research* **12** (2024) 1–11.

(This project was completed as part of a summer internship with Carbon Solutions LLC in Energy Research, as opposed to Mathematics, and lists co-authors in order of most contribution to least instead of alphabetically.)

E. CAIRNCROSS, J. CARLSON, P. HOLLANDER, B. KITCHEN, E. LOPEZ, AND A. ZHUANG. Throttling for standard zero forcing on directed graphs. *The Australasian Journal of Combinatorics* **84** (2022) Part 1, 1–27.

E. CAIRNCROSS, S. FORD, E. GARCIA, AND K. JABBUSCH. Classifying toric surface codes of dimension 7. *Involve* **14** (2021) No. 4, 605–616.

Selected Talks

2025 Special Session on Recent Advances in Graph Theory
AMS Sectional Meeting, Tulane University, New Orleans, LA.

2025 Special Session on Graphs, Hypergraphs, and Extremal Combinatorics
Special Session on Sparse Graphs: Colorings, Randomness, Ramsey Theory
Joint Mathematics Meeting, Seattle, WA.

2024 Early Career Conference in Combinatorics, UIUC, Urbana, IL.

2024 9th Lake Michigan Workshop on Combinatorics and Graph Theory
Western Michigan University, Kalamazoo, MI.

2023 Discrete Math Seminar, Iowa State University, Ames, IA.

2023 Combinatorics and Probability Seminar, UIC, Chicago, IL.