2022

Polychromatic colorings of certain subgraphs of complete graphs and maximum densities of substructures of a hypercube

Ryan Tyler Hansen
West Virginia University, rhansen@math.wvu.edu

Polychromatic colorings of certain subgraphs of complete graphs and maximum densities of substructures of a hypercube

Ryan Hansen

Dissertation submitted
to the Eberly College of Arts and Sciences
at West Virginia University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in
Mathematics

John Goldwasser, Ph.D., Chair
Hon-Jian Lai, Ph.D.
Jerzy Wojciechowski, Ph.D.
Cun-Quan Zhang, Ph.D.
Department of Mathematics

Elaine M. Eschen, Ph.D.
Department of Computer Science

Morgantown, West Virginia
2022

Keywords: hypercube, inducibility, perfect cycle, polychromatic coloring, long cycles

Abstract

Polychromatic colorings of certain subgraphs of complete graphs and maximum densities of substructures of a hypercube

Ryan Hansen

If G is a graph and $\mathcal{H}$ is a set of subgraphs of G, we say that an edge-coloring of G is $\mathcal{H}$-polychromatic if every graph from $\mathcal{H}$ gets all colors present in G on its edges. The $\mathcal{H}$-polychromatic number of G, denoted by $\mathrm{poly}_{\mathcal{H}}(G)$, is the largest number of colors in an $\mathcal{H}$-polychromatic coloring. In Chapter 1 we determine $\mathrm{poly}_{\mathcal{H}}(G)$ exactly when G is a complete graph on n vertices, q is a fixed nonnegative integer, and $\mathcal{H}$ is one of three families: the family of all matchings spanning $n-q$ vertices, the family of all 2-regular graphs spanning at least $n-q$ vertices, and the family of all cycles of length precisely $n-q$. There are connections with an extension of results on Ramsey numbers for cycles in a graph.

Let H and K be subsets of the vertex set $V(Q_d)$ of the d-cube Q_d (we call H and K configurations in Q_d). We say K is an *exact copy* of H if there is an automorphism of Q_d which sends H to K. If d is a positive integer and H is a configuration in Q_d, we define $\lambda(H,d)$ to be the limit as n goes to infinity of the maximum fraction, over all subsets S of $V(Q_n)$, of sub-d-cubes of Q_n whose intersection with S is an exact copy of H.

In Chapter 2, we determine $\lambda(C_8,4)$ and $\lambda(P_4,3)$ where C_8 is a "perfect" 8-cycle in Q_4 and P_4 is a "perfect" path with 4 vertices in Q_3, and make conjectures about $\lambda(C_{2d},d)$ and $\lambda(P_{d+1},d)$ for larger values of d. In Chapter 3, we determine $\lambda(H,d)$ for several configurations in Q_2, Q_3, and Q_4 as well as for an infinite family of configurations. The proofs contained in Chapters 2 and 3 include connections with counting the number of sequences with certain properties and with the inducibility of certain small graphs. In particular, we needed to determine the inducibility of two vertex disjoint edges in the family of bipartite graphs. Further, there are strong connections with the inducibility of other graphs.

Contents

List of Figures

List of Tables

Chapter 1

Polychromatic colorings of 1-regular and 2-regular subgraphs of complete graphs

Portions of the material for this chapter currently appear in publication in the Journal of Graph Theory, Volume 87, Issue 4, April 2018 [4]. This chapter also includes material that was accepted for publication in Discrete Mathematics, Volume 345, Issue 8, August 2022 [27].

1.1 Introduction

If G is a graph and $\mathcal{H}$ is a set of subgraphs of G, we say that an edge-coloring of G is *$\mathcal{H}$-polychromatic* if every graph from $\mathcal{H}$ has all colors present in G on its edges. The $\mathcal{H}$-polychromatic number of G, denoted by $\text{poly}_{\mathcal{H}}(G)$ is the largest number of colors in an $\mathcal{H}$-polychromatic coloring. If an $\mathcal{H}$-polychromatic coloring of G uses $\text{poly}_{\mathcal{H}}(G)$ colors, it is called an *optimal* $\mathcal{H}$-polychromatic coloring of G.

Alon *et al.* [2] found a lower bound for $\text{poly}_{\mathcal{H}}(G)$ when $G = Q_n$, the n-dimensional hypercube, and $\mathcal{H}$ is the family of all subgraphs isomorphic to Q_d, where d is fixed. Offner [39] showed this lower bound is, in fact, the exact value for all d and sufficiently large n. Bialostocki [7] showed that if $d = 2$, then the polychromatic number is 2 and that any optimal coloring uses each color about half the time. Goldwasser *et al.* [28] considered the case when $\mathcal{H}$ is the family of all subgraphs isomorphic to Q_d minus an edge or Q_d minus a vertex.

Bollobas *et al.* [10] treated the case where G is a tree and $\mathcal{H}$ is the set of all paths of length at least r, where r is fixed. Goddard and Henning [25] considered vertex colorings of graphs such that each open neighborhood gets all colors.

For large n, it makes sense to consider $\text{poly}_{\mathcal{H}}(K_n) = \text{poly}_{\mathcal{H}}(n)$ only if $\mathcal{H}$ consists of sufficiently large graphs. Indeed, if the graphs from $\mathcal{H}$ have at most a fixed number s of vertices, then $\text{poly}_{\mathcal{H}}(n) = 1$ for sufficiently large n by Ramsey's theorem, since even with only two colors there exists a monochromatic clique with s vertices.

Axenovich *et al.* [4] considered the case where $G = K_n$ and $\mathcal{H}$ is one of three families of spanning subgraphs: perfect matchings (so n must be even), 2-regular graphs, and Hamiltonian cycles. They determined $\text{poly}_{\mathcal{H}}(n)$ precisely for the first of these and to within a small additive constant for the other two. In this chapter, we determine the exact $\mathcal{H}$-polychromatic number of K_n, where q is a fixed nonnegative integer and $\mathcal{H}$ is one of three families of graphs: matchings spanning precisely $n - q$ vertices, $(n - q)$-cycles, and 2-regular graphs spanning at least $n - q$ vertices (so $q = 0$ gives the results of Axenovich *et al.* in [4] without the undetermined constant.)

This chapter is organized as follows. We give a few definitions and state the main results in Section 1.2. We give some more definitions in Section 1.3. The optimal polychromatic colorings in this paper are all based on a type of ordering, and in Section 1.4 we state and prove the technical ordering lemmas we will need. In Section 1.5 we prove Theorem 1.2.1, a result about matchings. In Section 1.6 we use some classical results on Ramsey numbers for cycles to take care of polychromatic numbers 1 and 2 for cycles. In Section 1.7 we prove Theorem 1.2.6, a result about coloring cycles, and use some results on long cycles in the literature to prove a necessary lemma. In

Section 1.8 we give the rather long proofs of the three main lemmas that we require. In Section 1.9 we describe precisely the various simply-ordered and nearly simply-ordered optimal polychromatic colorings of K_n. In Section 1.10 we show how our results can be reconstituted in a context which generalizes the classical results on Ramsey numbers of cycles presented in Section 1.6. In Section 1.11 we state a general conjecture of which, if true, most of our results are special cases.

1.2 Main Results

We call an edge coloring φ of K_n *ordered* if there exists an ordering $v_1, v_2, \dots, v_n$ of $V(K_n)$ such that $\varphi(v_i v_j) = \varphi(v_i v_m)$ for all $1 \le i < j < m \le n$. Moreover this coloring is *simply-ordered* if for all $i < j < m$, $\varphi(v_i v_m) = \varphi(v_j v_m) = a$ implies that $\varphi(v_t v_m) = a$ for all $i \le t \le j$. simply-ordered colorings play a fundamental role in this paper. An ordered edge coloring φ induces a vertex coloring φ' on $V(K_n)$ called the *φ-inherited coloring*, defined by $\varphi'(v_i) = \varphi(v_i v_m)$ for $i < m \le n$ and $\varphi'(v_n) = \varphi'(v_{n-1})$. We can represent the induced vertex coloring φ' by the sequence $c_1, c_2, \dots, c_n$ of colors, where $c_i = \varphi'(v_i)$ for each i. A *block* in this sequence is a maximal set of consecutive vertices of the same color. If φ is simply-ordered then the vertices in each color class appear in a single block, so in that case, the number of blocks equals the number of colors.

Let q be a fixed nonnegative integer. We define four families of subgraphs of K_n as follows.

1. $F_q = F_q(n)$ is the family of all matchings in K_n spanning precisely $n - q$ vertices (so $n - q$ must be even).
2. $C_q = C_q(n)$ is the family of all cycles of length precisely $n - q$.
3. $R_q = R_q(n)$ is the family of all 2-regular subgraphs spanning at least $n - q$ vertices.
4. $C_q^* = C_q^*(n)$ is the family of all cycles of length precisely $n - q$ where n and q are such that $\mathrm{poly}_{C_q}(n) \ge 3$.

Our main result is that for F_q, R_q, and C_q there exist optimal polychromatic colorings which are simply-ordered, or almost simply-ordered (except for C_q if $\varphi_{\mathrm{C_q}}(n) = 2$). Once we know there exists an optimal simply-ordered (or nearly simply-ordered) coloring, it is easy to construct it and to determine a formula for the polychromatic number. Our main results are the following.

Theorem 1.2.1. *For all integers q and n such that q is nonnegative and $n-q$ is positive and even, there exists an optimal simply-ordered F_q-polychromatic coloring of K_n.*

Theorem 1.2.2. *[4] If $n \geq 3$, then there exist optimal R_0-polychromatic and C_0-polychromatic colorings of K_n which can be obtained from simply-ordered colorings by recoloring one edge.*

Theorem 1.2.3. *If $n \geq 4$, then there exist optimal R_1-polychromatic and C_1-polychromatic colorings of K_n which can be obtained from simply-ordered colorings by recoloring two edges.*

Theorem 1.2.4. *Let $q \geq 2$ be an integer. If $n \geq q+3$, then there exists an optimal simply-ordered R_q-polychromatic coloring of K_n. If $n \geq q+4$, then there exists an optimal simply-ordered C_q-polychromatic coloring except if $n \in [2q+2, 3q+2]$ and $n-q$ is odd.*

Theorem 1.2.5. *Suppose $q \geq 2$ and $n \geq 6$.*

(a) If $n-q$ is even then there exists a C_q-polychromatic 2-coloring of K_n if and only if $n \geq 3q+3$.

(b) If $n-q$ is odd then there exists a C_q-polychromatic 2-coloring of K_n if and only if $n \geq 2q+2$.

Theorem 1.2.5 follows from results of Bondy and Erdős [11] and Faudree and Schelp [23].

The following result, which is needed for the proof of Theorem 1.2.4, may be of independent interest, so we state it as a theorem:

Theorem 1.2.6. *Let n and j be integers with $4 \leq j \leq n$, and let φ be an edge-coloring of K_n with at least three colors so that every j-cycle gets all colors. Then every cycle of length at least j gets all colors under φ.*

The statements about cycles in Theorems 1.2.2–1.2.5 can be used to get an extension of the result of Faudree and Schelp [23] in the following manner. Let s and t be integers with $t \geq 2, s \geq 3$, and $s \geq t$. The t-polychromatic cyclic Ramsey number $\mathrm{PR}_t(s)$ is the smallest integer $N \geq s$ such that in any t-coloring of the edges of K_N there exists an s-cycle whose edges do not contain all t colors. Note that in the special case $t = 2$, this is the classical Ramsey number for cycles, the smallest integer N such that in any 2-coloring of the edges of K_N there exists a monochromatic s-cycle. These numbers were determined for all s by Faudree and Schelp [23], confirming a conjecture of Bondy and Erdős [11].

Theorem 1.2.7. *Let $\mathrm{PR}_t(s)$ be the smallest integer $n \geq s \geq 3$ such that in any t-coloring of the edges of K_n there exists an s-cycle whose edges do not contain all t colors. If $t \geq 3$,*

$$\mathrm{PR}_t(s) = \begin{cases} s, & \text{if } 3 < s \leq 3 \cdot 2^{t-3} \\ s+1, & \text{if } s \in [3 \cdot 2^{t-3} + 1, 5 \cdot 2^{t-2} - 2] \\ s+2, & \text{if } s \in [5 \cdot 2^{t-2} - 1, 5 \cdot 2^{t-1} - 4] \\ s + \mathrm{Round}\left(\frac{s-2}{2^t-2}\right), & \text{if } s \geq 5 \cdot 2^{t-1} - 3 \end{cases}$$

where $\mathrm{Round}\left(\frac{s-2}{2^t-2}\right)$ is the closest integer to $\frac{s-2}{2^t-2}$, rounding up if it is $\frac{1}{2}$ more than an integer. Note that, as we mention in Section 1.10, $\mathrm{Round}\left(\frac{s-2}{2^t-2}\right) \geq 3$ when $s \geq 5 \cdot 2^{t-1} - 3$ so $\mathrm{PR}_t(s) \geq s + 3$ when $s \geq 5 \cdot 2^{t-1} - 3$.

1.3 Definitions

Recall that if φ is an ordered edge coloring of K_n with respect to the ordering $v_1, \ldots, v_n$ of its vertices, we say that φ' is the *φ-inherited coloring* (or just *inherited coloring*) if it is the vertex coloring of K_n defined by $\varphi'(v_i) = \varphi(v_i v_j)$ for $1 \leq i < j \leq n$ and $\varphi'(v_n) = \varphi'(v_{n-1})$. Given an ordering of $V(K_n)$, any vertex coloring φ' such that $\varphi'(v_{n-1}) = \varphi'(v_n)$ uniquely determines a corresponding ordered coloring. We define a *color class M_i of color i* to be the set of all vertices v where $\varphi'(v) = i$. In this paper, we shall always think of the ordered vertices as arranged on a horizontal line with v_i to the left of v_j if $i < j$. We say that an edge $v_i v_m$, $i < m$ goes from v_i to the right and from v_m to the left. If X is a (possibly empty) subset of $V(K_n)$, we say that the edge-coloring φ of K_n is

- *X-constant* if for any $v \in X$, $\varphi(vu) = \varphi(vw)$ for all $u, w \in V \setminus X$.
- *X-ordered* if it is X-constant and the vertices of X can be ordered $x_1, \ldots, x_m$ such that for each $i = 1, \ldots, m$, $\varphi(x_i x_p) = \varphi(x_i x_m) = \varphi(x_i w)$ for all $i < p \leq m$ and all $w \in V \setminus X$.

If Z is a nonempty subset of $V(K_n)$ we say φ is

- *Z-quasi-ordered* if

    1. φ is Z-constant.

2. Each vertex v_i in Z is incident to precisely $n-2$ edges of one color, which we call the *main color* of v_i, and one edge v_iv_j of another color, where $v_j \in Z$. If that other color is t, then v_j is incident to precisely $n-2$ edges of color t.

It is not hard to show that there are only two possibilities for the set Z in a Z-quasi-ordered coloring (see Figure 1.1):

1. $|Z| = 3$, the three vertices in Z have different main colors, and there is one edge in Z of each of these colors.

2. $|Z| = 4$, with two vertices u, v in Z with one main color, say 1, and two vertices y, z in Z with another main color, say 2, and $\varphi(uv) = \varphi(uy) = \varphi(vz) = 1, \varphi(yz) = \varphi(yv) = \varphi(zu) = 2$.

- *quasi-ordered* if it is Z-quasi-ordered for some subset Z of V and the restriction of φ to $V \setminus Z$ is ordered,

- *quasi-simply-ordered* if it is Z-quasi-ordered for some subset Z of V and the restriction of φ to $V \setminus Z$ is simply-ordered and does not use any of the main colors of Z,

- *nearly X-ordered* if it is Z-quasi-ordered and the restriction of φ to $V \setminus Z$ is T-ordered for some (possibly empty) subset T of $V \setminus Z$ and $X = Z \cup T$. (If φ is nearly X-ordered then one or two edges could be recolored to get an X-ordered coloring.)

It is easy to check that if φ is quasi-ordered (quasi-simply-ordered) for some set Z then if $|Z| = 3$ one edge can be recolored, and if $|Z| = 4$, then two edges can be recolored to get an ordered (simply-ordered) coloring.

To see this, suppose φ is Z-quasi-ordered and quasi-ordered (quasi-simply-ordered). Suppose $Z = \{x, y, z\}$ with x, y, z having main colors 1,2,3 respectively, and with $\varphi(xy) = 1$, $\varphi(yz) = 2$, $\varphi(zx) = 3$, as in (1) above (see Figure 1.1A). If we recolor zx so that $\varphi(zx) = 1$, then all edges incident with x have color 1, all edges incident with y, except xy, have color 2, and all edges incident with z, except zx and yz have color 3, so the modified coloring is ordered (simply-ordered). Suppose $Z = \{u, v, y, z\}$ with u and v having main color 1 and y and z having main color 2, with the colors of the edges in Z as in (2) above (see Figure 1.1B). If we recolor uz and vy so that $\varphi(uz) = \varphi(vy) = 1$, then all edges incident with u and v will have color 1 and all edges incident with y or z, but not incident with u or v have color 2, so the modified coloring will be ordered (simply-ordered).

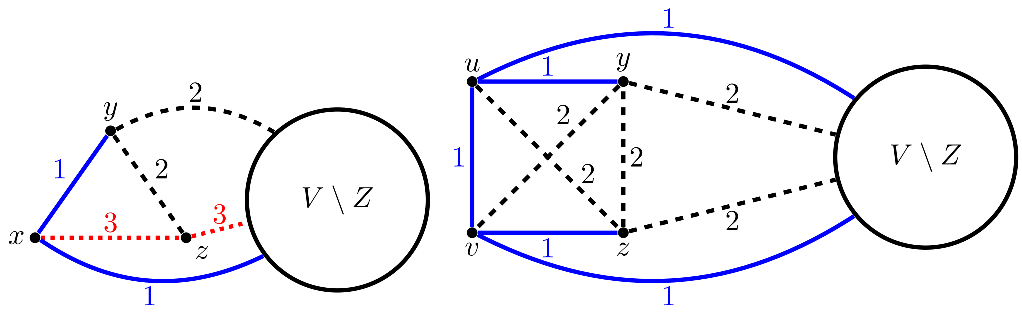


Figure 1.1: Z-quasi-orderings

The *maximum monochromatic degree* of an edge coloring of K_n is the maximum number of edges of the same color incident with a single vertex. If the maximum monochromatic degree of a coloring is d, and the vertex v is incident with d edges of color t, and the other $n-1-d$ edges incident with v have color s, we say v is a *t-max vertex* and also a *(t,s)-max vertex* with *majority color* t and *minority color* s.

We extend the notion of inherited coloring to quasi-ordered colorings as follows. If φ is a quasi-ordered coloring with ψ the ordered coloring which is the restriction of φ to $V \setminus Z$, we define φ', the φ-inherited coloring, by letting $\varphi'(x)$ equal the main color of x if $x \in Z$ and $\varphi(y) = \psi'(y)$ if $y \notin Z$. We think of the vertices in Z preceding those not in Z, in the order left to right, and if $|Z| = 4$ we list two vertices in Z with the same main color first, then the other two vertices with the same main color.

1.4 Ordering Lemmas

Let φ be an ordered edge coloring of K_n with vertex order $v_1, v_2, \ldots, v_n$, colors $1, \ldots, k$, and φ' be the inherited coloring of $V(K_n)$. For each $t \in [k]$ and $j \in [n]$, let M_t be a color class t of φ' and $M_t(j) = M_t \cap \{v_1, v_2, \ldots, v_j\}$. The next lemma is a key structural lemma that characterizes ordered polychromatic colorings.

Lemma 1.4.1. *Let $\varphi : E(K_n) \to [k]$ be an ordered or quasi-ordered coloring with vertex order $v_1, v_2 \ldots, v_n$.*

Then the following statements hold:

(I) φ is F_q-polychromatic $\iff \forall t \in [k]\ \exists j \in [n]$ such that $|M_t(j)| > \frac{j+q}{2}$,

(II) φ is C_q-polychromatic $\iff \forall t \in [k]$ either

(a) $\exists j \in [q+1, n-1]$ *such that* $|M_t(j)| \geq \frac{j+q}{2}$ *or*

(b) $q = 0$, φ *is* Z*-quasi-ordered and* t *is the color of some edge in* Z *or*

(c) $q = 1$, φ *is* Z*-quasi-ordered with* $|Z| = 4$ *and* t *is the color of some edge in* Z.

(III) φ *is* R_q*-polychromatic* $\Longleftrightarrow \forall t \in [k]$ *either*

(a) $\exists j \in [n]$ *such that*

(i) $|M_t(j)| > \frac{j+q}{2}$ *or*

(ii) $|M_t(j)| = \frac{j+q}{2}$ *and* $j \in \{2+q, n-2\}$ *or*

(iii) $|M_t(j)| = \frac{j+q}{2}$ *and* $|M_t(j+2)| = \frac{j+q+2}{2}$ *where* $j \in [4+q, n-3]$.

(b) $q = 0$, φ *is* Z*-quasi-ordered and* t *is the color of some edge in* Z

(c) $q = 1$, φ *is* Z*-quasi-ordered with* $|Z| = 4$ *and* t *is the color of some edge in* Z

Proof. Note that to prove the lemma, it is sufficient to consider an arbitrary color t and show for $\mathcal{H} \in \{F_q, C_q, R_q\}$ and for each $H \in \mathcal{H}$, that the given respective conditions are equivalent to H containing an edge of color t.

(I) Let j be an index such that $|M_t(j)| = m_j > (j+q)/2$ and let H be a 1-factor. Let $x_1, \ldots, x_{m_j}$ be the vertices of M_t in order and let $y_1, \ldots, y_{j-m_j}$ be the other vertices of $\{v_1, v_2, \ldots, v_j\}$ in order. Since $j - m_j < \frac{j-q}{2}$ and $m_j - q > \frac{j-q}{2}$, then at least one edge of H with an endpoint in $M_t(j)$ must go to the right, and thus, have color t.

On the other hand, by way of contradiction, assume that for each $j \in [n]$, $|M_t(j)| \leq (j+q)/2$. Letting $m = |M_t|$, we have $m \leq (n+q)/2$. Consider a 1-factor that spans all vertices except for q vertices in M_t. Let $x_1, \ldots, x_{m-q}$ be the $m-q$ vertices remaining from M_t in order and let $y_1, \ldots, y_{n-m}$, be the vertices outside of M_t in order. Note that since $m \leq (j+q)/2$, it follows that $n - m \geq m - q$ since if $n - m < m - q$ then $n < 2m - q$ and so $j > n$ which is impossible. Now, let H consist of the edges $x_1y_1, x_2y_2, \ldots, x_{m-q}y_{m-q}$ and a perfect matching on $\{y_{m-q+1}, \ldots, y_{n-m}\}$ (if this set is non-empty). We will show that y_i precedes x_i in the order $v_1, v_2, \ldots, v_n$ for each $i \in [m-q]$, so H has no edge of color t.

By way of contradiction, assume x_i precedes y_i for some $i \in [m-q]$. Letting $j = 2i - 1 + q$, y_i cannot be among the first j vertices in the order $v_1, v_2, \ldots, v_n$, because if it were there would be at least $i + q$ vertices of color t among these j vertices, so a total of at least $2i + q > j$ vertices. Hence

$$\frac{j+q}{2} = \frac{2i+2q-1}{2} < i + q \leq |M_t(j)| \leq \frac{j+q}{2}$$

which is impossible. Hence y_i precedes x_i for each i and φ is not F_q polychromatic.

(II) If t is a color such that (a) holds with strict inequality, the argument in (I) shows there is an edge of H with color t. If $|M_t(j)| = \frac{j+q}{2}$ for some $j \in [q+1, n-1]$ and every edge in H incident to a vertex in $M_t(j)$ goes to the left then, since each of these edges has its other vertex not in $M_t(j)$, H contains $\frac{j-q}{2}$ vertices in $M_t(j)$ and the same number not in $M_t(j)$. If $\frac{j-q}{2} = 1$, then the vertex in $M_t(j)$ is incident with at least one edge which goes to the right, and if $\frac{j-q}{2} \geq 2$ then H contains a 2-regular subgraph, which is impossible because an $n-q$ cycle cannot have a 2-regular subgraph on less than $n-q$ vertices.

If t is such that (b) holds, then note that t must be the main color of a vertex in Z and that the cycle must contain 2 edges incident with each vertex in Z. Any choice of these edges will contain an edge of color t since only one edge incident with each vertex in Z is not the main color of that vertex.

If t is such that (c) holds, then note that t must be the main color of a vertex in Z and any cycle on $n-1$ vertices must contain 2 edges incident with at least three of the four vertices in Z. Any choice of these edges will contain an edge of color t since only one edge incident with each vertex in Z is not the main color of that vertex.

On the other hand, suppose that for each $j \in [q+1, n-1]$, $|M_t(j)| = m < \frac{j+q}{2}$ and φ is not Z-quasi-ordered with t a main color. In particular, when $j = n-2$, we have that $|M_t(j)| = m < \frac{n+q}{2} - 1$. Consider a cycle that spans all vertices except for q vertices in M_t. Let $x_1, \ldots, x_{m-q}$ be the other $m-q$ vertices in M_t in order and $y_1, \ldots, y_{n-m}$ be the vertices outside of M_t in order. Note that if $m < \frac{j+q}{2}$, then $n-m > m-q$ since $n-m \leq m-q \implies j > n$ which is impossible. Consider the cycle $y_1 x_1 y_2 x_2 \cdots y_{m-q} x_{m-q} y_{m-q+1} \cdots y_{n-m} y_1$. Suppose y_i is to the right of x_i for some i. Then at most i of the first $j = 2i+q$ vertices are not in $M_t(j)$, so $|M_t(j)| \geq i+q = \frac{j+q}{2}$, which is impossible. Hence y_i and y_{i+1} are to the left of x_i for each $1 \leq i \leq m$, all edges of H incident to M_t go to the left, and thus are not of color t.

Observation. If H is a 2-regular subgraph that has no edge of color t, and M is any subset of M_t, then all edges of H incident to M go to the left, so at most half the vertices in H are in M_t and if $|M_t(j)| = \frac{j+q}{2}$, then of the first j vertices, precisely $j-q$ are in H, precisely half of these in M_t, and if $j-q \geq 4$ then these $j-q$ vertices induce a 2-regular subgraph of H.

(III) Let j be an index such that (III)(a) (i), (ii), or (iii) holds. Assume first that (i) holds, i.e., that $|M_t(j)| > \frac{j+q}{2}$ and let H be a 2-factor. Then the argument given in (I) shows that at least one edge of H with an endpoint in $M_t(j)$ must go to the right, and thus, have color t. Assume that (ii) holds. If $j = 2+q$, then M_t contains $q+1$ of the first $q+2$ vertices, so H contains a vertex in M_t which has an edge that goes to the right, so there is an edge of color t in H. If $j = n-2$ and H has no edges of color t, then (by the above observation) the subgraph of H induced by $[n-2]$ is a 2-regular graph spanning $n-2-q$ vertices. Since $\{v_{n-1}, v_n\}$ do not induce a cycle, H is not a 2-factor, a contradiction. Finally, assume that (iii) holds. If H does not have an edge of color t, then by the previous observation, H has a 2-regular subgraph spanning $j-q+2$ vertices, which has a 2-regular subgraph spanning $j-q$ vertices, which is impossible.

If (III)(b) or (III)(c) holds, by an argument identical to those for (II)(b) and (II)(c), H has an edge of color t.

On the other hand, suppose that none of (III)(a), (III)(b), or (III)(c) hold. We shall construct a 2-factor that does not have an edge of color t. If $|M_t(j)| < \frac{j+q}{2}$ for each $j \in [q+1, n-1]$, then there is a cycle with no color t edge as described in (II). If not, let $i_1, i_2, \ldots, i_k$ be the values of j in $[4+q, n-3]$ for which $|M_t(j)| = \frac{j+q}{2}$. Since (III)(a)(iii) is not satisfied, $i_{q+1} - i_q$ is at least 4 and even for $q = 1, 2, \ldots, k-1$. As before, suppose there are m vertices of color t. Let $x_1, x_2, \ldots, x_{m-q}$ be the last $m-q$ of these, in order, and let $y_1, y_2, \ldots, y_{n-m}$ be the other vertices, in order. Note that since $m \leq \frac{n+q}{2}$ we have $m - q \leq \frac{n-q}{2}$ and $n - m \geq \frac{n-q}{2}$. For each q in $[1, k-1]$, moving left to right within the interval $[i_q + 1, i_{q+1}]$, there are always more y's than x's (except an equal number of each at the end of the interval), since otherwise there would have been another value of j between i_q and i_{q+1} where $|M_t(j)| = \frac{j+q}{2}$. Form an $(i_{q+1} - i_q)$-cycle by alternately taking y's and x's, starting with the y with the smallest subscript. Also form an $i_1 - q$ cycle using the first $\frac{i_1-q}{2}$ y's and the same number of x's, and an $n - i_k$ cycle at the end, first alternating the y's and x's, putting any excess y's at the end.

■

Lemma 1.4.2. *Let $\mathcal{H} \in \{F_q, R_q, C_q\}$. If there exists an ordered (quasi-ordered) $\mathcal{H}$-polychromatic coloring of K_n with k colors, then there exists one which is simply-ordered (quasi-simply-ordered) with k colors.*

Proof. Let $V(K_n) = [n]$ with the natural order. If c' is a coloring of $[n]$, a *block* of c' is a maximal interval of integers from $[n]$ which all have the same color. So a simply-ordered k-polychromatic coloring has precisely k blocks. We define a *block shift operation* as follows. Assume that $t \in [k]$ is a color for which there are at least 2 blocks. Let $j(t) = j$ be the smallest integer so that $M_t(j) > (j+q)/2$ if such exists. If there is a block $[m, s]$ in M_t where $m > j$, delete this block, then take the color of the last vertex in the remaining sequence, and add $s - m + 1$ more vertices with this color at the end of the sequence. If each block of color t has its smallest element less than or equal to j, consider the block B of color t that contains j and consider another block B_1 of color

t that is strictly to the left of B. Form a new coloring by "moving" B_1 next to B. We see that the resulting coloring has at least one less block.

Let c be a ordered (quasi-ordered) F_q-polychromatic coloring of K_n on vertex set $[n]$ with k colors such that the inherited vertex coloring c' has the smallest possible number of blocks. Assume that color t has at least 2 blocks. Let $j(t) = j$ be the smallest integer so that $M_t(j) > (j+q)/2$. Such j exists by Lemma 1.4.1(I), and the color of j is t. Apply the block shifting operation. The condition from part (I) of Lemma 1.4.1 is still valid for all color classes, so the new coloring is F_q-polychromatic using k colors. This contradicts the choice of c having the smallest number of blocks.

If c is an ordered (quasi-ordered) C_q-polychromatic coloring of K_n, an argument very similar to the one above shows if (II)(a), (b), or (c) hold, there exists a simply-ordered (quasi-simply-ordered) coloring that uses the same number of colors and that is C_q-polychromatic.

Finally, let c be an ordered (quasi-ordered) R_q-polychromatic coloring of K_n on vertex set $[n]$ with k colors such that the inherited vertex coloring c' has the minimum possible number of blocks. Assume that $t \in [k]$ is a color for which there are at least 2 blocks. If (III)(b) or (III)(c) hold, then the block shifting operation gives a coloring that is still R_q-polychromatic with the same number of colors and fewer blocks.

Thus, by Lemma 1.4.1(III) there exists j such that

(1) $|M_t(j)| > (j+q)/2$ or

(2) $|M_t(2+q)| = 1+q$ or

(3) $|M_t(n-2)| = (n+q-2)/2$ or

(4) $|M_t(n-1)| = (n+q-1)/2$ or

(5) $|M_t(j)| = (j+q)/2$ and $|M_t(j+2)| = (j+q+2)/2$ and $4+q \le j \le n-3$.

If (1) holds, then we apply the block shifting operation and observe, as in the case of F_q, that the resulting coloring is still R_q-polychromatic with the same number of colors and fewer blocks. The case when (2) applies is similar.

Assume neither (1) nor (2) holds. If (3) holds then, since $c'(v_{n-1}) = c'(v_n)$, neither v_{n-1} nor v_n can have color t. Hence there is another block of color t vertices to the left of the one containing v_{n-2}, so we can do a block shift operation ot reduce the number of blocks, a contradiction.

The same argument works if (4) holds.

Finally, assume that none of (1)–(4) holds, but (5) holds. This implies that $c'(j) = c'(j+2) = t$ and $c'(j+1) = u \neq t$. Now define c'' by $c''(i) = c'(i)$ if $i \notin \{j+1, j+$

$2\}, c''(j+1) = t$, and $c''(j+2) = u$. Clearly c'' has at least one fewer block than c'. Since $j+q+1$ is odd, the only situation where c'' would not be R_q-polychromatic is if $M_u(j+1) > \frac{j+q+1}{2}$. However, then $|M_u(j-1)| = |M_u(j+1)| - 1 > \frac{j+q-1}{2}$, so c'' is R_q-polychromatic after all. ■

1.5 Proof of Theorem 1.2.1 on Matchings

We prove Theorem 1.2.1. This proof is similar to the proof of Theorem 1 in [4]. Let $k = \text{poly}_{F_q}(n)$ be the polychromatic number for 1-factors spanning $n-q$ vertices in $G = K_n = (V, E)$. Among all F_q-polychromatic colorings of K_n with k colors we choose ones that are X-ordered for a subset X (possibly empty) of the largest possible size, and, of these, choose a coloring c whose restriction to $V \setminus X$ has the largest possible maximum monochromatic degree. Let v be a vertex of maximum monochromatic degree, r, in c restricted to $G[V \setminus X]$, let the majority color on the edges incident to v in $V \setminus X$ be color 1. By the maximality of $|X|$, there is a vertex u in $V \setminus X$ such that $c(uv) \neq 1$. Assume $c(uv) = 2$. If every 1-factor spanning $n-q$ vertices containing uv had another edge of color 2, then the color of uv could be changed to 1, resulting in a F_q-polychromatic coloring where v has a larger maximum monochromatic degree in $V \setminus X$, a contradiction. Hence, there is a 1-factor F spanning $n-q$ vertices in which uv is the only edge with color 2 in c.

Let $c(vy_i) = 1$, $y_i \in V \setminus X$, $i = 1, \ldots, r$. Note that for each $k \in [r]$, y_k must be in F. If not, then $F - uv + vy_k$ is a 1-factor spanning $n-q$ vertices with no edge of color 2 (since uv was the unique edge of color 2 in F and vy_k is color 1). For each $i \in [r]$, let y_iw_i be the edge of F containing y_i (perhaps $w_i = y_j$ for some $j \neq i$). See Figure 1.2. We can get a different 1-factor F_i by replacing the edges uv and y_iw_i in F with edges vy_i and uw_i. Since F_i must have an edge of color 2 and $c(vy_i) = 1$, we must have $c(uw_i) = 2$ for each $i \in [r]$.

If $w_i \in X$ for some i then, since c is X-constant, $c(w_iy_i) = c(w_iu) = 2$, so y_iw_i and uv are two edges of color 2 in F, a contradiction. So, $w_i \in V \setminus X$. Thus $c(uv) = c(uw_1) = \cdots = c(uw_r) = 2$, and the monochromatic degree of u in $V \setminus X$ is at least $r+1$, larger than that of v, a contradiction. Hence $X = V$, c is ordered, and, by Lemma 1.4.2, there exists a simply-ordered F_q-polychromatic coloring c_s with k colors. ■

A formula for $\text{poly}_{F_q}(n)$ appears in Section 1.9.

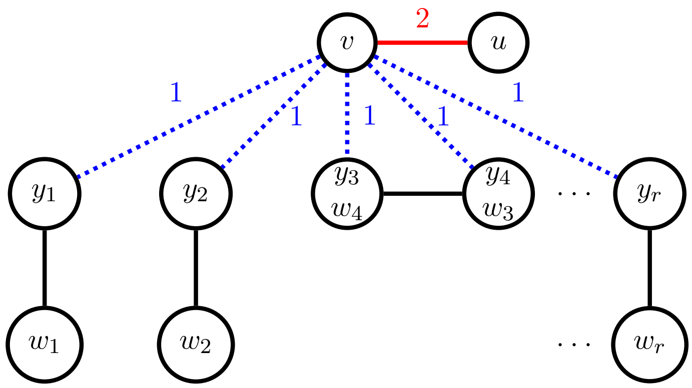


Figure 1.2: Maximum polychromatic degree in an F_q-polychromatic coloring

1.6 C_q-polychromatic Numbers 1 and 2

The following theorem is a special case of a theorem of Faudree and Schelp.

Theorem 1.6.1. *[23] Let $s \geq 5$ be an integer and let $c(s)$ denote the smallest integer n such that in any 2-coloring of the edges of K_n there is a monochromatic s-cycle. Then $c(s) = 2s - 1$ if s is odd and $c(s) = \frac{3}{2}s - 1$ if s is even.*

Faudree and Schelp actually determined all values of $c(r, s)$, the smallest integer n such that in any coloring of the edges of K_n with red and blue, there is either a red r-cycle or a blue s-cycle. Their theorem extended partial results and confirmed conjectures of Bondy and Erdős [11] and Chartrand and Schuster [13] (who showed $c(3) = c(4) = 6$). The coloring of K_{2s-2} to prove the lower bound for s odd is a copy of $K_{s-1,s-1}$ of red edges with all other edges blue, while for s even it's a red $K_{\frac{s}{2}-1,s-1}$ with all other edges blue.

Proof of Theorem 1.2.5. By Theorem 1.6.1, if $s \geq 5$ is odd then there is a polychromatic 2-coloring of K_n if and only if $n \leq 2s-2 = 2(n-q)-2$, so if and only if $n \geq 2q+2$. If $s \geq 5$ is even then there is a polychromatic 2-coloring if and only if $n \leq \frac{3}{2}s - 2 = \frac{3}{2}(n-q) - 2$, so if and only if $n \geq 3q + 4$. Hence if $n \in [2q + 2, 3q + 2]$ then $\varphi_{\mathrm{C_q}}(n) = 1$ if $n - q$ is even and $\varphi_{\mathrm{C_q}}(n) = 2$ if $n - q$ is odd. The smallest value of n for which there is a simply-ordered C_q-polychromatic 2-coloring is $n = 3q + 3$, so there does not exist one if $n - q$ is odd and $n \leq 3q + 2$. ■

We remark that the only values for $q \geq 2$ and n such that there is no optimal simply-ordered $C_q(n)$-polychromatic coloring of K_n are the ones given in Theorem 1.2.5

($n \in [2q+2, 3q+2]$ and $n-q$ is odd), and $q = 2$, $n = 5$ (two monochromatic C_5's is a coloring of K_5 with no monochromatic C_3's).

1.7 Proofs of Theorem 1.2.6 and Lemmas on Long Cycles

We will need some results on the existence of long cycles in bipartite graphs.

Theorem 1.7.1 (Jackson [34]). *Let G be a connected bipartite graph with bipartition $V(G) = S \cup T$ where $|S| = s$, $|T| = t$, and $s \leq t$. Let m be the minimum degree of a vertex in S and p be the minimum degree of a vertex in T. Then G has a cycle with length at least* $\min\{2s, 2(m+p-1)\}$.

Theorem 1.7.2 (Rahman, Kaykobad, Kaykobad [41]). *Let G be a connected m-regular bipartite graph with $4m$ vertices. Then G has a Hamiltonian cycle.*

Lemma 1.7.3. *Let B be a bipartite graph with vertex bipartition S, T where $|S| = s$, $|T| = t$, and $s \leq t$. Suppose each vertex in T has degree m and each vertex in S has degree $t - m$. Then B has a $2s$-cycle unless $s = t = 2m$ and B is the disjoint union of two copies of $K_{m,m}$.*

Proof. Suppose $s < t$. Summing degrees in S and T gives us $s(t-m) = tm$, so

$$m = \frac{st}{s+t} > \frac{st}{2t} = \frac{s}{2}$$

so B is connected. By Theorem 1.7.1, B has a $2s$-cycle, since $2[m + (t-m) - 1] = 2(t-1) \geq 2s$. If $s = t$, then B is an m-regular graph with $4m$ vertices. If B is connected then, by Theorem 1.7.2, it has a $2s$-cycle. If B is not connected then clearly it is the disjoint union of two copies of $K_{m,m}$. ■

We say that a cycle H' of length $n - q$ is obtained from a cycle H of length $n - q$ by a *twist* of disjoint edges e_1 and e_2 of H if $E(H) \setminus \{e_1, e_2\} \subseteq E(H')$, i.e. we remove e_1, e_2 from H and introduce two new edges to make the resulting graph a cycle. Note that the choice of the two edges to add is unique (due to connectedness), however, both choices would result in a 2-regular subgraph.

One main difference between the definitions of $C_q(n)$ and $R_q(n)$ is that for the former, we consider only cycles of length precisely $n - q$, whereas, in the latter, we consider all 2-regular subgraphs spanning *at least* $n - q$ vertices. This is because we can prove Theorem 1.2.6 for cycles, however, a similar result for 2-regular subgraphs remains elusive (see Conjecture 1.11.1).

1.7.1 Proof of Theorem 1.2.6

Suppose not. Let m be an integer in $[j, n-1]$ such that every m-cycle gets all colors but there is an $(m+1)$-cycle $H = v_1v_2\cdots v_{m+1}v_1$ which does not have an edge of color t. Then $c(v_iv_{i+2}) = t$ for all i, where the subscripts are read mod $(m+1)$, because otherwise, there is an m-cycle with no edge of color t.

Case 1. If $m+1$ is odd, then $v_1v_3v_5\cdots v_{m+1}v_2v_4\cdots v_{m-2}v_1$ is an m-cycle with at most two colors, since all edges except possibly $v_{m-2}v_1$ have color t. This is impossible.

Case 2. Suppose $m+1$ is even. Then $c_E = v_2v_4\cdots v_{m+1}v_2$ and $c_O = v_1v_3\cdots v_mv_1$ are $\frac{m+1}{2}$-cycles with all edges of color t. Suppose H has a chord v_jv_{j+r} with color t for some j and odd integer r in $[3, m-2]$. Then $v_{j+2}v_{j+4}\cdots v_{j-2}v_jv_{j+r}v_{j+r+2}\cdots v_{j+r-4}$ is a path with m vertices (missing v_{j+r-2}) and all edges of color t, so there is an m-cycle with at most two colors, which is impossible. Hence if v_i is a vertex in c_E and v_j is a vertex in c_O, then $v(v_iv_j) \neq t$.

We claim that for each j and even integer s, $c(v_jv_{j+s}) = t$. If not, then $v_jv_{j+s}v_{j+s+1}\cdots v_{j-3}v_{j-2}v_{j+s-1}v_{j+s-2}\cdots v_{j+1}v_j$ is an m-cycle (missing v_{j-1}) with no edge of color t (note $c(v_{j-2}v_{j+s-1}) \neq t$ because $j-2$ and $j+s-1$ have different parities). Hence, the vertices of c_E and c_O each induce a complete graph with $\frac{m+1}{2}$ vertices and all edges of color t, and there are no other edges of color t in K_n.

If there is a color w, different than t, such that there exist two disjoint edges of color w, then it is easy to find an m-cycle with two edges of color w and the rest of color t. If there do not exist two such edges of color w, then all edges of color w are incident to a single vertex x, so any m-cycle with x incident to two edges of color t does not contain an edge of color w (these exist since $\frac{m+1}{2} \geq 3$). ■

We remark that the statement in Theorem 1.2.6 would be false without the requirement that there be at least three colors. If $m \geq 3$ is odd, then two vertex disjoint complete graphs each with $\frac{m+1}{2}$ vertices and all edges of color t with all edges between them of color w has an $(m+1)$-cycle with all edges of color w, while every m-cycle has edges of both colors. This is the reason for the difference between odd and even values of $n-q$ in Theorem 1.2.5. The statement would also be false with three colors if $j = 3$ and $n = 4$.

1.8 Main Lemmas and Proofs of Theorems

We now state and prove the three main lemmas needed for the proofs of Theorems 1.2.2, 1.2.3, and 1.2.4.

Lemma 1.8.1.

(a) *Let $\mathcal{H} \in \{R_q(n), C_q^*(n)\}$. Of all optimal $\mathcal{H}$-polychromatic colorings, let φ be one which is X-ordered on a (possibly empty) subset X of $V(K_n)$ of maximum size and, of these, such that $G_m = K_n[Y]$ has a vertex $v \in Y$ of maximum possible monochromatic degree d in G_m where $Y = V(K_n) \setminus X$, $|Y| = m$, and $d < (m-1)$. If v is incident in G_m to d edges of color 1 and $u \in Y$ is such that $\varphi(vu) = 2$, then v is a $(1,2)$-max vertex in G_m and u is a $(2,t)$-max vertex in G_m for some color t (possibly $t = 1$).*

(b) *The same is true if $X \neq \emptyset$ and φ is nearly X-ordered.*

Proof of (a). Let $y_1, y_2, \ldots, y_d \in Y$ be such that $\varphi(vy_i) = 1$. Let $H \in C_q^*$ or $H \in R_q$ be such that uv is the only edge of color 2. There must be such an H otherwise we could change the color of uv from 2 to 1, giving an $\mathcal{H}$-polychromatic coloring with monochromatic degree greater than d in G_m. Orient the edges of H to get a directed cycle or 2-regular graph H' where $\overrightarrow{uv}$ is an arc.

If $y_i \in H'$ then the predecessor w_i of y_i in H' must be such that $\varphi(w_i u) = 2$, because otherwise we can twist uv and $w_i y_i$ to get an $(n-q)$-cycle (if $H \in C_q^*$) or a 2-regular graph (if $H \in R_q$) with no edge of color 2. Note that w_i must be in Y because otherwise, since φ is X-constant, $\varphi(w_i u) = \varphi(w_i y_i) = 2$, contradicting the assumption that uv is the only edge in H of color 2.

Suppose $y_i \notin H$ for some $i \in [d]$. If $\varphi(y_i u) \neq 2$, then $J = (H \setminus \{uv\}) \cup \{vy_i, y_i u\}$ has no edge of color 2. This is impossible if $H \in R_q$, because J is a 2-regular graph spanning $n-q+1$ vertices. If $H \in C_q^*$, then J is an $(n-q+1)$-cycle with no edge of color 2, so by Theorem 1.2.6, since the polychromatic number of H is at least 3, there exists an $(n-q)$-cycle which is not polychromatic, a contradiction. Hence $\varphi(y_i u) = 2$ in either case.

Thus, for each $i \in [d]$, either $y_i \notin H$ and $\varphi(y_i u) = 2$, or $y_i \in H$ and $\varphi(w_i u) = 2$ where w_i is the predecessor of y_i in H'. That gives us d edges in G_m of color 2 which are incident to u. Since v has maximum monochromatic degree in G_m, it follows that $v = w_i$ for some i (otherwise uv is a different edge of color 2 incident to u) and it also follows that no edge in G_m incident to v can have color t where $t \notin \{1, 2\}$. This is because if vz were such an edge, as shown above, then either $z \in H$ and $\varphi(w'u = 2)$ where w' is the predecessor of z in H', or $z \notin H$ and $\varphi(zu) = 2$. In either case we get $d+1$ edges of color 2 in G_m incident to u, a contradiction. So v is a $(1,2)$-max vertex and u is a $(2,t)$-max vertex for some color t.

The proof of (b) is exactly the same. ■

Lemma 1.8.2. *Let $n \geq 7$ and $\mathcal{H} \in \{R_q(n), C_q(n)\}$. If there does not exist an optimal $\mathcal{H}$-polychromatic coloring of K_n with maximum monochromatic degree $n-1$, then one* *of the following holds.*

(a) $\mathcal{H} = C_q(n)$, *$n-q$ is odd and* $n \in [2q+2, 3q+2]$ *(and* $\varphi_{\mathrm{C_q}}(n) = 2$*).*

(b) $q = 0$ *and there exists an optimal* $\mathcal{H}$*-polychromatic coloring which is* Z*-quasi-ordered with* $|Z| = 3$*.*

(c) $q = 1$ *and there exists an optimal* $\mathcal{H}$*-polychromatic coloring which is* Z*-quasi-ordered with* $|Z| = 4$*.*

Proof. First assume that $\mathcal{H} = C_q(n)$ and that $q \geq 2$ and n are such that $\varphi_{\mathrm{C_q}}(n) \leq 2$. If $n-q$ is even then, by Theorem 1.2.5, there is a C_q-polychromatic 2-coloring if and only if $n \geq 3q+3$. Since $3q+3$ is the smallest value of n such that the simply-ordered C_q-polychromatic coloring φ_{C_q} uses two colors, if $\varphi_{\mathrm{C_q}}(n) \leq 2$ and $n-q$ is even, then there is an optimal simply-ordered C_q-polychromatic coloring, and this coloring has a vertex (in fact $q+1$ of them) with monochromatic degree $n-1$.

If $n-q$ is odd then, by Theorem 1.2.5, there is a C_q-polychromatic 2-coloring if and only if $n \geq 2q+2$. Since there is a simply-ordered C_q-polychromatic 2-coloring if $n \geq 3q+3$, that means that if $n-q$ is odd, $\varphi_{\mathrm{C_q}}(n) \leq 2$ and $n \notin [2q+2, 3q+2]$ then there is a simply-ordered C_q-polychromatic coloring. Thus if $\varphi_{\mathrm{C_q}}(n) \leq 2$, there is an optimal simply-ordered C_q-polychromatic coloring, and hence one with maximum monochromatic degree $n-1$, unless $n-q$ is odd and $n \in [2q+2, 3q+2]$, which are the conditions for (a).

Now let $\mathcal{H} \in \{R_q(n), C_q^*(n)\}$ and suppose there does not exist an optimal $\mathcal{H}$-polychromatic coloring of K_n with maximum monochromatic degree $n-1$. Of all optimal $\mathcal{H}$-polychromatic colorings of K_n, let φ be the one with maximum possible monochromatic degree d (so $d < n-1$).

Claim 1. $d > \frac{n-1}{2}$.

Proof. Since there are only two colors at a max vertex, certainly $d \geq \frac{n-1}{2}$. Assume $d = \frac{n-1}{2}$ (so n is odd) and that x is a -max vertex where colors i and j appear. Then x is both an i-max and j-max vertex so, by Lemma 1.8.1, each vertex in V is a -max vertex.

Suppose there are more than 3 colors, say colors i, j, s, t are all used. If i and j appear at x then no vertex y can have colors s and t, because there is no color for xy. So the sets of colors on the vertices is an intersecting family of 2-sets. Since there are at least 4 colors, the only way this can happen is if some color, say i, appears at every vertex. Let n_{ij}, n_{is}, and n_{it} be the number of (i,j)-max, (i,s)-max, and (i,t)-max vertices with $n_{ij} \leq n_{is} \leq n_{it}$. Then $n_{ij} < \frac{n}{2}$ (in fact, $n_{ij} \leq \frac{n}{3}$). If x is an (i,j)-max vertex and y is an (i,s)-max vertex, then $c(xy) = i$. Hence the number of edges of color j incident to x is at most $n_{ij} - 1 < \frac{n-2}{2} < d$, a contradiction.

Now suppose there are precisely 3 colors. Let A, B, C be the set of all $(1,2)$-max, $(2,3)$-max, and $(1,3)$-max vertices, respectively, with $|A| = a, |B| = b$, and $|C| = c$. All edges from a vertex in A to a vertex in B have color 2, from B to C have color 3, from A to C have color 1; internal edges in A have color 1 or 2, in B have color 2 or 3, in C have color 1 or 3. We clearly cannot have a, b, or c greater than $\frac{n-1}{2}$ so, without loss of generality, we can assume $a \leq b \leq c \leq \frac{n-1}{2}$ and $a + b + c = n$.

Consider the graph F formed by the edges of color 1 or 2. Vertices of F in B or C have degree $\frac{n-1}{2}$, while vertices in A have degree $n-1$. Since $a \leq c$ we have $a \leq \frac{n-b}{2}$. The internal degree in F of each vertex in B is $\frac{n-1}{2} - a \geq \frac{n-1}{2} - \frac{n-b}{2} = \frac{b-1}{2}$. As is well known (Dirac's theorem), that means there is a Hamiltonian path within B. Similarly there is one within C. If $a \geq 2$, that makes it easy to construct a Hamiltonian cycle in F. If $a = 1$ we must have $b = c = \frac{n-1}{2}$, so F is two complete graphs of size $\frac{n+1}{2}$ which share one vertex. This graph has a spanning 2-regular subgraph if $n \geq 7$ (a 3-cycle and a 4-cycle if $n = 7$), so no R_q-polychromatic coloring with 3 colors for any $q \geq 0$ if $n \geq 7$.

If $a = 1$ and $b = c = \frac{n-1}{2}$ consider the subgraph of all edges of colors 1 or 3. It consists of a complete bipartite graph with vertex parts $A \cup B$ and C, with sizes $\frac{n+1}{2}$ and $\frac{n-1}{2}$, plus internal edges in C. Clearly this graph has an $(n-1)$-cycle, but no Hamiltonian cycle. Hence there can be a C_q-polychromatic 3-coloring only if $q = 0$. However, the C_0-polychromatic coloring φ_{C_0} uses at least 4 colors if $n \geq 7$, so there is no optimal one with maximum monochromatic degree $\frac{n-1}{2}$.

Claim 2. If $q = 0$, then, up to relabeling the colors, there is a $(1,2)$-max vertex, a $(2,3)$-max vertex and a $(3,1)$-max vertex.

Proof. Assume that every -max vertex has majority color either 1 or 2. Then u must be a $(2,1)$-max vertex. This is because by Lemma 1.8.1, if it were a $(2,t)$-max vertex for some third color t, and $c(uz) = t$, then z would have to be a t-max vertex, a contradiction. Hence, every -max vertex is either a $(1,2)$-max vertex or a $(2,1)$-max vertex. Let S be the set of all $(1,2)$-max-vertices, T be the set of all $(2,1)$-max-vertices, and $W = V \setminus (S \cup T)$. Edges within S and from S to W must have color 1 (because any minority color edge at a -max vertex is incident to a max vertex of that color), edges within T and from T to W must have color 2, and all edges between S and T must have color 1 or 2. If $|S| = s$ and $|T| = t$ and $m = n - 1 - d$, then each vertex in S is adjacent to m vertices in T by edges of color 2 (and adjacent to $t - m$ vertices in T by edges of color 1), and each vertex in T is adjacent to m vertices in S by edges of color 1.

Suppose $s < t$ and consider any edge ab from S to T of color 2. As before, there is an $H \in \mathcal{H}$ which contains ab, but no other edges of color 2. Hence H has no edges from T to W. Since $s < t$ there must be an edge of H with both vertices in T, so it does have another edge of color 2 after all, a contradiction. The same argument works if $t < s$ with an edge with color 1. To avoid this, we must have $s = t = 2m$. If there is an edge from S to W then, again, H has an internal edge in T, which is impossible. Hence if $\mathcal{H} = C_0^*$ then $W = \emptyset$ and every edge has color 1 or 2, which is impossible since H has at least 3 colors. If $\mathcal{H} = R_0$ then the subgraph of H induced by $S \cup T$ is the union of cycles. If $m = 1$ then $S \cup T$ induces a 4-cycle in H, two edges of each color, so ab is not the only edge with color 2. If $m \geq 2$ then two applications of Hall's Theorem gives two disjoint perfect matchings of edges of color 1 between S and T, whose union is a 2-factor of edges of color 1 spanning $S \cup T$, which together with the subgraph of H induced by W, produces a 2-factor $H' \in R_0$ with no edge of color 2.

We have shown that u is not a $(2,1)$-max vertex, so it must be a $(2,3)$-max vertex for some other color 3. Say $\varphi(uz) = 3$. Then, by Lemma 1.8.1, z is a 3-max vertex. If $\varphi(vz) = 2$, then z would be a 2-max vertex. So z would be both a 2-max and a 3-max vertex, and so $d = \frac{n-1}{2}$, a contradiction to Claim 1. Hence $\varphi(vz) = 1$, which means z must be a $(3,1)$-max vertex.

Claim 3. If $q = 0$ then V can be partitioned into sets A, B, D, E where the following properties hold (see Figure 1.3).

1. All vertices in A are $(1,2)$-max-vertices.
2. All vertices in B are $(2,3)$-max-vertices.
3. All vertices in D are $(3,1)$-max-vertices.
4. No vertex in E is a -max vertex.
5. All edges within A, from A to D, and from A to E have color 1.
6. All edges within B, from B to A, and from B to E have color 2.
7. All edges within D, from D to B, and from D to E have color 3.
8. $|A| = |B| = |D| = m = n - 1 - d$.

Proof. Let $A = \{x : x \text{ is a } (1,2)\text{-max vertex}\}$, $B = \{x : x \text{ is a } (2,3)\text{-max vertex}\}$, $D = \{x : x \text{ is a } (3,1)\text{-max vertex}\}$ and $E = V \setminus (A \cup B \cup D)$. Let $x \in A$. If $y \in A$, then $\varphi(xy) = 1$ because if $\varphi(xy) = 2$, then y would be a 2-max vertex. If $y \in B$, then

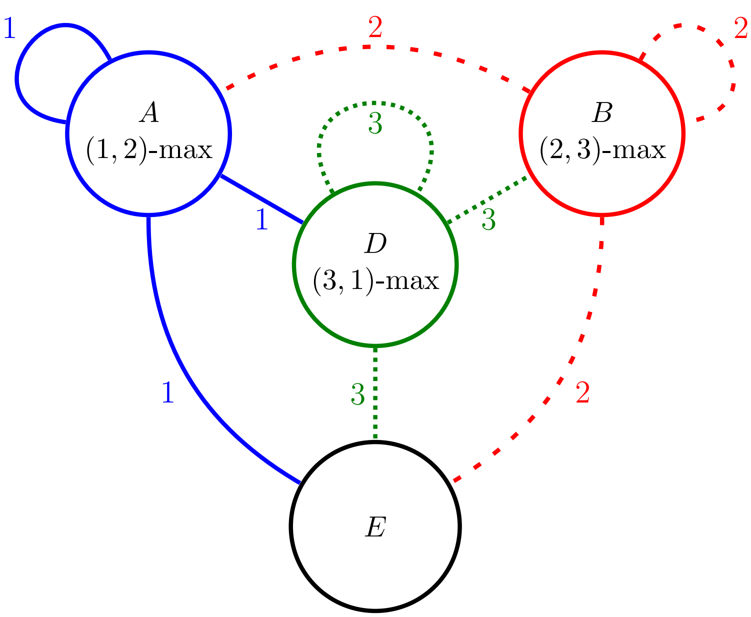


Figure 1.3: A C_0-polychromatic and R_0-polychromatic coloring.

$\varphi(xy) = 2$ because that is the only possible color for an edge incident to x and y and, similarly, if $y \in D$, then $\varphi(xy) = 1$.

Suppose w is a -max vertex in E. Then the two colors on edges incident to w must be a subset of $\{1, 2, 3\}$, because, otherwise, it would be disjoint from $\{1, 2\}$, $\{2, 3\}$, or $\{1, 3\}$, so there would be an edge incident to w for which there is no color. Say 1 and 2 are the colors at w. Since $w \notin A$, w is a $(2, 1)$-max vertex. Let z be a $(3, 1)$-max vertex. Then the edge wz must have color 1 so, by Lemma 1.8.1, z is a 1-max vertex, a contradiction. We have now verified (1)–(4). If $x \in A$ and $w \in E$ then $\varphi(xw) = 1$ because if $\varphi(xw) = 2$ then w would be a 2-max vertex. Similar arguments show that if $y \in B$ then $\varphi(yw) = 2$ and if $y \in D$ then $\varphi(yw) = 3$. We have now verified (1)–(7).

We have shown that if x is in A then $\varphi(xy) = 2$ if and only if $y \in B$. That means $|B| = m$, and by the same argument $|A| = |C| = m$ as well, completing the proof of Claim 3.

Claim 4. If $\mathcal{H} \in \{C_0^*, R_0\}$, and there exists an optimal $\mathcal{H}$-polychromatic coloring satisfying (1)–(8) with $m > 1$, then there exists one with $m = 1$, i.e. one that is Z-quasi-ordered with $|Z| = 3$.

Proof. Let $A = \{a_i : i \in [m]\}, B = \{b_i : i \in [m]\}, D = \{d_i : i \in [m]\}$. Define an edge coloring γ by

$$\begin{aligned}\gamma(a_1b_i) &= 1 \text{ if } i > 1\\ \gamma(b_1d_i) &= 2 \text{ if } i > 1\\ \gamma(d_1a_i) &= 3 \text{ if } i > 1\\ \gamma(uv) &= \varphi(uv) \text{ for all other } u, v \in V.\end{aligned}$$

It is easy to check that γ has the structure described above with $m = 1$. We have essentially moved $m-1$ vertices from each of A, B, and D, to E. Since a_1, b_1, and c_1 each have monochromatic degree $n-2$, any 2-factor must have edges of colors 1,2, and 3 under the coloring γ, so if it had all colors under φ, it still does under γ. ■

We remark that the coloring γ with $m = 1$ in Claim 4 is Z-quasi-ordered with $|Z| = 3$. As we have shown, if there exists such an R_0-polychromatic coloring φ with $m > 1$, then there exists one with $m = 1$. However, if $m > 1$ and $n > 6$, a coloring φ satisfying properties (1)–(8) might not be R_0-polychromatic. This is because if E has no internal edges with color 1, then any 2-factor with a $2m$-cycle consisting of alternating vertices from A and B has no edge with color 1. However, the modified coloring γ (with $m = 1$) is an R_0-polychromatic coloring because then colors 1, 2, and 3 must appear in any 2-factor.

Claim 5. If $q \geq 1$ then, up to relabelling colors, every max vertex is a $(1,2)$-max vertex or a $(2,1)$-max vertex.

Proof. As before, we assume v is a $(1,2)$-max vertex, that $\varphi(uv) = 2$ and that $H \in R_q$ (or $H \in C_q^*$) is such that uv is the only edge of color 2. We know that u is a $(2,t)$-max vertex for some color t. By way of contradiction, suppose u is a $(2,3)$-max vertex. Then we have the configuration of Figure 1.3, with $|A| = |B| = |D| = m$. If uw is also an edge of H then $w \in D$, since otherwise $\varphi(uw) = 2$. Let Q be the set of vertices not in H (so $|Q| = q > 0$) and suppose $p \in Q$ but $p \notin B$. Then we can replace u in H with p to get a 2-regular graph (cycle) with no edge of color 2. Hence $Q \subseteq B$. Orient the edges of H to get a directed graph H' where $\overrightarrow{uv}$ is an arc. Since $|B \setminus Q| < |D|$, and every vertex in D appears in H', for some $d \in D$ and $e \notin B$, $\overrightarrow{de}$ is an arc in H'. Since $\varphi(du) = 3$ and $\varphi(ev) = 1$, when you twist uv and de you get a 2-regular graph (cycle) with no edge of color 2, a contradiction. Hence every -max vertex is a $(1,2)$-max vertex or $(2,1)$-max vertex.

Claim 6. If $q = 1$ then, up to relabelling colors, the vertex set can be partitioned into S, T, W such that

1. S is the set of all $(1, 2)$-max vertices
2. T is the set of all $(2, 1)$-max vertices
3. W has no max vertices
4. All internal edges in S and all edges from S to W have color 1; all internal edges in T and all edges from T to W have color 2
5. The edges of color 1 between S and T form two disjoint copies of $K_{m,m}$, as do the edges of color 2 (so $|S| = |T| = 2m$, where $n - m - 1$ is the maximum monochromatic degree)

Proof. By Claim 5, if $q \geq 1$, then every max vertex is a $(1, 2)$ or $(2, 1)$-max vertex.

Let S be the set of all $(1, 2)$-max vertices and T be the set of all $(2, 1)$-max vertices, with $|S| = s$ and $|T| = t$, $s \leq t$, and let m be the maximum monochromatic degree. Let $W = V(G) \setminus (S \cup T)$ and let B be the complete bipartite graph with vertex bipartition S, T and edges colored as they are in G. So each vertex of B in S is incident with m edges of color 2 and $t - m$ edges of color 1, and each vertex of B in T is incident with m edges of color 1 and $s - m$ edges of color 2. All edges of G within S and between S and W have color 1 (otherwise there would be a $(2, 1)$-max vertex not in T) and all edges within T and between T and W have color 2.

We note that the edges of color 1 in B satisfy the conditions of Lemma 1.7.3, so B has a $2s$-cycle of edges of color 1 unless $s = t = 2m$ and the edges of color 1 (and those of color 2) form two disjoint copies of $K_{m,m}$.

Again, let $v \in S$ and $u \in T$ be such that $c(uv) = 2$, and let $H \in C_q^*(n)$ (or $H \in R_q(n)$), $q \geq 1$, be such that uv is the only edge of color 2. If uw is also an edge of H then $w \in S$, because otherwise $c(uw) = 2$. Hence if z is a vertex of G not in H then $z \in T$, because otherwise we can replace u with z in H to get $H'' \in C_q^*(n)$ (or $H'' \in R_q(n)$) with no edge of color 2. That means that if Q is the set of vertices of G not in H, then $Q \subseteq T$. Since uv is the only edge in H with color 2, each vertex in $T \setminus Q$ is adjacent in H to two vertices in S, so there are $2(t - q)$ edges in H between S and T, where $q = |Q| \geq t - s$.

Let M be the subgraph of H remaining when the $2(t - q)$ edges in H between S and T have been removed (along with any remaining isolated vertices). If $q = t - s$ then, since every edge in H incident to a vertex in T goes to S, either H is a $2s$-cycle and $W = \emptyset$ (if $H \in C_q^*(n)$) or the union of the components of H which have a vertex in T is a 2-regular graph spanning S and $s = t - q$ vertices in T. In either case, since $s < t$, we can replace the components of H which intersect T with the $2s$-cycle of edges of color 1 promised by Theorem 1.7.1, to get an $H'' \in C_q^*(n)$ (or $H'' \in R_q(n)$) with no edge of color 2. Hence $q > t - s$.

Each component of M is a path with at least one edge, both endpoints in S with interior points in S or W. If a component has $j > 2$ vertices in S, we split it into $j - 1$ paths which each have their endpoints in S with all interior points in W. If a vertex of S is an interior point in a component then it is an endpoint of two of these paths. The number of such paths is $\frac{2(s-(t-q))}{2} = s - (t - q) > 0$.

We denote these paths by $P_1, P_2, \ldots, P_r$ where $r = s-(t-q)$. For each i in $[r]$ where P_i has more than 2 vertices, we remove the edges containing the two endpoints (which are both in S), leaving a path W_i whose vertices are all in W (the union of the vertices in all the W_i's is equal to W).

We will now show that there cannot be a $2s$-cycle of edges of color 1 in B. Suppose J is such a $2s$-cycle. Let $R = \{x_1, x_2, \ldots, x_r\}$ be the set of any r vertices in $T \cap V(J)$ and let K be the subgraph of J obtained by removing the r vertices in R. For each $i \in [r]$ let y_{ia} and y_{ib} be the vertices adjacent to x_i in J. Both are in S and possibly $y_{ib} = y_{ja}$ if $i \neq j$. Now, for each $i \in [r]$, attach W_i to y_{ia} and y_{ib} (R_i can be oriented either way). More precisely, if W_i is the path $w_{i1}w_{i2}\ldots w_{id}$ in W, we attach it to K by adding the edges $y_{ia}w_{i1}$ and $y_{ib}w_{id}$, while if W_i is empty (meaning the i^{th} component of M has only two vertices, so none in W) we add the edge $y_{ia}y_{ib}$. The resulting graph H'' has no edge of color 2, since we constructed it using only edges from J and edges from H within $S \cup W$. Since $V(H'') = V(G) \setminus R$, H'' has $n - q$ vertices. Clearly H'' is 2-regular and, if H is a cycle, so is H'' (if H is not a cycle, H'' will still be a cycle if H does not have any components completely contained in W). Thus $H'' \in R_q(n)$ ($H'' \in C_q^*(n)$) and has no edge of color 2, a contradiction. Hence there is no $2s$-cycle of edges of color 1 in B.

By Lemma 1.7.3 it follows that $s = t = 2m$ with the edges of color 1 forming two vertex-disjoint copies of $K_{m,m}$. (If these two disjoint copies have vertex sets $S_1 \cup T_1$ and $S_2 \cup T_2$, where $S_1 \cup S_2 = S$ and $T_1 \cup T_2 = T$, then $S_1 \cup T_2$ and $S_2 \cup T_1$ are the vertex sets which induce two disjoint copies of $K_{m,m}$ with edges of color 2.) We have now verified that properties (1)–(5) hold if $q \geq 1$. We will now show we get a contradiction if $q \geq 2$.

Assume $q \geq 2$. Let T_1 and T_2 be the sets of vertices in T in the two s-cycles of edges of color 1 ($|T_1| = |T_2| = \frac{s}{2}$, $T_1 \cup T_2 = T$). Recall that $v \in S$, $u \in T$, and uv is the only edge of H of color 2. The subgraph M of H defined earlier still consists of paths which can be split into paths $P_1, P_2, \ldots, P_q$ (since $r = s - t + q = q$) with endpoints in S and interior points in W. Let J be the union of the two s-cycles of edges of color 1. Choose the subset Q of size q so that it has at least one vertex in each of T_1 and T_2, say $Q = \{x_1, x_2, \dots, x_q\}$ where $x_1 \in T_1$ and $x_q \in T_2$. Again, let K be the subgraph obtained from J by removing the vertices in Q. Then, as before, the paths $W_1, W_2, \dots, W_q$ (perhaps some of them empty) can be stitched into K. We attach W_i to y_{ia} and y_{ib} if $i \in [2, q-1]$ (just adding the edge $y_{ia}y_{ib}$ if W_i is empty). We attach W_1 to y_{1a} and y_{qb} and W_q to y_{1b} and y_{qa}, creating an $(n-q)$-cycle if no component of H is contained in W, and a 2-regular graph spanning $n-q$ vertices if H has a component contained in W. There is no edge of color 2 in this graph contradicting the assumption that if $q \geq 2$ and $\mathcal{H} \in \{R_q(n), C_q^*(n)\}$ then the maximum monochromatic degree in all optimal $\mathcal{H}$-polychromatic colorings is less than $n-1$.

Claim 7. If $\mathcal{H} \in \{C_1^*, R_1\}$ and there exists an $\mathcal{H}$-polychromatic coloring satisfying (1)–(5) in Claim 6 with $m > 1$, then there exists one with $m = 1$, i.e. one that is Z-quasi-ordered with $|Z| = 4$.

Proof. Assume there is an R_1-polychromatic coloring (C_1^*-polychromatic coloring) c with $q = 1$ satisfying (1) – (5) of Claim 6 where $s = t > 2$. Let v and x be vertices in S and u and y be vertices in T such that $c(vu) = c(xy) = 2$ and $c(xu) = c(vy) = 1$. Let c' be the coloring obtained from c by recoloring the following edges (perhaps they are recolored the same color they had under c):

$$\begin{aligned}
c'(vp) &= 1 \quad \text{for all} \quad p \in T \setminus \{u, y\} \\
c'(xp) &= 1 \quad \text{for all} \quad p \in T \setminus \{u, y\} \\
c'(zu) &= 2 \quad \text{for all} \quad z \in S \setminus \{v, x\} \\
c'(zy) &= 2 \quad \text{for all} \quad z \in S \setminus \{v, x\} \\
c'(zp) &= 3 \quad \text{for all} \quad p \in T \setminus \{u, y\} \text{ and } z \in S \setminus \{v, x\}
\end{aligned}$$

Since all but one edge incident to v and x have color 1 under c', certainly every $(n-1)$-cycle contains an edge of color 1. Similarly for u and y and edges of color 2. Every edge which was recolored had color 1 or 2 under c, so c' must be a polychromatic coloring with the same number of colors. It has the desired form with $|S| = |T| = 2$, so, in fact, is Z-quasi-ordered with $Z = \{v, x, u, y\}$. ■

We remark that a coloring c satisfying properties (1)–(5) of Claim 6 with $s = t > 2$ is actually not R_1-polychromatic. To see this, let $S_1 \cup T_1$ and $S_2 \cup T_2$ be the vertex sets of the two copies of $K_{m,m}$ of edges of color 1 ($S_1 \cup S_2 = S$, $T_1 \cup T_2 = T$) where $v \in S_1, u \in T_2$ and uv is the only edge of color 2 in $H \in R_1$. The subgraph M of H in the proof of Claim 6 has only one component (since $s-(t-q) = 1$), a path $dw_1w_2 \dots w_e z$ where $d \in S_1$, $z \in T_1$, and $\{w_1, w_2, \dots, w_e\} \subseteq W$. To construct a 2-regular subgraph with no edges of color 2 spanning $n-1$ vertices, remove a vertex x in T_2 from one of the two s-cycles of edges of color 1. If y_a and y_b are the two vertices in S_2 adjacent to x in the s-cycle, attach the path $w_1w_2\dots w_e$ to y_a and y_b to get a 2-regular subgraph with no edge of color 2 spanning $n-1$ vertices. However, this construction cannot be done when $m=1$, so in this case you do get an R_1-polychromatic coloring.

Lemma 1.8.3. *Let $\mathcal{H} \in \{R_q(n), C_q^*(n)\}$.*

(a) Suppose for some $X \neq \emptyset$ there exists an optimal X-ordered $\mathcal{H}$-polychromatic coloring of K_n. Then there is one which is ordered.

(b) Suppose there exists an optimal Z-quasi-ordered $\mathcal{H}$-polychromatic coloring of K_n. Then there is one which is quasi-ordered

Proof. Among all such $\mathcal{H}$-polychromatic colorings we assume φ is one such that

(a) if φ is X-ordered then X has maximum possible size

(b) if φ is Z-quasi-ordered then the restriction of φ to $V(K_n)\setminus Z$ is T-ordered for the largest possible subset T of $V(K_n)\setminus Z$. In this case, we let $X = Z\cup T$ so φ is nearly X-ordered (one or two edges could be recolored to make it X-ordered).

For both (a) and (b) we assume that φ is such that its restriction to $G_m = K_n[Y]$ has a vertex v of maximum possible monochromatic degree in G_m, where $Y = V(K_n)\setminus X$, $|Y| = m$, and the degree of v in G_m is $d < m-1$ (if $d = m-1$ then $|X|$ is not maximal).

Since v has maximum monochromatic degree d in G_m, by Lemma 1.8.1 it is a $(1,2)$-max vertex in G_m, for some colors 1 and 2, and if $u \in Y$ is such that $\varphi(uv) = 2$, then u is a $(2,t)$-max vertex for some color t (perhaps $t = 1$).

As before, let $y_1, y_2, \dots, y_d$ be vertices in Y such that $c(vy_i) = 1$ for $i = 1, 2, \dots, d$. As before, let $H \in \mathcal{H}$ be such that uv is its only edge with color 2. Let H' be a cyclic orientation of the edges of H such that $\overrightarrow{uv}$ is an arc, and let w_i be the predecessor of y_i in H' for $i = 1, 2, \dots, d$. As shown before, $c(w_iv) = 2$ for $i = 1, 2, \dots, d$.

Suppose there is an edge of H which has one vertex in X and one in Y. Then there exist $w \in Y$ and $x \in X$ such that $\overrightarrow{wx} \in H'$. Certainly w is not the predecessor in H' of any y_i in Y. Since φ is X-constant and uv is the only edge of color 2 in H, $\varphi(xv) = \varphi(xw) \neq 2$. Now twist xw, uv in H. Since $\varphi(xv) \neq 2$, we must have $\varphi(wu) = 2$, so u is incident in G_m to at least $d+1$ vertices of color 2, a contradiction. Hence H cannot have an edge with one vertex in X and one in Y.

Now suppose $x \in X$ and $x \notin H$. If $\varphi(xv) = \varphi(xu) \neq 2$ then $H \setminus \{uv\} \cup \{ux, xv\}$ is an $(n-q+1)$-cycle with no edge of color 2, which is clearly impossible if $\mathcal{H} = R_q(n)$, and is impossible if $\mathcal{H} = C_q^*(n)$ by Theorem 1.2.6. Hence $\varphi(xv) = \varphi(xu) = 2$ for each $x \in X$.

Since u is a $(2,t)$-max vertex for some color $t \neq 2$, we can repeat the above argument with u in place of v. That shows that $\varphi(xv) = \varphi(xu) = t$ for each $x \in X$, which is clearly impossible.

It remains to consider the possibility that $\mathcal{H} = R_q(n)$ and X is spanned by a union of cycles in H. Suppose xz is an edge of H contained in X. Then we can twist xz and uv to get another subgraph in R_q and, unless either x or z has main color 2, this subgraph has no edge of color 2. Hence at least half the vertices in X have main color 2 (and more than half would if H had an odd component in X).

The above argument can be repeated with u in place of v. If u is a $(2,t)$-max vertex then that would show that at least half the vertices in X have main color $t \neq 2$. So each vertex in X has main color 2 or t. Since φ is X-ordered or nearly X-ordered, some vertex $x \in X$ has monochromatic degree $n-2$ or $n-1$ and the main color of x must be 2 or t. Assume it is 2. Then every cycle containing x has an edge with color 2, contradicting the assumption that H has only one edge with color 2. Similarly, we get a contradiction if the main color of x is t. We have shown there is no vertex v with monochromatic degree $d < m-1$, so φ is ordered or quasi-ordered. ■

Now there is not much left to do to prove Theorems 1.2.2, 1.2.3, and 1.2.4.

1.8.1 Proof of Theorem 1.2.4

Theorem 1.2.5 takes care of the case of C_q-polychromatic colorings when $q \geq 2$ and $n \in [2q+2, 3q+2]$. The smallest value of n for which there is a simply-ordered C_q-polychromatic 2-coloring is $n = 3q+3$ (the coloring φ_{C_q} in Section 1.9.3). Hence if $q \geq 2$ and $\varphi_{\mathrm{C_q}} \leq 2$ then there exists an optimal simply-ordered C_q-polychromatic coloring except if $n-q$ is odd and $n \in [2q+2, 3q+2]$, or if $q = 2$ and $n = 5$ (the coloring of K_5 with two monochromatic 5-cycles has no monochromatic 3-cycle). So we need only consider $\mathcal{H} \in \{R_q(n), C_q^*(n)\}$ (when $q \geq 2$). Since none of (a), (b), or (c) of Lemma 1.8.2 are satisfied, there exists an optimal $\mathcal{H}$-polychromatic coloring with maximum monochromatic degree $n-1$. That means it is X-ordered, for some nonempty set X, so by Lemma 1.8.3 there exists an optimal $\mathcal{H}$-polychromatic coloring which is ordered, and then, by Lemma 1.4.2, one which is simply-ordered. ■

1.8.2 Proof of Theorem 1.2.2

If $\mathcal{H} \in \{R_0(n), C_0(n)\}$ then, by Lemma 1.8.2, if there does not exist an optimal $\mathcal{H}$-polychromatic coloring with maximum monochromatic degree $n-1$, then there exists one which is Z-quasi-ordered with $|Z| = 3$. Hence by Lemma 1.8.3(b), there exists one which is quasi-ordered, and then, by Lemma 1.4.2, one which is quasi-simply-ordered with $|Z| = 3$. Such a coloring is one candidate to be an optimal $\mathcal{H}$-polychromatic coloring.

If there does exist an optimal $\mathcal{H}$-polychromatic coloring with maximum monochromatic degree $n-1$, then, since this coloring is X-ordered with $|X| = 1$, by Lemma 1.8.3(a), there is an optimal ordered $\mathcal{H}$-polychromatic coloring, and by Lemma 1.4.2, one which is simply-ordered. This is the other candidate to be an optimal $\mathcal{H}$-polychromatic coloring.

For each of these candidates, the conditions in Lemma 1.4.1 ((II)(a) for C_0-polychromatic and (III)(a) for R_0-polychromatic) provide lower bounds for the sizes of the successive coloring classes. For fixed n we clearly will get the maximum number of colors if we make the successive classes as small as possible, while satisfying the required inequalities, so it is a simple matter to determine which candidate is better.

If $\mathcal{H} = R_0(n)$, the sizes of the successive color classes for simply-ordered are $1, 1, 3, 6, 12, 24, \ldots$ ($|M_t| > \sum_{i=1}^{t-1} |M_i|$ if $t \geq 3$), while for quasi-simply-ordered the sizes are $1, 1, 1, 4, 8, 16, \ldots$ (with $|Z| = 3$ the inequality is required only for $t \geq 4$, since the main colors in Z will automatically appear in every 2-factor). Hence the quasi-simply-ordered coloring is always at least as good. For example, if $n - 29$, then both will use 5 colors (color class sizes $1, 1, 3, 6, 18$ for simply-ordered and $1, 1, 1, 4, 22$ for quasi-simply-ordered), while if $n = 35$ the simply-ordered coloring will still use 5 colors (color class sizes $1, 1, 3, 6, 24$), while the quasi-simply-ordered coloring will use 6 colors (color class sizes $1, 1, 1, 4, 8, 20$). A formula for $\text{poly}_{R_0}(n)$ appears in Section 1.9.

The situtation is similar if $\mathcal{H} = C_0$: The color class sizes for the simply-ordered candidate are $1, 1, 2, 4, 8, 16, \ldots$ and for the quasi-simply-ordered candidate are $1, 1, 1, 3, 6, 12, \ldots$. Again, the quasi-simply-ordered candidate is at least as good for any value of n. A formula for $\text{poly}_{C_0}(n)$ appears in Section 1.9.

We have already remarked that these optimal quasi-simply-ordered $\mathcal{H}$-polychromatic colorings can be obtained by recoloring one edge of a simply-ordered coloring (which is not $\mathcal{H}$-polychromatic).

■

1.8.3 Proof of Theorem 1.2.3

As in the proof of Theorem 1.2.2, there are two candidates to be an optimal $\mathcal{H}$-polychromatic coloring wiht $\mathcal{H} \in \{R_1(n), C_1(n)\}$, one of them simply-ordered and the other quasi-simply-ordered with $|Z| = 4$. If $\mathcal{H} = R_1(n)$, the successive color class sizes are $2, 4, 8, 16, 32, \ldots$ for simply-ordered and $2, 2, 6, 12, 24, \ldots$ for quasi-simply-ordered, so the quasi-simply-ordered coloring is at least as good. If $\mathcal{H} = C_1(n)$, the color class sizes are $2, 3, 6, 12, 24, \ldots$ for simply-ordered and $2, 2, 5, 10, 20, \ldots$ for quasi-simply-ordered, so again the quasi-simply-ordered coloring is at least as good. We have already remarked that these optimal quasi-simply-ordered colorings with $|Z| = 4$ can be obtained from a (non-$\mathcal{H}$-polychromatic) simply-ordered coloring by recoloring two edges. Formulae for $\text{poly}_{R_1}(n)$ and $\text{poly}_{C_1}(n)$ appear in Section 1.9. ■

1.9 Optimal Polychromatic Colorings

The seven following colorings are all optimal F_q, R_q, or C_q polychromatic colorings for various values of q and n. Each of them is simply-ordered or quasi-simply-ordered. We describe the color classes for each, and give a formula for the polychromatic number k in terms of q and n.

1.9.1 F_q-polychromatic coloring $\varphi_{\mathrm{F_q}}$ of $E(K_n)$ (even $n - q \geq 2$).

Let q be nonnegative and $n - q$ positive and even with k a positive integer such that

$$(q+1)(2^k - 1) \leq n < (q+1)(2^{k+1} - 1). \tag{1.9.1}$$

Let φ_{F_q} be the simply-ordered edge k-coloring with colors $1, 2, \ldots, k$ with the inherited vertex k-coloring φ'_{F_q} having successive color classes $M_1, M_2, \ldots, M_k$, moving left to right such that $|M_i| = 2^{i-1}(q+1)$ if $i < k$ and $|M_k| = n - \sum_{i=1}^{k-1} |M_i| = n - (2^{k-1} - 1)(q+1)$. We have $k \leq \log_2 \frac{n+q+1}{q+1} < k + 1$ so $\text{poly}_{F_q}(n) = k = \lfloor \log_2 \frac{n+q+1}{q+1} \rfloor$.

1.9.2 R_q-polychromatic coloring φ_{R_q} $(q \geq 2)$

If $q \geq 2$, $n \geq q + 3$ and n and k are such that (1.9.1) is satisfied, we let $\varphi_{R_q} = \varphi_{F_q}$ (same color classes), giving us the same formula for k in terms of n.

1.9.3 C_q-polychromatic coloring φ_{C_q} $(q \geq 2)$.

If $q \geq 2$, $n \geq q + 3$ and

$$(2^k - 1)q + 2^{k-1} < n \leq (2^{k+1} - 1)q + 2^k \tag{1.9.2}$$

let φ_{C_q} be the simply-ordered edge k-coloring with colors $1, 2, \ldots, k$ and the inherited vertex k coloring φ'_{C_q} with successive color classes $M_1, M_2, \ldots, M_k$ of sizes given by:

$$
\begin{aligned}
|M_1| &= q+1, \\
|M_i| &= 2^{i-1}q + 2^{i-2} \text{ if i} \in [2, \text{k}-1], \\
|M_k| &= n - \sum_{i=1}^{k-1} |M_i| = n - 2^{k-1}q - 2^{k-2}.
\end{aligned}
$$

From equation (1.9.2) we get $\text{poly}_{C_q}(n) = k = \lfloor \log_2 \frac{2(n+q-1)}{2q+1} \rfloor$.

1.9.4 R_0-polychromatic coloring φ_{R_0} $(q = 0)$.

If $n \geq 3$ and $2^{k-1} - 1 \leq n < 2^{k-1}$ let φ_{R_0} be the quasi-simply-ordered coloring with $|X| = 3$ and color class sizes $|M_1| = |M_2| = 1$ and $|M_3| = n - 2$ if $3 \leq n \leq 6$, and if $n \geq 7$:

$$
\begin{aligned}
|M_1| &= |M_2| = |M_3| = 1, \\
|M_i| &= 2^{i-2} \text{ if i} \in [4, \text{k}-1], \\
|M_k| &= n - \sum_{i=1}^{k} -1\, |M_i| = n - 2^{k-2} + 1.
\end{aligned}
$$

From this, we get $\text{poly}_{R_0}(n) = k = 1 + \lfloor \log_2(n+1) \rfloor$ where $n \geq 3$.

1.9.5 C_0-polychromatic coloring φ_{C_0} $(q = 0)$

If $n \geq 3$ and $3 \cdot 2^{k-3} < n \leq 3 \cdot 2^{k-2}$ let φ_{C_0} be the quasi-simply-ordered coloring with $|X| = 3$ and color class sizes $|M_1| = |M_2| = 1$ and $|M_3| = n - 2$ if $3 \leq n \leq 6$, and if $n \geq 7$:

$$
\begin{aligned}
|M_1| &= |M_2| = |M_3| = 1, \\
|M_i| &= 3 \cdot 2^{i-4} \text{ if i} \in [4, \text{k}-1], \\
|M_k| &= n - \sum_{i=1}^{k-1} |M_i| = n - 3 \cdot 2^{k-4}.
\end{aligned}
$$

From this, we get $\text{poly}_{C_0}(n) = k = \lfloor \log_2 \frac{8(n-1)}{3} \rfloor$ where $n \geq 4$.

1.9.6 R_1-polychromatic coloring φ_{R_1} $(q = 1)$

If $n \geq 4$ and $3 \cdot 2^{k-1} - 2 \leq n < 3 \cdot 2^k - 2$ let φ_{R_1} be the quasi-simply-ordered coloring with $|X| = 4$ and color class sizes $|M_1| = 2$ and $|M_2| = n-2$ if $4 \leq n \leq 9$, and if $n \geq 10$:

$$|M_1| = |M - 2| = 2,$$
$$|M_i| = 3 \cdot 2^{i-2} \text{ if i} \in [3, \text{k} - 1],$$
$$|M_k| = n - \sum_{i=1}^{k-1} |M_i| = n - 3 \cdot 2^{k-2} + 2.$$

From this, we get $\text{poly}_{R_1}(n) = k = \lfloor \log_2 \frac{2(n+2)}{3} \rfloor$ where $n \geq 4$.

1.9.7 C_1-polychromatic coloring φ_{C_1} $(q = 1)$

If $n \geq 4$ and $5 \cdot 2^{k-2} \leq n < 5 \cdot 2^{k-1}$ let φ_{C_1} be the quasi-simply-ordered coloring with $|X| = 4$ and color class sizes $|M_1| = |M_2| = 2$ and $|M_3| = n - 4$ if $4 \leq n \leq 9$ and change every edge of color 3 to color 2, and if $n \geq 10$:

$$|M_1| = |M_2| = 2,$$
$$|M_i| = 5 \cdot 2^{i-3} \text{ if i} \in [3, \text{k} - 1],$$
$$|M_k| = n - \sum_{i=1}^{k-1} |M_i| = n - 5 \cdot 2^{k-3} + 1.$$

From this, we get $\text{poly}_{C_1}(n) = k = \lfloor \log_2 \frac{4n}{5} \rfloor$ where $n \geq 4$.

1.10 Polychromatic cyclic Ramsey numbers

Let s, t, and j be integers with $t \geq 2, s \geq 3, s \geq t$, and $1 \leq j \leq t-1$. We define $\text{CR}(s, t, j)$ to be the smallest integer n such that in any t-coloring of the edges of K_n there exists an s-cycle that uses at most j colors. Erdős and Gyárfás [19] defined a related function for cliques instead of cycles. So $\text{CR}(s, t, 1)$ is the classical t-color Ramsey number for s-cycles and $\text{CR}(s, 2, 1) = c(s)$, the function in Theorem 1.6.1. While it may be difficult to say much about the function $\text{CR}(s, t, j)$ in general, if $j = t-1$ we get $\text{CR}(s, t, t-1) = \text{PR}_t(s)$ the smallest integer $n \geq s$ such that in any t-coloring of K_n there exists an s-cycle that does not contain all t colors. This is the function of Theorem 1.2.7 if $t \geq 3$, while

$\mathrm{PR}_2(s) = c(s)$.

1.10.1 Proof of Theorem 1.2.7

Let $q \geq 0, s \geq 3$, and n be integers with $n = q+s$. Assume $q \geq 2$. By Theorem 1.2.4 and the properties of the coloring φ_{C_q} (see Section 1.9.3), there exists a C_q-polychromatic t-coloring of K_n if and only if

$$\begin{aligned} q + s = n &\geq (2^t - 1)q + 2^{t-1} + 1, \\ s &\geq (2^t - 2)q + 2^{t-1} + 1, \\ q &\leq \frac{s - 2^{t-1} - 1}{2^t - 2} = \frac{s-2}{2^t-2} - \frac{1}{2} \end{aligned}$$

Since $q \geq 2$, we want to choose s so that the right-hand side of the last inequality is at least 2, so

$$\begin{aligned} s - 2 &\geq \frac{5}{2}(2^t - 2) = 5 \cdot 2^{t-1} - 5 \\ s &\geq 5 \cdot 2^{t-1} - 3 \end{aligned}$$

So if $s \geq 5 \cdot 2^{t-1} - 3$, then the smallest n for which there does not exist a C_q-polychromatic k-coloring is $n = q + s$ where $q > \frac{s-2}{2^t-2} - \frac{1}{2}$, so $n = s + \left\lfloor \frac{s-2}{2^t-2} + \frac{1}{2} \right\rfloor = s + \mathrm{Round}\left(\frac{s-2}{2^t-2}\right)$.

We note that if $s \geq 5 \cdot 2^{t-1} - 3$ then $\mathrm{Round}\left(\frac{s-2}{2^t-2}\right) \geq \mathrm{Round}\left(\frac{5}{2}\right) = 3$, so $\mathrm{PR}_t(s) \geq s+3$ if $s \geq 5 \cdot 2^{t-1} - 3$.

Now we assume that $\mathrm{PR}_t(s) = s + 2$. So $s + 2$ is the smallest value of n for which in any t-coloring of the edges of K_n there is an s-cycle which does not have all colors, which means there is a polychromatic t-coloring when $n = s + 1$. Since $q = 1$ in such a coloring, by Theorem 1.2.3 and the properties of the coloring φ_{C_1}, $n \geq 5 \cdot 2^{t-2}$. Hence if $s \in [5 \cdot 2^{t-2} - 1, 5 \cdot 2^{t-1} - 4]$, then $\mathrm{PR}_t(s) = s + 2$.

Now we assume that $\mathrm{PR}_t(s) = s + 1$. So $n - s$ is the largest value of n such that in any t-coloring of K_n, every s-cycle gets all colors. So $q = n - s = 0$ and, by Theorem 1.2.2 and properties of the coloring φ_{C_0}, $n \geq 3 \cdot 2^{t-3} + 1$.

Finally, since the t-coloring φ_{C_0} requires $n \geq 3 \cdot 2^{t-3} + 1$ where $t \geq 4$ if $n \leq 3 \cdot 2^{t-3}$ and $t \geq 4$, then in any t-coloring of K_n, some Hamiltonian cycle will not get all colors, so $\mathrm{PR}_t(s) = s$ if $3 < s \leq 3 \cdot 2^{t-3}$. ■

1.11 Conjectures

We mentioned that we have been unable to prove a result for 2-regular graphs analogous to Theorem 1.2.6 for cycles. In fact, we think it holds even for two colors, except for a few cases with j and n small.

Conjecture 1.11.1. Let $n \geq 6$ and j be integers such that $3 \leq j < n$, and if $j = 5$ then $n \geq 9$, and let φ be an edge-coloring of K_n so that every 2-regular subgraph spanning j vertices gets all colors. Then every 2-regular subgraph spanning at least j vertices gets all colors under φ.

This does not hold for $j = 3, n = 4$, and 3 colors (proper edge 3-coloring) or for $n = 5, j = 3$, and 2 colors (two monochromatic K_5s).

We can extend the notions of Z-quasi-ordered, quasi-ordered, and quasi-simply-ordered to sets Z of larger size, allowing a main color to have degree less than $n - 2$. Let $q \geq 0$ and $r \geq 1$ be integers such that $q \leq 2r - 3$. Hence $\frac{2r-2}{q+1} \geq 1$, and we let $k = \left\lfloor \frac{2r-2}{q+1} \right\rfloor + 1 \geq 2$ and $z = k(q + 1)$. Let Z be a set of z vertices. We define a *seed-coloring* φ with k colors on the edges of the complete graph K_z with vertex set Z as follows. Partition the z vertices into k sets $S_1, S_2, \ldots, S_k$ of size $q+1$. For $j = 1, 2, \ldots, k$, all edges within S_j have color j, all edges between S_i and S_j $(i \neq j)$ have color i or j, and for each j and each vertex v in S_j, v is incident to $\left\lceil \frac{(q+1)(k-1)}{2} \right\rceil$ or $\left\lfloor \frac{(q+1)(k-1)}{2} \right\rfloor$ edges with colors other than j (so, within round off, half of the edges from each vertex in S_j to vertices in other parts have color j). We say each vertex in S_j has main color j.

If $n \geq z$, we get a Z-quasi-ordered coloring c of K_n which is an extension of the coloring φ on Z if for each j and each $v \in S_j$, $c(vy) = j$ for each $y \in V(K_n) \setminus Z$. If c is Z-quasi-ordered then it is quasi-ordered if c restricted to $V(K_n) \setminus Z$ is ordered, and quasi-simply-ordered if c restricted to $V(K_n) \setminus Z$ is simply-ordered.

If $r > 0$ and $q \geq 0$ are integers, we let $\mathscr{R}(n, r, q)$ be the set of all r-regular subgraphs of K_n spanning precisely $n - q$ vertices (assume $n - q$ is even if r is odd, so the set is nonempty), and if $r \geq 2$ let $\mathscr{C}(n, r, q)$ be the set of all such subgraphs which are connected.

Since $k - 1 = \left\lfloor \frac{2r-2}{q+1} \right\rfloor \leq \frac{2r-2}{q+1}$, we have $r \geq \frac{(q+1)(k-1)}{2} + 1 > \left\lceil \frac{(q+1)(k-1)}{2} \right\rceil$. So if H is in $\mathscr{R}(n, r, q)$ or $\mathscr{C}(n, r, q)$, then H contains an edge with each of the k colors on edges within Z, because it contains at least one vertex in S_j for each j, and fewer than r of the edges incident to this vertex have colors other than j. We can get an $\mathscr{R}(n, r, q)$-polychromatic or $\mathscr{C}(n, r, q)$-polychromatic quasi-simply-ordered coloring of K_n with $m > k$ colors by making the color classes M_t on the vertices in $V(K_n) \setminus Z$ for $t = k + 1, k + 2, \ldots, m$

sufficiently large. If $H \in \mathscr{R}(n,r,q)$, for each $t \in [k+1,m]$ we will need the size of M_t to be at least $q+1$ more than the sum of the sizes of all previous color classes, while if $H \in \mathscr{C}(n,r,q)$ we will need the size of M_t to be at least q more than the sum of the sizes of all previous classes, with an extra vertex in M_m. To try to get optimal polychromatic colorings we make the sizes of these color classes as small as possible, yet satisfying these conditions.

For example, if $r=2$ and $q=0$ then $k = \left\lfloor \frac{2r-2}{q+1} \right\rfloor + 1 = 3$ and $z = k(q+1) = 3$, and we get the quasi-simply-ordered colorings φ_{R_0} and φ_{C_0} with $|Z| = 3$ of Theorem 1.2.2. If $r=2$ and $q=1$ then $k=2$ and $z=4$, and we get the colorings φ_{R_1} and φ_{C_1} with $|Z|=4$ of Theorem 1.2.3.

Example 1 ($r=3, q=0$, so $k=5, z=5$)**.** Let φ be the edge coloring obtained where $\{v_1,v_2,v_3,v_4,v_5\} = Z$ such that v_iv_{i+1} and v_iv_{i+2} (mod 5) have color i. The edges connecting v_i to the remaining vertices in $V(K_n) \setminus Z$ are color i. See Figure 1.4.

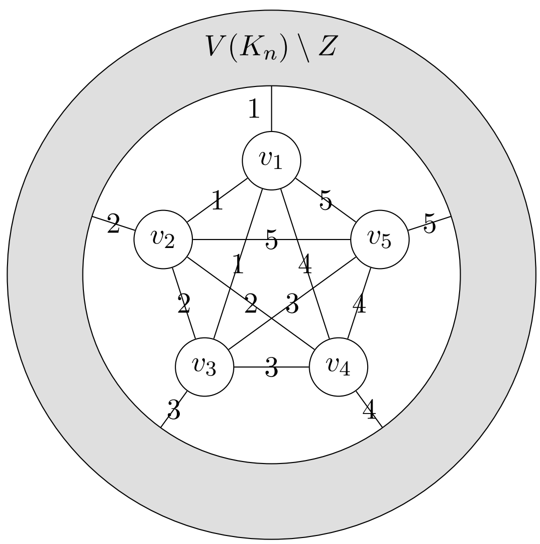


Figure 1.4: The coloring for Example 1.

Example 2 ($r=3, q=3, k=2, z=8$)**.** Z has two color classes, 4 vertices in each. The complete bipartite graph between these two sets of vertices could have two vertex disjoint copies of $K_{2,2}$ of one color and also of the other color, or could have an 8-cycle of each color.

Example 3 ($r=4, q=2, k=3, z=9$)**.** So S_1, S_2, S_3 each have size $q+1=3$. One way to color the edges between parts is for $j=1,2,3$, each vertex in S_j is incident with 2 edges of color j to vertices in S_{j+1} and 1 edge of color j to a vertex in S_{j-1} (so is incident with one edge of color $j+1$ and two edges of color $j-1$, cyclically). The smallest value of n for which this seed can generate a quasi-simply-ordered $\mathscr{R}(n,4,2)$-polychromatic coloring with 5-colors is $n = 45$ (the 4^{th} and 5^{th} color classes would have sizes $9+2+1 = 12$ and $21+2+1 = 24$ respectively), while to get a simply-ordered $\mathscr{R}(n,4,2)$-polychromatic coloring with 5 colors you would need $n \geq 69$ (color class sizes $3, 3, 9, 18, 36$ works).

Conjecture 1.11.2. Let $r \geq 1$ and $q \geq 0$ be integers such that $q \leq 2r-3$. Let $k = \left\lfloor \frac{2r-2}{q+1} \right\rfloor + 1 \geq 2$ and $z = k(q+1)$. If $n \geq z$ and $n-q$ is even if r is odd, then there exist optimal quasi-simply-ordered $\mathscr{R}(n,r,q)$ and $\mathscr{C}(n,r,q)$-polychromatic colorings with seed Z with parameters r, q, k, z.

It is not hard to check that each of these quasi-simply-ordered colorings does at least as well as a simply-ordered coloring for those values of r and q. The only quesiton is whether some other coloring does better and the conjecture says no.

What if $\frac{2r-2}{q+1} < 1$? Then $k = \left\lfloor \frac{2r-2}{q+1} \right\rfloor + 1 = 1$, which seems to be saying no seed Z exists with at least 2 colors.

Conjecture 1.11.3. Let $r \geq 1$ and $q \geq 0$ be integers with $q \geq 2r-2$, $n \geq q+r+1$, and not both r and $n-q$ are odd. Then there exists an optimal simply-ordered $\mathscr{R}(n,r,q)$-polychromatic coloring of K_n. If $r \geq 2$ then there exists a $\mathscr{C}(n,r,q)$-polychromatic coloring of K_n (unless $r = 2$, $q \geq 2$, $n-q$ is odd, and $n \in [2q+2, 3q+1]$).

Theorem 1.2.1 says this conjecture is true for $r = 1$. Theorem 1.2.4 says it is true for $\mathscr{C}(n,r,q)$ for $r = 2$ and that it would be true for $\mathscr{R}(n,r,q)$ for $r = 2$ if Theorem 1.2.6 held for 2-regular graphs.

Chapter 2

Maximum density of vertex-induced perfect cycles and paths in the hypercube

The material for this chapter currently appears in publication in Discrete Mathematics, Volume 344, Issue 11, November 2021 [26].

2.1 Background

The n-cube, which we denote by Q_n, is the graph whose vertex set $V_n = V(Q_n)$ is the set of all binary n-tuples, with two vertices adjacent if and only if they differ in precisely one coordinate (so Hamming distance 1). Let $[n] = \{1, 2, \ldots, n\}$. We sometimes denote a vertex $(x_1, x_2, \ldots, x_n)$ of Q_n by the subset S of $[n]$ such that $i \in S$ if and only if $x_i = 1$. So if $n = 4$, then $\emptyset$ denotes (0000), and $\{1, 3\}$ or 13, denotes (1010) and $\{\{1\}, \{1, 3\}\}$ (or $\{1, 13\}$) denotes $\{(1000), (1010)\}$. The weight of a vertex is the number of 1s. For each positive integer d less than or equal to n, Q_n has $\binom{n}{d}2^{n-d}$ subgraphs which are isomorphic to Q_d (d coordinates can vary, while $n-d$ coordinates are fixed).

Let H and K be subsets of $V(Q_d)$ (we call H and K configurations in Q_d). We say K is an *exact copy* of H if there is an automorphism of Q_d which sends H to K. For example, $\{\emptyset, 12\}$ is an exact copy of $\{2, 123\}$ in Q_3, but $\{2, 13\}$ is not (the vertices are distance 3 apart). So if K is an exact copy of H then they induce isomorphic subgraphs of Q_d, but the converse may not hold.

Let d and n be positive integers with $d \le n$, let H be a configuration in Q_d and let S be a subset of V_n. We let $G(H, d, n, S)$ denote the number of sub-d-cubes R of Q_n in which $S \cap R$ is an exact copy of H, $g(H, d, n, S) = \frac{G(H,d,n,S)}{\binom{n}{d}2^{n-d}}$, $G_{\max}(H, d, n) = \max_{S \subseteq V_n} G(H, d, n, S)$ and

$$\lambda(H, d, n) = \frac{G_{\max}(H, d, n)}{\binom{n}{d}2^{n-d}} = \max_{S \subseteq V_n} g(H, d, n, S).$$

Note that $\lambda(H, d, n)$ is the average of $2n$ densities $g(H, d, n-1, S_j)$, each of them the fraction of sub-d-cubes R in a sub-$(n-1)$-cube of Q_n in which $R \cap S_j$ is an exact copy of H, where S_j is the intersection of a maximizing subset S of V_n with one of the $2n$ sub-$(n-1)$-cubes. Hence $\lambda(H, d, n)$ is the average of $2n$ densities, each of them less than or equal to $\lambda(H, d, n-1)$, which means $\lambda(H, d, n)$ is a nonincreasing function of n, so we can define the d-cube density $\lambda(H, d)$ of H by

$$\lambda(H, d) = \lim_{n \to \infty} \lambda(H, d, n).$$

So $\lambda(H, d)$ is the limit as n goes to infinity of the maximum fraction, over all $S \subseteq V_n$, of "good" sub-d-cubes – those whose intersection with S is an exact copy of H.

As far as we know, our paper [26], which is the basis for this chapter's material, was the first to define the notion of d-cube density. There have been many papers on Turán and Ramsey type problems in the hypercube. There has been extensive research on the maximum fraction of edges of Q_n one can take with no cycle of various lengths [16, 18, 24, 42] and a few papers on vertex Turán problems in Q_n [35–37]. There has also been extensive work on which monochromatic cycles must appear in any edge-coloring of a large hypercube with a fixed number of colors [3, 5, 16, 17], and a few results on which vertex structures must appear [29]. Leader and Long [38] showed that if the average degree in a subgraph of Q_n is at least d, then there must be a *geodesic* of length d, and their geodesic is what we call a perfect path.

In [2, 28, 39] results were obtained on the polychromatic number of Q_d in Q_n, the maximum number of colors in an edge coloring of a large Q_n such that every sub-d-cube gets all colors.

We wanted to investigate a different extremal problem in the hypercube: the maximum density of a small structure within a subgraph of a large hypercube. Instead of using graph isomorphism to determine if two substructures are the same, it seemed to capture the essesnce of a hypecube better if the small structure was "rigid" within a sub-d-cube, and that is what motivated our definition of d-cube density. It is not quite the same thing as "isomorphism preserving Hamming distance" either. If $H = \{(0000), (1100), (1010), (0110)\}$ and $K = \{(0000), (1100), (1010), (1001)\}$ then H and K are each 4 isolated vertices, each pair of them Hamming distance 2 apart, but K is not an exact copy of H (H embeds in a 3-cube and K does not).

There are strong connections between d-cube density and *inducibility* of a graph, a notion of extensive study over the past few years. Given graphs G and H, with $|V(G)| = n$ and $|V(H)| = k$, the *density* of H in G, denoted $d_H(G)$, is defined by

$$d_H(G) = \frac{\text{\# of induced copies of } H \text{ in } G}{\binom{n}{k}}$$

Pippenger and Golumbic [40] defined the *inducibility* $I(H)$ of H by

$$I(H) = \lim_{n\to\infty} \max_{|V(G)|=n} d_H(G).$$

Within the past few years, $I(H)$ has been determined for all graphs H with 4 vertices except the path P_4 [20, 21, 33].

Given a graph H, a natural candidate for maximizing the number of induced copies of H is a balanced blow-up of H. Equipartition the n vertices into $|V(H)| = k$ classes corresponding to the vertices of H and add all possible edges between each pair of parts corresponding to an edge of H. Any k-subset which has one vertex in each part will induce a copy of H, so $I(H) \geq \frac{k!}{k^k}$ for any graph H with k vertices. Iterating blow-ups of H within each part improves the bound to $I(H) \geq \frac{k!}{k^k - k}$.

A natural generalization of $I(H)$ is to restrict the host graph G to a particular class of graphs. Let $\mathscr{G}$ be a class of graphs. The inducibility of H in $\mathscr{G}$ is defined by

$$I(H,\mathscr{G}) = \lim_{n\to\infty} \max_{|V(G)|=n, G\in\mathscr{G}} d_H(G),$$

if the limit exists (if $\mathscr{G}$ is all graphs the limit always exists). Let $\mathscr{T}$ be the family of all triangle-free graphs. Hatami et al. [32] and Grzesik [31] used flag algebras to show that $I(C_5,\mathscr{T}) = \frac{5!}{5^5} = \frac{24}{625}$, achieving the non-iterated blow-up lower bound.

In [15], Choi, Lidicky, and Pfender consider the inducibility of oriented graphs (directed graphs with no 2-cycles). For the directed path $\overrightarrow{P_k}$ they conjectured that

$$I(\overrightarrow{P_k}) = \frac{k!}{(k+1)^{k-1}-1}$$

the lower bound provided by an iterated blow-up of the directed cycle $\overrightarrow{C_{k+1}}$. To eliminate the possibility of iterated blow-ups, they considered the family $\vec{\mathcal{T}}$ of oriented graphs with no transitive tournament on three vertices (so every 3-cycle is directed). They conjectured that

$$I(\overrightarrow{P_k},\vec{\mathcal{T}}) = \frac{k!}{(k+1)^{k-1}}$$

Again, the lower bound is provided by a blow-up of $\overrightarrow{C_{k+1}}$ (no iterations). They used flag algebras to prove their conjecture for $k=4$:

$$I(\overrightarrow{P_4},\vec{\mathcal{T}}) = \frac{4!}{5^3} = \frac{24}{125}.$$

It has been shown [8,12] that if H is a complete bipartite graph then the graph that maximizes $I(H)$ can be chosen to be complete bipartite. There are also a few results on inducibility of 3-graphs [22].

A different kind of blow-up can be used to produce a lower bound for $\lambda(H,d)$ for any configuration H in Q_d (Proposition 2.3.1). As with inducibility, d-cube density is exceedingly difficult to determine for all but a few configurations H. We have some results in Chapter 3 for certain configurations H when d is equal to 2, 3, or 4, and for a couple of infinite families with d any integer greater than 2. If H is two opposite vertices in Q_2, clearly $\lambda(H,2)=1$ (let S be all vertices in Q_n of even weight). A more interesting example is when H is two adjacent vertices in Q_2. Then it is not hard to show that $\lambda(H,2)=\frac{1}{2}$. (For the lower bound, take S to be all vertices in Q_n such that the sum of coordinates 1 through $\lfloor \frac{n}{2} \rfloor$ is even. Any sub-2-cube which has one varying coordinate in and one out of $\left[1,\lfloor \frac{n}{2} \rfloor\right]$ will have an exact copy of H.) A single vertex seems to be one of the hard ones. Let W_d be the configuration in Q_d consisting of a single vertex.

We have been unable to determine $\lambda(W_d, d)$ for any $d \geq 2$. Letting S be the set of all vertices in Q_n with weight a multiple of 3 shows that $\lambda(W_2, 2) \geq \frac{2}{3}$. Using flag algebras Rahil Baber [6] has shown that $\lambda(W_2, 2) \leq .686$. We conjecture that $\lambda(W_2, 2) = \frac{2}{3}$ and that for sufficiently large d, one cannot do better than choosing vertices randomly with uniform probability $\left(\frac{1}{2}\right)^d$, which gives d-cube density $\frac{1}{e}$ in the limit as d goes to infinity. This has the same flavor as a special case of the edge-statistics conjecture of Alon et. al. [1] which says (though formulated differently) that the limit as k goes to infinity of the inducibility of a graph with k vertices and one edge is $\frac{1}{e}$.

In this chapter, we determine the d-cube density of a "perfect" path with 4 vertices in Q_3 and a "perfect" 8-cycle in Q_4.

2.2 Results

Let P_{d+1} denote the vertex set of a path in Q_d with $d+1$ vertices whose endpoints are Hamming distance d apart. We call P_{d+1} a *perfect path*. For example, $\{\emptyset, 1, 12, 123, 1234\}$ and $\{13, 3, \emptyset, 4, 24\}$ are both perfect paths in Q_4, while $\{13, 3, \emptyset, 4, 14\}$ is not, even though these 5 vertices do induce a graph-theoretic path.

Let C_{2d} denote the vertex set of a $2d$-cycle in Q_d where all d opposite pairs of vertices are distance d apart. We call C_{2d} a *perfect* $2d$*-cycle*. The only graph-theoretic induced 6 cycle in Q_3 is perfect, but while $\{\emptyset, 1, 12, 123, 1234, 234, 34, 4\}$ is a perfect 8-cycle, $\{\emptyset, 1, 12, 123, 23, 234, 34, 4\}$ and $\{\emptyset, 1, 12, 123, 1234, 134, 34, 3\}$ induce 8-cycles in Q_4 which are not perfect (and are not exact copies of each other).

The main results in this paper are the two following theorems.

Theorem 2.2.1. $\lambda(C_8, 4) = \frac{3}{32}$

Theorem 2.2.2. $\lambda(P_4, 3) = \frac{3}{8}$

These are special cases of the following conjectures.

Conjecture 2.2.3. $\lambda(C_{2d}, d) = \frac{d!}{d^d}$ for all $d \geq 4$.

Conjecture 2.2.4. $\lambda(P_{d+1}, d) = \frac{d!}{(d+1)^{d-1}}$ for all $d \geq 3$.

Note that the formulas in these two conjectures are the same as in the conjectures about the inducibility of directed cycles and paths in oriented graphs. Conjecture 2.2.3 is significant because, as we show in Proposition 2.3.1, $\lambda(H, d) \geq \frac{d!}{d^d}$ for all configurations H in Q_d for all $d \geq 1$. So Conjecture 2.2.3 says that the perfect $2d$-cycle has the minimum possible d-cube density for all $d \geq 4$, and Theorem 2.2.1 says the conjecture is correct for $d = 4$. To show $\frac{d!}{d^d}$ is also an upper bound when $d = 4$ we needed to find the inducibility of two vertex disjoint edges in the family of all bipartite graphs. To prove both Theorem 2.2.1 and Theorem 2.2.2, we show that the d-cube density we are trying to determine is equal to the fraction of d-sequences of an n-set which have certain properties and we then solve the sequence problems.

2.3 Constructions

Consider the following construction which gives a lower bound for the d-cube density of any configuration H in Q_d, for any d. Recall that $[n]$ denotes the set $\{1, 2, \ldots, n\}$. We partition $[n]$ into $A_1, A_2, \ldots, A_d$ and let B be the set of binary d-tuples representing H. For each vertex $v = (v_1, v_2, \ldots, v_n)$ in Q_n we let $v(A_i)$ equal 0 or 1 according to $v(A_i) \equiv \sum_{j \in A_i} v_j \mod 2$. We put v in S if and only if the d-tuple $(v(A_j))_{j\in[d]}$ is in B. For example, for a perfect 8-cycle in Q_4, we could have $B = \{0000, 1000, 1100, 1110, 1111, 0111, 0011, 0001\}$ and v would be in S if and only if its number of 1s in coordinates in A_1, A_2, A_3, A_4 is either even,even,even,even, or odd,even,even,even, and so on. We observe that if a sub-d-cube has one coordinate in each of $A_1, A_2, \ldots, A_d$, then it will contain an exact copy of H. By taking an equipartition of $[n]$, we find the following lower bound:

Proposition 2.3.1. $\lambda(H, d) \geq \dfrac{d!}{d^d}$ *for all configurations H in Q_d for all positive integers d.*

We call a set S constructed in this way a *blow-up of H*. This notion of blow-up is clearly related to, but not the same as, the blow-up of a graph G (for one thing a blow-up of a graph has one part for each vertex, whereas a blow up of a configuration in Q_d has d parts). In Q_2, the only configuration H for which equality holds in Proposition 2.3.1 is two adjacent vertices. The smallest upper bound for the 3-cube density of any of the 22 possible configurations in Q_3 as computed by Rahil Baber using flag algebras is .3048 (when H is two adjacent vertices in Q_3, see Chapter 3), so it is highly unlikely that any configuration in Q_3 has 3-cube density equal to $\frac{2}{9}$, the lower bound provided by Proposition 2.3.1. Of the 238 possible configurations in Q_4, only three have flag algebra calculated upper bound 4-cube densities less than .1: one is the perfect 8-cycle, for which Theorem 2.2.1 says the exact value is $\frac{3}{32} = .09375$ and another is a graph theoretic, but not perfect, induced 8-cycle, with flag algebra 4-cube density upper bound .094205. So there seems to be something special about the perfect 8-cycle.

For the perfect path P_{d+1} in Q_d it turns out that a blow-up of C_{2d+2} gives a better lower bound than that provided by Proposition 2.3.1:

Proposition 2.3.2. $\lambda(P_{d+1}, d) \geq \dfrac{d!}{(d+1)^{d-1}}$ *for all positive integers d.*

Proof. Let S be a blow up of C_{2d+2}. That is we partition $[n]$ into $A_1, A_2, \ldots, A_{d+1}$ and let B be the set of binary $(d+1)$-tuples in a copy of C_{2d+2}.

For each vertex $v = (v_1, \ldots, v_n)$ in Q_n, we let $v(A_i)$ equal 1 or 0 according to $v(A_i) \equiv \sum_{j \in A_i} v_j \mod 2$. We put v in S if and only if the $(d+1)$-tuple $(v(A_j))_{j\in[d+1]}$ is in B. If a sub-d-cube has one coordinate in each of d parts (and none in the other), then it will contain an exact copy of P_{d+1}. For example, if $d = 3$ and $B = \{0000, 1000, 1100, 1110, 1111, 0111, 0011, 0001\}$ and we select a sub-3-cube with one coordinate in each of A_1, A_2, and A_4 (so each coordinate in A_3 is fixed) then if $v(A_3) = 0$ the 4-tuples $0001, 0000, 1000, 1100$ in B give us an exact copy of P_4 in any such sub-3-cube, while if $v(A_3) = 1$, then $1110, 1111, 0111, 0011$ does the same. If it is an equipartition, selecting the coordinates of the sub-d-cube one-by-one shows that

$$\lambda(P_{d+1}, d) \geq \frac{(d+1)!}{(d+1)^d} = \frac{d!}{(d+1)^{d-1}}.$$

■

2.4 Local density, perfect cycles, and sequences

Let H be a configuration in Q_d and S be a subset of V_n. For each vertex v in S, we let $G_{\text{v(in)}}(H, d, n, S)$ be the number of sub-d-cubes R of Q_n containing v in which $S \cap R$ is an exact copy of H, $G_{\max(\text{in})}(H, d, n) = \max_{v\in S} G_{\text{v(in)}}(H, d, n, S)$ where the max is over all v and S such that $v \in S$ and $\lambda_{\text{local(in)}}(H, d, n) = \frac{G_{\max(\text{in})}(H,d,n)}{\binom{n}{d}}$. Since there are $\binom{n}{d}$ sub-d-cubes which contain v, $\lambda_{\text{local(in)}}(H, d, n)$ is the maximum fraction, over all $v \in S \subseteq V_n$, of sub-d-cubes containing v which have an exact copy of H. As with $\lambda(H, d, n)$, a simple averaging argument shows that $\lambda_{\text{local(in)}}(H, d, n)$ is a nonincreasing function of n, so we define $\lambda_{\text{local(in)}}(H, d)$ by

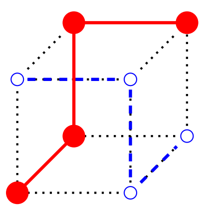
$$\lambda_{\text{local(in)}}(H, d) = \lim_{n\to\infty} \lambda_{\text{local(in)}}(H, d, n).$$

For each vertex $v \notin S$, a similar procedure defines the functions $G_{\text{v(out)}}(H, d, n, S)$, $G_{\max(\text{out})}(H, d, n)$, $\lambda_{\text{local(out)}}(H, d, n)$, and $\lambda_{\text{local(out)}}(H, d)$. This means $\lambda_{\text{local(in)}}(H, d)$ and $\lambda_{\text{local(out)}}(H, d)$ are the limit as n goes to infinity of the maximum fraction of sub-d-cubes of Q_n containing a particular vertex v which have an exact copy of H, for $v \in S$ and $v \notin S$ respectively. Finally, we define $\lambda_{\text{local}}(H, d)$ as $\max\{\lambda_{\text{local(in)}}(H, d), \lambda_{\text{local(out)}}(H, d)\}$.

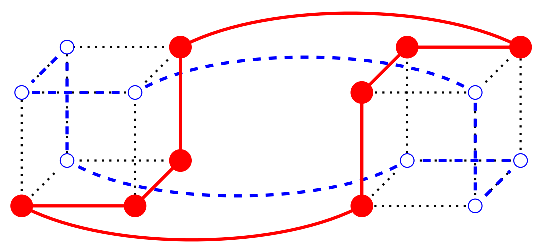
Since the global density cannot be more than the maximum local density, we must have $\lambda(H,d) \le \lambda_{\text{local}}(H,d)$.

For most configurations H for which we have been able to determine $\lambda(H,d)$, our procedure has been to prove an upper bound for $\lambda_{\text{local}}(H,d)$ which matches the density of a construction.

If H is a configuration in Q_d we let $\overline{H}$ denote $V(Q_d) \setminus H$. Clearly $\lambda(H,d) = \lambda(\overline{H},d)$ and $\lambda_{\text{local(in)}}(H,d) = \lambda_{\text{local(out)}}(\overline{H},d)$. If H is self-complementary in Q_d, i.e. $\overline{H}$ is an exact copy of H, then $\lambda_{\text{local(in)}}(H,d) = \lambda_{\text{local(out)}}(\overline{H},d) = \lambda_{\text{local(out)}}(H,d) = \lambda_{\text{local}}(H,d)$. Each of the six distinct configurations in Q_3 with 4 vertices is self-complementary, including P_4 (see Figure 2.1A), and C_8 is self-complementary in Q_4 (see Figure 2.1B). The complements of the two non-perfect induced 8-cycles in Q_4 are not 8-cycles.



A: P_4 and its complement in Q_3.



B: C_8 and its complement in Q_4.

Figure 2.1: Two self complementary configurations.

We now pose and solve a different maximization problem whose answer we will show to be $\lambda(C_8,4)$. Let S be a set of size n and d a positive integer. Let $M(d,n)$ be the set of all sequences of d distinct elements of S. Given a sequence w in $M(d,n)$ an *end-segment of* w is the set of the first j elements of w or the set of the last j elements of w, for some j in $[1,d)$. We say a subset $A(d,n)$ of $M(d,n)$ has Property U if the two following conditions are satisfied:

1. For each pair of sequences w and x in $A(d,n)$, if L is an end-segment of w and all elements of L are in the sequence x, then L is an end-segment of x with elements in the same order as in w (so if abc is the beginning of w, and a, b, and c all appear in x, then either x begins abc or ends cba).

2. A sequence and its reversal are not both in $A(d,n)$ (unless $d = 1$).

For example, if x and w are sequences in a set $A(5,n)$ with Property U and if x is $abcde$, then w cannot be $abceg$ (or its reversal), $abegh$ (or its reversal), or $ghiaj$ (or its reversal), but could be $fbdcg$ (or its reversal) or $edgbh$ (or its reversal). It is easy to see that no two sequences in $A(d,n)$ can have the same set of d elements.

Let $T(d,n)$ denote the maximum size of a family $A(d,n)$ with Property U.

Proposition 2.4.1. $G_{\max(\mathrm{in})}(C_{2d},d,n) = T(d,n)$ *for all* $d \geq 2$.

Proof. Without loss of generality, we can assume that $\emptyset$ is a vertex where the local d-cube density of C_{2d} is a maximum and that $G_{\emptyset(\mathrm{in})}(C_{2d},d,n,S) = G_{\max(\mathrm{in})}(C_{2d},d,n)$. Now we construct a set $A(d,n)$ of d-sequences.

The sequence $a_1, a_2, \ldots, a_d$ or its reversal is in $A(d,n)$ if and only if the sub-d-cube R containing $\emptyset$ where $a_1, a_2, \ldots, a_d$ are the nonconstant coordinates contains an exact copy of C_{2d}, say $S \cap R = \{\emptyset, a_1, a_1a_2, a_1a_2a_3, \ldots, a_1a_2\cdots a_k, a_2a_3\cdots a_d, \ldots, a_{d-1}a_d, a_d\}$. Note that $S \cap R$ contains $\emptyset$ and all the end-segments of the sequence $a_1, a_2, \ldots, a_d$.

We claim that $A(d,n)$ has Property U. Suppose it does not, say $x = a_1a_2\ldots a_d$ and $w = b_1b_2\ldots b_d$ are sequences in $A(d,n)$ with $b_1b_2\cdots b_j$ an end-segment in w all of whose elements are in x but not an end-segment in x. Then $\{b_1, b_2, \ldots, b_j\}$ is a subset of $\{a_1, a_2, \ldots, a_d\}$, so is another vertex in S which is in the sub-d-cube containing the perfect $2d$-cycle $\{\emptyset, a_1, a_1a_2, \ldots, a_{d-1}a_d, a_d\}$, a contradiction because this sub-d-cube has an exact copy of C_{2d}.

Similarly, by reversing the procedure, a family of sequences with Property U and size $T(d,n)$ can be used to construct $T(d,n)$ sub-d-cubes containing $\emptyset$ with exact copies of C_{2d}. ■

We define $t(d,n)$ to be $\frac{T(d,n)}{\binom{n}{d}}$. Hence $t(d,n) = \frac{G_{\max(\mathrm{in})}(C_{2d},d,n)}{\binom{n}{d}} = \lambda_{\mathrm{local(in)}}(C_{2d},d,n)$ is a nonincreasing function of n, so we can define $t(d)$ by setting $t(d) = \lim_{n\to\infty} t(d,n) = \lim_{n\to\infty} \lambda_{\mathrm{local(in)}}(C_{2d},d,n) = \lambda_{\mathrm{local(in)}}(C_{2d},d)$. Hence we have

Proposition 2.4.2. *For all* $d \geq 2$

$$\lambda_{\mathrm{local(in)}}(C_{2d},d) = t(d).$$

We now calculate $t(3)$.

Let $A(3,n)$ be a set of 3-sequences with Property U. No symbol can appear at the end in one sequence and in the middle of another, so we let D be the set of symbols which appear at the beginning or end and E be the set of symbols which appear in the middle. If $|D| = m$ and $|E| = p \leq n-m$, then, since a sequence and its reversal cannot both be in $A(3,n)$, the total number of sequences is at most $\binom{m}{2}p \leq \frac{(n-m)m(m-1)}{2}$ which is maximized when $m = \lceil \frac{2n}{3} \rceil$. Hence,

$$\lambda_{\mathrm{local(in)}}(C_6,3) = t(3) \leq \lim_{n\to\infty} \frac{\left(\frac{2}{3}n\right)^2 \left(\frac{1}{3}n\right)}{2\binom{n}{3}} = \frac{4}{9}.$$

We can construct a set $A(3,n)$ with Property U by partitioning $[n]$ into sets D and E with $|D| = \lceil \frac{2n}{3} \rceil$ and $|E| = \lfloor \frac{n}{3} \rfloor$ by putting one of the sequences abc and cba in $M(3,n)$ into $A(3,n)$ if and only if $a, c \in D$ and $b \in E$. Since $|A(3,n)| = \frac{\binom{|D|}{2}|E|}{\binom{n}{3}}$, we have $t(3) \geq \lim_{n\to\infty} \frac{|A(3,n)|}{\binom{n}{3}} = \lim_{n\to\infty} \frac{\frac{1}{2}\left(\frac{2n}{3}\right)^2 \frac{n}{3}}{\frac{n^3}{6}} = \frac{4}{9}$. Hence

$$\lambda_{\text{local(in)}}(C_6, 3) = t(3) = \frac{4}{9}.$$

To find $\lambda_{\text{local(out)}}(C_6, 3)$, we just note that if S is the set of all vertices in Q_n with weight not divisible by 3, then every Q_3 containing $\emptyset$ has an exact copy of the 6-cycle (the unique vertex with weight 3 in each Q_3 containing $\emptyset$ is also not in the 6-cycle), so $\lambda_{\text{local(out)}}(C_6, 3) = 1$. Using this same set S, it is not hard to show that $\lambda(C_6, 3) \geq \frac{1}{3}$ (any Q_3 whose smallest weight vertex is a multiple of 3 has an exact copy of C_6). We have been unable to show equality, but Baber's flag algebra upper bound of .3333333336 would seem to imply equality should hold.

To prove Theorem 2.2.1, we will prove a result about inducibility in bipartite graphs.

Theorem 2.4.3. *Let G be a bipartite graph with n vertices. Then the limit as n goes to infinity of the maximum fraction of sets of 4 vertices of G which induce two disjoint edges is equal to $\frac{3}{32}$. The unique optimizing graph when n is divisible by 4 is two disjoint copies of $K_{\frac{n}{4},\frac{n}{4}}$.*

Proof. Suppose M, P is a bipartition of $V(G)$ where $|M| = m$ and $|P| = p$. Let $\{u_1, u_2, \ldots, u_m\}$ and $\{v_1, v_2, \ldots, v_p\}$ be the vertices of M and P with respective degrees $r_1, r_2, \ldots, r_m$ and $c_1, c_2, \ldots, c_p$. For $i \neq j$, let $t_{i,j}$ denote the number of vertices in P which are adjacent to both u_i and u_j. Hence the total number of "good" sets of 4 vertices is

$$N = \sum_{i<j} (r_i - t_{i,j})(r_j - t_{i,j})$$

where the sum is over all pairs i, j such that $1 \leq i < j \leq m$. To get an upper bound for this we first get an upper bound on the sum S of all pairs of the factors in the products:

$$S = \sum_{i<j} [(r_i - t_{i,j}) + (r_j - t_{i,j})] = (m-1)\sum_{i=1}^{m} r_i - 2\sum_{i=1}^{p} \binom{c_j}{2}.$$

This is because each r_i appears in a sum with each r_j where $j \neq i$ and because $\sum_{i=1}^{m} t_{i,j} = \binom{c_j}{2}$ since each pair of edges adjacent to v_j is counted precisely once in the sum.

Let $w = \sum_{i=1}^{m} r_i = \sum_{j=1}^{p} c_j$. Then

$$\begin{aligned} S &= (m-1)\sum_{i=1}^{m} r_i - \sum_{j=1}^{p} c_j^2 + \sum_{j=1}^{p} c_j \\ &= mw - \sum_{j=1}^{p} c_j^2 \\ &\le mw - \sum_{j=1}^{p} \left(\frac{w}{p}\right)^2 \\ &= mw - \frac{w^2}{p} \end{aligned}$$

where the inequality is by Cauchy-Schwartz.

The function $f(w) = mw - \frac{w^2}{p}$ is maximized when $w = \frac{mp}{2}$, so $S \le \frac{m^2 p}{4}$.

Now, we return to our consideration of $N = \sum_{i<j}(r_i - t_{i,j})(r_j - t_{i,j})$.

The product $(r_i - t_{i,j})(r_j - t_{i,j})$ is at most $\left(\frac{p}{2}\right)^2$, achieved when $r_i = r_j = \frac{p}{2}$ and $t_{i,j} = 0$, in which case $(r_i - t_{i,j}) + (r_j - t_{i,j}) = p$. Since $S \le \frac{m^2 p}{4}$, to maximize N the two factors in each product should be equal, which reduces the problem to maximizing

$$\sum_{k=1}^{\binom{m}{2}} x_k^2 \text{ where } x_k \in \left[0, \frac{p}{2}\right] \text{ and } \sum x_k = \frac{m^2 p}{8}.$$

To do this, we clearly want each x_k to be equal to either 0 or $\frac{p}{2}$, so we want to have $\frac{m^2}{4}$ products of $\left(\frac{p}{2}\right)^2$, with all other products being $0 \cdot 0$. Hence $N \le \frac{m^2 p^2}{16}$. Since $m + p = n$, this is maximized when $m = p = \frac{n}{2}$, so $N \le \frac{n^4}{256}$. Equality can hold only if n is divisible by 4 and there are $\frac{n^2}{16}$ summands, each of them equal to $\left(\frac{n}{4}\right)^2$, so G must be two disjoint copies of $K_{\frac{n}{4},\frac{n}{4}}$.

This gives a fraction of "good" sets of 4 vertices as

$$\frac{\frac{n^4}{256}}{\binom{n}{4}} = \frac{n^3}{(n-1)(n-2)(n-3)} \cdot \frac{3}{32}.$$

∎

Interestingly, two disjoint copies of $K_{\frac{n}{4},\frac{n}{4}}$ is also the graph which maximizes, among all graphs with n vertices, the number of induced subgraphs with 4 vertices consisting of two edges which share a vertex and an isolated vertex [20].

We remark that the inducibility of $2K_2$ (among all host graphs G, not just bipartite) is $\frac{3}{8}$, the maximum density achieved when $G = 2K_{\frac{n}{2}}$ [20]. Stated differently, Theorem 2.4.3 says that if $m + p = n$ then the $m \times p$ (0,1)-matrix with the maximum number of 2×2 submatrices which have precisely two 1s in different rows and columns is an equi-blow-up of I_2.

Proof of Theorem 2.2.1. By Proposition 2.3.1, $\lambda(C_8, 4) \geq \frac{3}{32}$. Since C_8 is self-complementary in Q_4, $\lambda(C_8, 4) \leq \lambda_{\text{local}}(C_8, 4) = \lambda_{\text{local(in)}}(C_8, 4) = t(4)$ the last equality by Proposition 2.4.2. So to complete the proof we just need to show that $t(4) \leq \frac{3}{32}$.

Let $A(4, n)$ be a maximum size set of 4-sequences with Property U with elements from $[n]$. Let $A = \{i \in [n] : i$ is the first or last element in a sequence in $A(4, n)\}$ and let $B = [n] \setminus A$. We construct a bipartite graph G with vertex bipartition A, B. If $a \in A$ and $b \in B$, then $[a, b]$ is an edge of G if and only if a and b are consecutive elements in some sequence in $A(4, n)$ (so some sequence begins ab or ends ba).

Suppose $a_1 b_1 b_2 a_2$ is a sequence in $A(4, n)$. Then $[a_1, b_1]$ and $[a_2, b_2]$ are edges of G. Suppose $a_1 b_2$ (or $a_2 b_1$) is also an edge. Then $a_1 b_2$ is an end-segment of some sequence in $A(4, n)$, which is impossible because $\{a_1, b_2\} \subseteq \{a_1, b_1, b_2, a_2\}$, but $a_1 b_2$ is not an end-segment in $a_1 b_1 b_2 a_2$. Hence the size $T(4, n)$ of $A(4, n)$ is at most the number of sets of 4 vertices in G which induce two disjoint edges, and by Theorem 2.4.3,

$$t(4) = \lim_{n \to \infty} \frac{T(4, n)}{\binom{n}{4}} \leq \frac{3}{32}.$$

■

So Conjecture 2.2.3 is true for $d = 4$. Zongchen Chen [14] has shown that $t(d) = \frac{d!}{d^d}$ for all $d \geq 4$, so we know that $\lambda_{\text{local(in)}}(C_{2d}, d) = t(d) = \frac{d!}{d^d}$ for all $d \geq 4$. However, we have been unable to show that $\lambda_{\text{local(out)}}(C_{2d}, d) = \lambda_{\text{local(in)}}(C_{2d}, d)$ if $d \geq 5$ (our proof for $d = 4$ used the fact that C_8 is self-complementary in Q_4), which would complete a proof of Conjecture 3.

We have seen that equality in Conjecture 2.2.3 does not hold for $d = 3$ (since $\lambda(C_6, 3) \geq \frac{1}{3}$).

2.5 Perfect Paths

To determine the d-cube density of P_4 in Q_3, as mentioned in Section 2.4, our procedure is to prove an upper bound for $\lambda_{\text{local}}(P_4, 3)$ which matches the density of the construction given in Proposition 2.3.2. We will show

Proposition 2.5.1. $\lambda_{\text{local}}(P_4, 3) \leq \dfrac{3}{8}$.

Let S be a set of size n and d a positive integer. Let $P(d,n)$ be the family of pairs of sequences $x = \{x_1; x_2\}$ of elements in S, which we call *d-bisequences*, one sequence of length k and the other of length $d-k$, where $k \in [0,d]$, and where the d elements in the pair of sequences $\{x_1, x_2\}$ are distinct. Given a bisequence $x = \{x_1; x_2\}$ in $P(d,n)$ an *initial segment of* x is the set of the first j elements of either x_i where $j \in [1,d]$ and $i = 1$ or 2. We also say these j elements are an initial segment in x_i. We say that a subset $R(d,n)$ of $P(d,n)$ has Property V if the following conditions are satisfied:

1. For each pair of bisequences $w = \{w_1; w_2\}$ and $x = \{x_1; x_2\}$ in $R(d,n)$, if L is an initial segment of w and all elements of L are in x, then L is an initial segment of x_1 or x_2, with elements in the same order as in w.

2. The bisequences $\{x_1; x_2\}$ and $\{x_2; x_1\}$ are not both in $R(d,n)$.

Let $B(d,n)$ denote the maximum size of a family of bisequences $R(d,n)$ with property V and let $b(d,n) = \frac{B(d,n)}{\binom{n}{d}}$.

Proposition 2.5.2. $\lambda_{\text{local(in)}}(P_{d+1}, d, n) = b(d,n)$.

Proof. Without loss of generality, we can assume that $\emptyset \in S$ and $\emptyset$ is a vertex where the local d-cube density of P_{d+1} is a maximum. Now we construct a set $R(d,n)$ of bisequences.

The bisequence $\{(a_1, a_2, \ldots, a_j); (b_1, b_2, \ldots, b_i)\}$ or its reversal is in $R(d,n)$ if and only if the intersection of S and the sub-d-cube containing $\emptyset$ where $a_1, a_2, \ldots, a_j, b_1, b_2, \ldots, b_i$ are the nonconstant coordinates is precisely equal to $\{a_1 a_2 \cdots a_j, a_1 a_2 \cdots a_{j-1}, \ldots, a_1, \emptyset, b_1, b_1 b_2, \ldots, b_1 b_2 \cdots b_i\}$. Note that i or j could be equal to 0. We claim that $R(d,n)$ has property V.

Suppose it does not, say w and x are bisequences in $R(d,n)$ with $a_1 a_2 \cdots a_l$ an initial-segment of w all of whose elements are in x but not an initial-segment (or not in the same order as in w) of either of the sequences in $x = \{(b_1, b_2, \ldots, b_j); (c_1, c_2, \ldots, c_i)\}$. Then $\{\emptyset, a_1, a_1 a_2, \ldots, a_1 a_2 \cdots a_l\}$ is a path contained in the sub-d-cube containing the path $\{b_1 b_2 \cdots b_j, b_1 b_2 \cdots b_{j-1}, \ldots, b_1, \emptyset, c_1, c_1 c_2, \ldots, c_1 c_2 \cdots c_i\}$, but is not a subpath, a contradiction.

Similarly, reversing the procedure, a family of bisequences with Property V and size $B(d,n)$ can be used to construct $B(d,n)$ sub-d-cubes containing $\emptyset$ with exact copies of P_{d+1}. Hence $G_{\max(\text{in})}(P_{d+1}, d, n) = B(d,n)$, and dividing by $\binom{n}{d}$ gives the desired equality. ∎

Since $b(d,n) = \lambda_{\text{local(in)}}(P_{d+1}, d, n)$ is a non-increasing function of n, we can define $b(d)$ to be equal to $\lim\limits_{n\to\infty} b(d,n)$. So $b(d)$ is the limit as n goes to infinity of the maximum fraction of d-subsets of n which can be the sets of elements of a family $R(d,n)$ of bisequences with Property V. We have

$$\lambda_{\text{local(in)}}(P_{d+1}, d) = \lim_{n\to\infty} \lambda_{\text{local(in)}}(P_{d+1}, d, n) = \lim_{n\to\infty} b(d,n) = b(d)$$

and we can find $\lambda_{\text{local(in)}}(P_{d+1}, d)$ by finding $b(d)$.

Clearly if $R(d,n)$ is a family of d-bisequences with Property V, then no symbol can appear as the first element of some sequence and not as the first element of another. Furthermore, the following properties are easy to verify.

Lemma 2.5.3. *Suppose $R(3,n)$ is a family of 3-bisequences with Property V.*

(i) If $\{bxy;\emptyset\}$ and $\{bxz;\emptyset\}$ are in $R(3,n)$, then $\{byz;\emptyset\}$ is not.

(ii) If $\{bxz;\emptyset\}$ and $\{byz;\emptyset\}$ are in $R(3,n)$ then $\{bxy;\emptyset\}$ is not.

(iii) If $\{bx;c\}$ and $\{bx;d\}$ are in $R(3,n)$, then $\{dx;c\}$ is not.

(iv) If $\{bx;c\}$ and $\{dx;c\}$ are in $R(3,n)$, then $\{dx;b\}$ is not.

Proof. Let $A = \{i \in [n] : i$ is the first element of some sequence in a 3-bisequence in $R(3,n)\}$ and let $W = [n] \setminus A$. Let $a = \frac{|A|}{n}$ and $w = \frac{|W|}{n} = 1 - a$.

(i) If the assumption in (i) holds, then $\{byz;\emptyset\}$ cannot be in $B(3,n)$ because it has "by" as an initial segment, and that is a subset of one of the sequences in $\{bxy;\emptyset\}$ but is not an initial segment, violating property V.

Statements (ii), (iii), and (iv) are just as easy to verify. ∎

Proof of Proposition 2.5.1. By Proposition 2.5.2 it suffices to show that $b(3) \le \frac{3}{8}$. Let $R(3,n)$ be a family of 3-bisequences with Property V. Let A be the set of elements in $[n]$ which appear as the first element of some sequence in $R(3,n)$, let $W = [n] \setminus A$, and let $a = |A|$ and $w = |W|$. For each $e \in A$, let G_e be the graph with vertex set W and edge set $\{[x,y] : \{eyx;\emptyset\}$ or $\{exy;\emptyset\}$ or one of their reversals is in $R(3,n)\}$. By statements (i) and (ii) in Lemma 2.5.3, G_e is a triangle free graph for each $e \in A$. Hence by Turán's theorem, at most $\frac{w^2}{4}$ unordered pairs (x,y) of element x and y in W can appear as edges in G_e. That means that the total number of 3-subsets of $[n]$ which can be the set of elements of a 3-bisequence in $B(3,n)$ with one element in A and two in W is at most $\frac{w^2}{4} \cdot a$. Similarly, for each element x in W we let G_x be the graph with vertex set A and edge set $\{[b,c] : \{bx,c\}$ or $\{bc,x\}$ or either of their reversals is in $R(3,n)\}$. By statements (iii) and (iv) in Lemma 2.5.3, G_x is triangle free, and an identical argument to the one for G_e shows that the total number of 3-subsets of $[n]$ which can be the set of elements of a 3-bisequence in $B(3,n)$ with two elements in A and one in W is at most $w \cdot \frac{a^2}{4}$. If $B(3,n)$ is the size of $R(3,n)$ then

$$\begin{aligned} B(3,n) &\leq a \cdot \frac{w^2}{4} + w \cdot \frac{a^2}{4} \\ &= \frac{aw}{4} \cdot n \end{aligned}$$

This is maximized when $a = w = \frac{n}{2}$, so $B(3,n) \leq \frac{n^3}{16}$ and $b(3) = \lim_{n\to\infty} \frac{B(3,n)}{\binom{n}{3}} \leq \frac{3}{8}$. ∎

Further, this also shows Theorem 2.2.2 holds.

Proof of Theorem 2.2.2. By Proposition 2.3.2 and Proposition 2.5.1

$$\frac{3}{8} \leq \lambda(P_4, 3) \leq \lambda_{\text{local}}(P_4, 3) \leq \frac{3}{8}.$$

∎

So Conjecture 2.2.4 holds for $d = 3$. Lending credence to this conjecture is that Baber's flag algebra upper bound for $\lambda(P_5, 4)$ is .19200000058, while the conjecture with $d = 4$ gives $\frac{24}{125} = .192$.

2.6 Open Problems

In this section, we racapitulate the main conjectures and open problems suggested in this chapter.

Conjecture 2.2.3. $\lambda(C_{2d}, d) = \frac{d!}{d^d}$ for all $d \geq 4$.

As mentioned in Section 2.4 Zongchen Chen [14] has shown that $t(d) = \frac{d!}{d^d}$ for all $d \geq 4$, so $\lambda_{\text{local(in)}}(C_{2d}, d) = \frac{d!}{d^d}$ for all $d \geq 4$.

To prove Conjecture 2.2.3 it would suffice to show that $\lambda_{\text{local(out)}}(C_{2d}, d) = \frac{d!}{d^d}$. The difficulty in doing this is that a vertex $v \notin S$ with maximum local density could lie at different distances from the perfect $2d$-cycles in two good sub-d-cubes in which it lies.

As mentioned in the discussion before Theorem 2.4.3, a simple construction shows that $\lambda(C_6, 3) \geq \frac{1}{3}$ and Baber's flag algebra result is that $\lambda(C_6, 3) \leq .3333333336$, so the obvious conjecture is:

Conjecture 2.6.1. $\lambda(C_6, 3) = \frac{1}{3}$.

In Section 2.1, we mentioned that we have shown that $\frac{2}{3} \leq \lambda(W_2, 2) \leq .686$, the upper bound from Baber's flag algebra result.

Conjecture 2.6.2. $\lambda(W_2, 2) = \frac{2}{3}$.

Conjecture 2.2.4. $\lambda(P_{d+1}, d) = \frac{d!}{(d+1)^{d-1}}$ for all $d \geq 3$.

In Section 2.1 we mentioned that we believe that for sufficiently large d, choosing vertices to be in S with uniform probability $\left(\frac{1}{2}\right)^d$ is an optimal construction to maximize the number of exact copies of W_d in Q_d where W_d is a single vertex.

Conjecture 2.6.3. $\lim_{d\to\infty} \lambda(W_d, d) = \frac{1}{e}$.

Recall that the perfect 6-cycle in Q_3 can be described by saying it is the set of all vertices in $V(Q_d)$ with weight 1 or 2. That it can be described in this way led to two phenomena:

1. $\lambda_{\text{local(out)}}(C_6, 3) = 1$
2. The weight pattern in C_6 suggested a set S in Q_n which seems to maximize the number of exact copies of C_6: $S = \{v \in V_n : \operatorname{wt} v \text{ is not divisible by } 3\}$.

Let W be a subset of $[d]$ and H be the configuration in Q_d defined by $H = \{v \in V(Q_d) : \operatorname{wt} v \in W\}$. The two phenomena mentioned above will still hold. What more can be said about this type of configuration?

Chapter 3

Maximum densities of other vertex-induced substructures in a hypercube

The material for this chapter has not currently been submitted for journal publication.

3.1 Introduction

In Chapter 2 and [26] we initiated the investigation of d-cube-density. Using a kind of *blow–up*, we showed that $\lambda(H, d) \geq \frac{d!}{d^d}$ for every configuration H in Q_d. We defined a perfect $2d$-cycle C_{2d} in Q_d to be a cycle with d pairs of vertices each Hamming distance d apart. We showed that $\lambda(C_8, 4) = \frac{4!}{4^4}$, achieving the smallest possible value for any configuration in Q_4. We also showed $\lambda(P_4, 3) = \frac{3}{8}$ where P_4 is the induced path in Q_3 with 4 vertices.

Finding d-cube density seems to be very difficult even for most small configurations. In this paper, we find the d-cube density for one configuration in Q_3 and two configurations in Q_4. We find a construction to produce a lower bound and then find a matching upper bound by using known results on the inducibility of small graphs to show the local density cannot be larger.

In Section 3.2 we again discuss local d-cube density, the notion we use to find the upper bounds. In Section 3.3 we consider the possible configurations in Q_2. In Section 3.4 we summarize the results on inducibility of graphs which we will use for configurations in Q_3 and Q_4. In Section 3.5 we consider d-cube density for configurations in Q_3, and in Section 3.6 we consider several configurations in Q_4. In Section 3.7 we find d-cube density for a nontrivial infinite family of configurations. In Section 3.8 we discuss *layered* configurations in Q_d, those that are defined in terms of the weights of the d-vectors.

3.2 Local d-cube density

As in 2.4, we let H be a configuration in Q_d and S be a subset of V_n. For each vertex $v \in S$, we let $G_{\mathrm{v(in)}}(H, d, n, S)$ be the number of sub-d-cubes R of Q_n containing v in which $S \cap R$ is an exact copy of H, $G_{\max(\mathrm{in})}(H, d, n) = \max_{v \in S} G_{\mathrm{v(in)}}(H, d, n, S)$ where the max is over all v and S such that $v \in S$, $g_{\mathrm{v(in)}}(H, d, n, S) = \frac{G_{\mathrm{v(in)}}(H,d,n,S)}{\binom{n}{d}}$ denote the fraction of sub-d-cubes R of Q_n containing v in which $S \cap R$ is an exact copy of H, and $\lambda_{\mathrm{local(in)}}(H, d, n) = \frac{G_{\max(\mathrm{in})}(H,d,n)}{\binom{n}{d}}$. As with $\lambda(H, d, n)$, a simple averaging argument shows that $\lambda_{\mathrm{local(in)}}(H, d, n)$ is a nonincreasing function of n, so we define $\lambda_{\mathrm{local(in)}}(H, d)$ by

$$\lambda_{\mathrm{local(in)}}(H, d) = \lim_{n\to\infty} \lambda_{\mathrm{local(in)}}(H, d, n)$$

For each vertex $v \notin H$, a similar procedure defines the functions $G_{\mathrm{v(out)}}(H, d, n, S)$, $G_{\max(\mathrm{out})}(H, d, n)$, $g_{\mathrm{v(out)}}(H, d, n, s)$, and $\lambda_{\mathrm{local(out)}}(H, d)$. This menas $\lambda_{\mathrm{local(in)}}(H, d)$ and $\lambda_{\mathrm{local(out)}}(H, d)$ are the maximum local densities of sub-d-cubes with an exact copy of H among all sub-d-cubes containing v in S and out of S respectively. Finally, we define $\lambda_{\text{local}}(H,d)$ to be $\max\{\lambda_{\text{local(in)}}(H,d), \lambda_{\text{local(out)}}(H,d)\}$. Since the global density cannot be more than the maximum local density, we must have $\lambda(H,d) \leq \lambda_{\text{local}}(H,d)$.

3.3 Configurations in Q_2

The following type of construction is referred to as a *partition-modular construction*. These are constructions generated by choosing a partition of $[n] = A_1 \cup A_2 \cup \cdots \cup A_i$ and taking as S the set of vertices such that their binary n-tuples satisfy a chosen set of congruences for the weight of the coordinates within the partitions. Sometimes we denote this as a list of i-tuples along with a list of their moduli for convenience. For example, $A \cup B$ taking 01 mod $(2,2)$ would indicate a partitioning of $[n] = A \cup B$ and taking all vertices with weight 0 mod 2 in A and weight 1 mod 2 in B. The fractional sizes of the A_i which maximize the number of Q_ds having the configuration may also be indicated.

Note that the sets in the partition may be of any sizes, however, when $i = 1$ we call such a construction *layered* since it is equivalent to choosing all vertices of particular weights modulo a (i.e. entire "levels" of Q_n).

It is obvious that $\lambda(H) = \lambda\left(\overline{H}\right)$, thus we may restrict our consideration to only one configuration in each of the complementary pairs.

A list of all of the configurations in Q_2, subject to the above restriction, are given in Figure 3.1. In the figure, red vertices are in the configuration and open blue are not. The red and blue edges have been added for emphasis but edge choices are not the focus of this chapter as the configurations are sets of vertices.

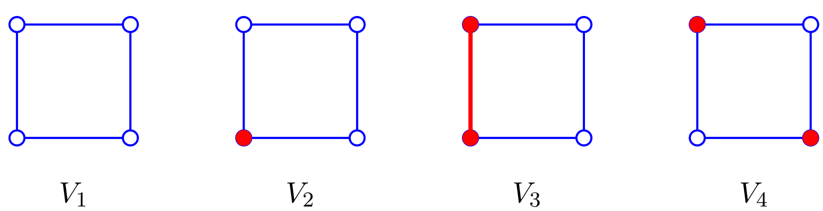


Figure 3.1: Configurations in Q_2.

3.3.1 Lower Bounds by Construction

Showing $\lambda(V_1) = 1$ is trivial, since we would simply consider $S = \emptyset$. To show $\lambda(V_4) = 1$, we consider S to be the layered construction $0 \mod 2$. This leaves only V_2 and V_3.

We can find a lower bound for V_2 by considering the layered construction given by $0 \mod 3$. This gives $\frac{2}{3} \leq \lambda(V_2)$. However, an upper bound other than the flag algebra bound provided in Table 3.1 for this configuration remains elusive at this time.

A construction for V_3 is given by considering $[n] = A \cup B$ then taking S to be the set of all vertices given by binary n-tuples with weight $0 \mod 2$ in A. This gives a "good" Q_2 for each Q_2 with one flip bit in each of A and B. If $|A| = a$ and $|B| = b$, we then want to maximize ab which occurs when $|A| = \lfloor \frac{n}{2} \rfloor$ and $|B| = \lceil \frac{n}{2} \rceil$. This results in $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil \approx \frac{n^2}{4}$ many "good" Q_2s which shows $\frac{1}{2} \leq \lambda(V_3)$.

Table 3.1 summarizes the best results obtained in Q_2.

Configuration	Construction	Lower Bound	Upper Bound
V_1	$\emptyset$	1	1
V_2	Layered: $0 \mod 3$	2/3	.685714[1]
V_3	$A \cup B$ taking $0 \mod 2$ in A	1/2	1/2 (Theorem 3.3.1)
V_4	Layered: $0 \mod 2$	1	1

Table 3.1: Summary of the best results for configurations in Q_2.

3.3.2 Upper Bounds

In order to show that $\lambda(V_3) = \frac{1}{2}$, we use an argument that will be applied, in a slightly more general form, to an infinite family of configurations in Section 3.7.

Theorem 3.3.1. $\lambda(V_3) = \frac{1}{2}$.

Proof. Recall in Section 3.3.1 we showed $\lambda(V_3) \geq \frac{1}{2}$.

Now let S be a set which achieves $\lambda(V_3)$, Let $\alpha_x = \dfrac{|\{v \in N(x) \cap S\}|}{n}$, the density of neighbors of x in S. Consider $s \in S$ and let R_0, R_1, R_2 be the fraction of $K_{1,2}$ subgraphs of Q_n containing s in which s is degree 2 and has 0, 1, or 2 chosen neighbors, respectively.

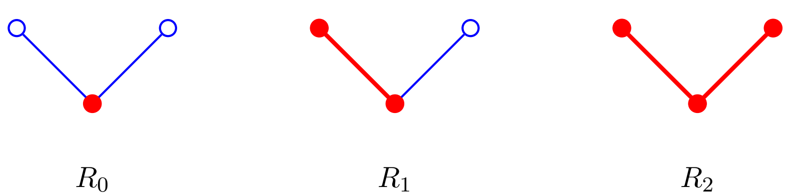


Figure 3.2: The red vertices are vertices in S and the blue are vertices not in S.

[1]The upper bound for V_2 is given by a flag algebra bound calculated by Rahil Baber [6].

Note that $R_0 + R_1 + R_2 = 1$ and we want to maximize R_1. We do this by minimizing $R_0 + R_2$ which is given by $f(\alpha_s) = (1-\alpha_s)^2 + \alpha_s^2 = 2\alpha_s^2 - 2\alpha_s + 1$ and so it is clear that $\alpha_s = \frac{1}{2}$. This gives a minimum value of $f\left(\frac{1}{2}\right) = \frac{1}{2}$. This means that $R_1 \leq \frac{1}{2}$.

Since V_3 is self-complimentary, $\lambda_{\text{local(out)}}(V_3, 2) = \lambda_{\text{local(in)}}(V_3, 2) = \lambda_{\text{local}}(V_3, 2)$ and so $\lambda(V_3) \leq \frac{1}{2}$. ■

3.4 Inducibility

As mentioned in Section 2.1, there are strong connections between d-cube density and *inducibility* of a graph. Recall that, given graphs G and H, with $|V(G)| = n$ and $|V(H)| = k$, the *density* of H in G, denoted $\mathrm{d}_H(G)$, is defined by

$$\mathrm{d}_H(G) = \frac{\#\text{ induced copies of } H \text{ in } G}{\binom{n}{k}}$$

Pippinger and Golumbic [40] defined the inducibility $i(H)$ of H by

$$i(H) = \lim_{n\to\infty} \max_{|V(G)|=n} \mathrm{d}_H(G)$$

Clearly $i(H) = i(\overline{H})$ where $\overline{H}$ is the complement of H. We summarize a few inducibility results, some of which we will use to prove upper bounds for d-cube density.

1. $i(K_{1,2}) = \frac{3}{4}$. The optimizing graph G is $K_{\frac{n}{2},\frac{n}{2}}$. That it cannot be larger than $\frac{3}{4}$ follows immediately from a theorem of Goodman [30] that says that in any 2-coloring of the edges of K_n, at least $\frac{1}{4}$ (asymptotically) of the K_3s are monochromatic.

2. $i(K_{2,2}) = \frac{3}{8}$. In [8], Bollobás et. al. showed that the graph on n vertices which has the most induced copies of $K_{r,r}$, for any $r \geq 2$, is $K_{\lceil\frac{n}{2}\rceil,\lfloor\frac{n}{2}\rfloor}$.

3. $i(K_{1,3}) = \frac{1}{2}$. In [12], Brown and Siderenko showed that the graph on n vertices which has the most induced copies of $K_{r,s}$, for any r, s (except $r = s = 1$), is complete bipartite. The optimizing graph for $K_{1,3}$ is not equibipartite; the sizes of the parts are roughly $\frac{n}{2} \pm \sqrt{n}$.

4. In [33], Hirst used flag algebras to show that $i(K_{1,1,2}) = \frac{72}{125} = .576$ and $i(K_{\text{PAW}}) = \frac{3}{8}$ where $K_{1,1,2}$ is K_4 minus an edge and K_{PAW} is K_3 plus a pendant edge, leaving the path as the only graph on 4 vertices whose exact inducibility has yet to be determined.

5. In [20], Even-Zohar and Linial improve earlier best bounds [21, 43] for $i(P_4)$ and find the inducibility of some graphs on 5 vertices.

3.5 Configurations in Q_3

Recalling $\lambda(H) = \lambda\left(\overline{H}\right)$, we may restrict ourselves to only considering one configuration in each complementary pair, a list of all of the configurations in Q_3, subject to this restriction, are given in Figure 3.3.

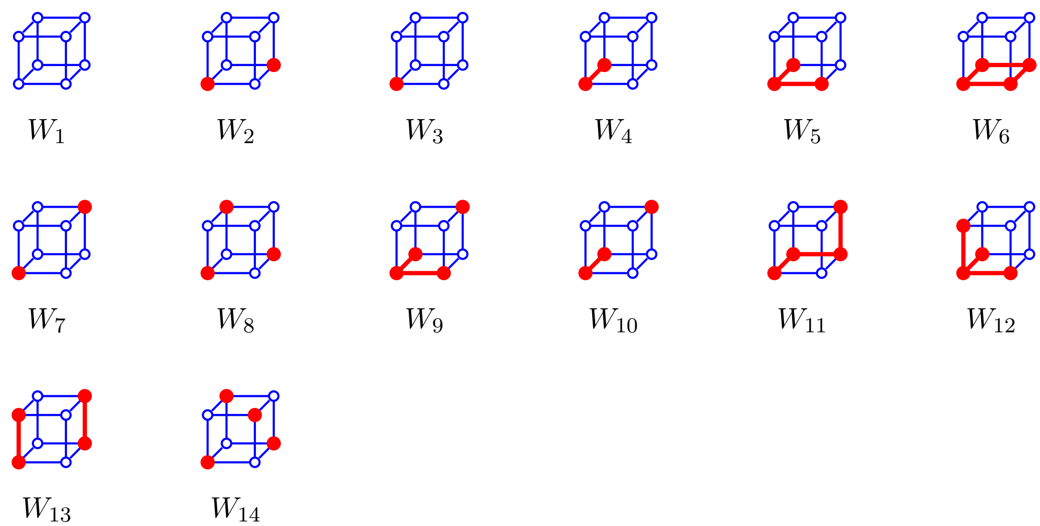


Figure 3.3: Configurations in Q_3.

3.5.1 Trivial configurations

Showing $\lambda(W_1) = 1$ is trivial since we take $S = \emptyset$. Further, $\lambda(W_{14}) = 1$ since we can consider S given by the layered construction given by $0 \mod 2$.

3.5.2 Layered constructions

Recalling the definition of "layered" from Section 3.3 these constructions choose all vertices of particular weights modulo m, for some m (i.e. entire "levels" of Q_n).

The layered constructions given in Table 3.2 for W_7, W_8, and W_{12} provide lower bounds which agree with Baber's flag algebra upper bounds to within 10^{-9}, so they are likely to be exact. We do not know if our lower bound for $\lambda W_3, 3$ is the actual value.

3.5.3 Other partition modular constructions

For W_4, we partition $[n] = A \cup B$ and take S to be the set of all vertices given by binary n-tuples with weight 0 mod 3 in A. Suppose a Q_3 contains precisely two flip bits in A. If the sum of the other bits in A is m, then the Q_3 will have configuration W_4 precisely when $m = 0$, or 1. When $m = 2$, it will have configuration W_{13}. Thus we will have a "good" Q_3 for 2/3 of the Q_3s with exactly two flip bits in A. If we let $|A| = a$ and $|B| = b$, then we want to maximize the function $\frac{2}{3} \cdot \frac{1}{2}a^2 b$ which occurs when $|A| = \lceil \frac{2n}{3} \rceil$ and $|B| = \lfloor \frac{n}{3} \rfloor$. This results in $\frac{1}{3} \lceil \frac{2n}{3} \rceil^2 \lfloor \frac{n}{3} \rfloor \approx \frac{4n^3}{81}$ many "good" Q_3s which shows $\frac{8}{27} = \left(\frac{2}{3}\right)^3 \leq \lambda(W_4)$.

When considering W_9, we partition $[n] = A \cup B$ and take S to be the set of all vertices given by binary n-tuples with weight 0 or 1 mod 3 in A and weight 2 mod 3 in B or vice versa. Suppose a Q_3 contains precisely two flip bits in A. If the sum of the other bits in A is m and in B is p, then the Q_3 will have configuration W_9 precisely when $m = 0, 2$ and $p = 1, 2$. This means that $\frac{4}{9}$ of these will be "good" Q_3s. Similarly, $\frac{4}{9}$ of the Q_3s will be "good" with precisely two flip bits in B. Thus, we want to maximize the function $\frac{2}{9}a^2 b + \frac{2}{9}ab^2$ which occurs when $|A| = \lceil \frac{n}{2} \rceil$ and $|B| = \lfloor \frac{n}{2} \rfloor$ giving $\frac{2}{9}n \lceil \frac{n}{2} \rceil \lfloor \frac{n}{2} \rfloor \approx \frac{n^3}{18}$ "good" Q_3s which shows $\frac{1}{3} \leq \lambda(W_9)$. This is likely not best possible since the flag algebra upper bound is just above $\frac{4}{9}$.

The constructions for W_6 and W_{13} use similar partition modular constructions to each other. For both, we take $[n] = A \cup B$ and S to be the set of all vertices given by binary n-tuples with weight $0 \mod 2$ in A.

For W_6, this gives a "good" Q_3 when we have exactly two flip bits in A. If we let $|A| = a$ and $|B| = b$, then to maximize the number of "good" Q_3s we want to maximize the function $\frac{1}{2}a^2 b$ which occurs when $|A| = \lceil \frac{2n}{3} \rceil$ and $|B| = \lfloor \frac{n}{3} \rfloor$. This results in $\frac{1}{2} \lceil \frac{2n}{3} \rceil^2 \lfloor \frac{n}{3} \rfloor \approx \frac{2n^3}{27}$ many "good" Q_3s which shows $\frac{4}{9} \leq \lambda(W_6)$.

For W_{13}, this construction gives a "good" Q_3 when we have exactly one flip bit in A. If we let $|A| = a$ and $|B| = b$, then to maximize the number of "good" Q_3s we want to maximize the function $\frac{1}{2}ab^2$ which occurs when $|A| = \lceil \frac{n}{3} \rceil$ and $|B| = \lfloor \frac{2n}{3} \rfloor$. This results in $\frac{1}{2} \lceil \frac{n}{3} \rceil \lfloor \frac{2n}{3} \rfloor^2 \approx \frac{2n^3}{27}$ many "good" Q_3s which shows $\frac{4}{9} \leq \lambda(W_{13})$.

Equality for the densities of W_6 and W_{13} will follow from Theorem 3.7.1 in Section 3.8 which is a generalization of Theorem 3.3.1.

A construction for W_2 is found by considering $[n] = A \cup B$ and taking S to be the set of all vertices given by binary n-tuples with weight 0 mod 2 in both A and B (i.e. weight 00). This gives a "good" Q_3 for each Q_3 with exactly two flip bits in A or exactly two in B. If we let $|A| = a$ and $|B| = b$, we then want to maximize the number of "good" Q_3s. This means we want to maximize $\frac{1}{2}a^2 b + \frac{1}{2}ab^2$ and so we find $|A| = \lfloor \frac{n}{2} \rfloor$ and $|B| = \lceil \frac{n}{2} \rceil$.

This results in $\frac{1}{2}\lfloor\frac{n}{2}\rfloor^2\lceil\frac{n}{2}\rceil+\frac{1}{2}\lfloor\frac{n}{2}\rfloor\lceil\frac{n}{2}\rceil^2 \approx \frac{n^3}{8}$ many "good" Q_3s which shows $\frac{3}{4} \leq \lambda(W_2)$. Theorem 3.5.1 shows equality holds.

Table 3.2 summarizes the best results obtained in Q_3. In Table 3.2, the upper bounds written in decimal form are flag algebra bounds calculated by Rahil Baber.

Configuration	Construction	Lower Bound	Upper Bound
W_1	$\emptyset$	1	1
W_2	$A \cup B$ taking 00 mod (2,2)	3/4	3/4 (Theorem 3.5.2)
W_3	Layered: 0 mod 4	1/2	.610043
W_4	$A \cup B$ taking 0 mod 3 in A where $a = \frac{2}{3}$	$(2/3)^3 \approx 0.2963$	.304762
W_5	$A \cup B$ taking 00,01,10,11 mod $(3,3)$	1/3	.333398
W_6	$A \cup B$ taking 0 mod 2 in A where $a = \frac{1}{3}$	4/9	4/9 (Theorem 3.7.1)
W_7	Layered: 0 mod 3	1/3	.333333
W_8	Layered: 0 mod 3	2/3	.666667
W_9	$A \cup B$ taking 02, 12, 20, 21 mod $(3,3)$	1/3	.444444
W_{10}			.416667
W_{11}	Perfect 8-cycle blow-up	3/8	3/8 (Chapter 2)
W_{12}	Layered: 0 and 1 mod 4	1/2	.500000
W_{13}	$A \cup B$ taking 0 mod 2 in A where $a = \frac{2}{3}$	4/9	4/9 (Theorem 3.7.1)
W_{14}	Layered: 0 mod 2	1	1

Table 3.2: Summary of the best results for configurations in Q_3.

The following Lemma is used in the proof that $\lambda(W_2) = \frac{3}{4}$.

Lemma 3.5.1. *Let G be a graph with n vertices where n is even. If $|E(G)| = e$, then G has at most* $\min\left\{n\binom{\frac{n}{2}}{2}, \frac{e}{2}(n-2)\right\}$ *induced copies of $K_{1,2}$.*

Proof. That it has at most $n\binom{\frac{n}{2}}{2}$ was proved in [40]. The optimizing graph is $K_{\frac{n}{2},\frac{n}{2}}$.

Each $uv \in E(G_s)$ can be in at most $n-2$ induced $K_{1,2}$s and summing over all edges uv counts each $K_{1,2}$ twice. ∎

Theorem 3.5.2. $\lambda(W_2) = \dfrac{3}{4}$.

Proof. Recall that we have a construction which shows that $\lambda(W_2) \geq \frac{3}{4}$.

Suppose $\emptyset \in S$, and let $\mathscr{M}$ be the set of good Q_3s containing $\emptyset$. Construct a graph G_s with $V(G_s) = [n]$ and $E(G_s) = \{uv : \emptyset, uv \text{ are the vertices in } S \text{ for some } M \in \mathscr{M}\}$. If u, v, x are flip bits for some M in $\mathscr{M}$, and if uv is in M, then neither ux nor vx can be in $E(G_s)$, so $|\mathscr{M}|$ is less than or equal to the number of induced copies of the graph with three vertices and a single edge. Equivalently, this is less than or equal to the number of induced copies of $K_{1,2}$ in a graph on n vertices. This means $\lambda_{\text{local(in)}}(W_2) \leq i(K_{1,2}) = \frac{3}{4}$.

Now suppose $\emptyset \notin S$. Let $A = \{i \in [n] : i \in S\}$, $B = [n] \setminus A$, $|A| = a$, and $|B| = b$. Let $\mathscr{M}$ be the set of all good Q_3s containing $\emptyset$. If $M \in \mathscr{M}$, then the two vertices of $M \in S$ have the structure of Type I, II, or III as in Figure 3.4.

ixy

j i i uv vx

Type I Type II Type III

Figure 3.4: The three structures of vertices in S for $M \in \mathscr{M}$ where $\emptyset \notin S$.

Define a graph G_s by $V(G_s) = B$ and $E(G_s) = \{uv : uv \text{ and } vx \text{ are the vertices in } S \text{ of some Type III } M \in \mathscr{M} \text{ with flip bits } u, v, x\}$. For such an M, ux cannot be in S so the number of Type III Q_3s in $\mathscr{M}$ is at most the number of induced copies of $K_{1,2}$ in G_s.

If L is a Type I Q_3 in $\mathscr{M}$ with flip bits i, j, x and with i, j the vertices of L in S, then $i, j \in A$ and $x \in B$. So the number of Type Is in $\mathscr{M}$ is at most $b\binom{a}{2}$. If L is a Type II Q_3 in $\mathscr{M}$ where i, ixy are the vertices of L in S, then $i \in A$ and $x, y \in B$, but $xy \notin E(G_s)$. So, if $e = |E(G_s)|$, then the number of Type II Q_3s in $\mathscr{M}$ is at most $a\left[\binom{b}{2} - e\right]$. By Lemma 3.5.1, we have that the number of Type III Q_3s in $\mathscr{M}$ is at most $\min\left\{b\binom{\frac{b}{2}}{2}, \frac{e}{2}(b-2)\right\}$.

One good candidate to maximize $|\mathscr{M}|$ is for $\mathscr{M}$ to have no Type IIIs (i.e. $|E(G_s)| = 0$), which would give $|\mathscr{M}| = b\binom{a}{2} + a\binom{b}{2}$. Another good candidate would be to have G_s be $K_{\frac{b}{2},\frac{b}{2}}$, so as to maximize the number of Type IIIs. This would mean that $|\mathscr{M}| = b\binom{a}{2} + a\left[\binom{b}{2} - \frac{b^2}{4}\right] + b\binom{\frac{b}{2}}{2} = b\binom{a}{2} + (2a+b)\binom{\frac{b}{2}}{2}$.

If $|E| = e$, then we have

$$|\mathscr{M}| \leq b\binom{a}{2} + a\left[\binom{b}{2} - e\right] + \min\left\{b\binom{\frac{b}{2}}{2}, \frac{e}{2}(b-2)\right\}. \qquad (\star)$$

If $e \geq \frac{b^2}{4}$ then $\min\{b\binom{\frac{b}{2}}{2}, \frac{e}{2}(b-2)\} = b\binom{\frac{b}{2}}{2}$, so the right-hand side of inequality ($\star$) is a decreasing function of e. Hence to maximize $|\mathscr{M}|$ we can assume $e \leq \frac{b^2}{4}$.

Case 1: If $\frac{b-2}{2} \leq a$, then

$$\begin{aligned}
|M| &\leq b\binom{a}{2} + a\left[\binom{b}{2} - e\right] + \frac{e}{2}(b-2) \\
&\leq b\binom{a}{2} + a\left[\binom{b}{2} - e\right] + ea \\
&= b\binom{a}{2} + a\binom{b}{2}
\end{aligned}$$

which is the size of $\mathscr{M}$ in the first good candidate above.

Case 2: If $\frac{b-2}{2} > a$, then

$$\begin{aligned}
|M| &\leq b\binom{a}{2} + a\left[\binom{b}{2} - e\right] + \frac{e}{2}(b-2) \\
&= b\binom{a}{2} + a\binom{b}{2} + e\left(\frac{b-2}{2} - a\right) \\
&\leq b\binom{a}{2} + a\binom{b}{2} + \frac{b^2}{4}\left(\frac{b-2}{2} - a\right) \\
&= b\binom{a}{2} + a\binom{b}{2} + b\binom{\frac{b}{2}}{2} - \frac{b^2}{4}a \\
&= b\binom{a}{2} + a\left(\binom{b}{2} - \frac{b^2}{4}\right) + b\binom{\frac{b}{2}}{2} \\
&= b\binom{a}{2} + 2a\binom{\frac{b}{2}}{2} + b\binom{\frac{b}{2}}{2} \\
&= b\binom{a}{2} + (2a+b)\binom{\frac{b}{2}}{2}
\end{aligned}$$

which is the size of $\mathscr{M}$ in the second good candidate above.

This expression can be rewritten as

$$\frac{b}{2}\binom{a}{2} + \frac{b}{2}\binom{a}{2} + a\binom{\frac{b}{2}}{2} + \frac{b}{2}\binom{\frac{b}{2}}{2} + \frac{b}{2}\binom{\frac{b}{2}}{2}$$

which is equal to

$$x\binom{y}{2} + x\binom{z}{2} + y\binom{x}{2} + y\binom{z}{2} + z\binom{x}{2} + z\binom{y}{2} \qquad (\star\star)$$

when $x = z = \frac{b}{2}$ and $y = a$. The expression in ($\star\star$) is the number of induced $K_{1,2}$s in a complete tripartite graph with part sizes x, y, and z. We know that $K_{\frac{n}{2},\frac{n}{2}}$ is the graph with n vertices which has the maximum number of induced $K_{1,2}$s, so ($\star\star$) attains its maximum value when $x = z = \frac{n}{2}$ and $y = 0$, so $b = n$ and $a = 0$. The size of $\mathscr{M}$ for the first candidate $a\binom{b}{2} + b\binom{a}{2}$ is the value of ($\star\star$) when $x = a$, $y = b$, and $z = 0$, so it attains its maximum value when $a = b = \frac{n}{2}$ and both good candidates have size

$$2 \cdot \frac{n}{2}\binom{\frac{n}{2}}{2} = \frac{n^2(n-2)}{8} = \frac{3}{4}\binom{n}{3}\frac{n}{n-1}$$

and $|\mathscr{M}|$ cannot be bigger. Hence

$$\frac{3}{4} \leq \lambda(W_2, 3) \leq \lambda_{\text{local}}(W_2, 3) \leq \frac{3}{4}.$$

■

We remark that in the construction we have with density $\frac{3}{4}$, of the vertices not in S, $\frac{2}{3}$ of them are in good Q_3's only of the type of the first good candidate (those vertices which have an odd sum in precisely one of A or B) and $\frac{1}{3}$ are in good Q_3's only of the type of the second candidate (those vertices with an odd sum in both A and B). The local density at all of these vertices is $\frac{3}{4}$.

3.6 Configurations in Q_4

In [26] we initiated the investigation of d-cube-density and considered two specific configurations, one of which was in Q_4. Using a kind of *blow–up*, we showed that $\lambda(H, d) \geq \frac{d!}{d^d}$ for every configuration H in Q_d. We defined a perfect $2d$-cycle C_{2d} in Q_d to be a cycle with d pairs of vertices each Hamming distance d apart. We showed that $\lambda(C_8, 4) = \frac{4!}{4^4}$, achieving the smallest possible value for any configuration in Q_4.

Theorem 3.6.1. *If Y is the configuration $\{0000, 1100, 0011, 1111\}$ in Q_4 (see Figure 3.5), then $\lambda(Y) = \dfrac{3}{8}$.*

Proof. Suppose $\emptyset \in S$ and let $\mathscr{M}$ be the set of good Q_4s containing $\emptyset$. We construct a graph G_s with $V(G_s) = [n]$ and $E(G_s) = \{uv : \emptyset, uv, xy, uvxy$ are the vertices in S of some $M \in \mathscr{M}\}$. If uv and xy are in $M \in \mathscr{M}$, then neither ux, uy, vx, nor vy can be in $E(G_s)$, so $|\mathscr{M}|$ is less than or equal to the number of induced copies of $2K_2$ in G_s. That means $\lambda_{\text{local(in)}}(Y) \leq i(2K_2) = \dfrac{3}{8}$.

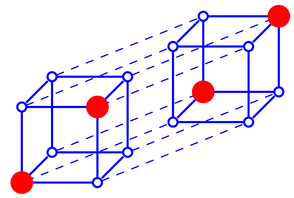


Figure 3.5: The configuration Y.

Now suppose $\emptyset \notin S$. Let $A = \{i \in [n] : i \in S\}$, $B = [n] \setminus A$, $|A| = a$, and $|B| = b$. Let $\mathscr{M}$ be the set of all good Q_4s containing $\emptyset$. If $M \in \mathscr{M}$, then the four vertices of M in S have the structure of Type I or Type II as in Figure 3.6.

iux jux uv vx xy yu

i j

Type I Type II

Figure 3.6: The two structures of vertices in S for $M \in \mathscr{M}$ where $\emptyset \notin S$.

Define a graph G_s by $V(G_s) = B$ and $E(G_s) = \{uv : uv, vx, xy, yu$ are the vertices in S of some Type II $M \in \mathscr{M}\}$. For such an M, neither ux, nor vy can be in S, so the number of Type II Q_4s in $\mathscr{M}$ is at most the number of induced copies of $K_{2,2}$ in G_s.

Lemma 3.6.2. *Let G be a graph with n vertices where n is even. If $|E(G)| = e$, then G has at most* $\min\left\{\binom{\frac{n}{2}}{2}^2, \frac{e}{4}\frac{(n-2)^2}{4}\right\}$ *induced copies of $K_{2,2}$.*

Proof. That it has at most $\binom{\frac{n}{2}}{2}^2$ copies of $K_{2,2}$ is proved in [12] and [9] (the optimizing graph is $K_{\frac{n}{2},\frac{n}{2}}$). If $uv \in E(G)$, define an auxiliary graph F with $V(F) = V(G) \setminus \{u, v\}$ and $E(F) = \{xy : \{u, v, x, y\}$ induces $K_{2,2}\}$. The graph F is triangle-free since if $\{u, v, x, y\}$ and $\{u, v, x, z\}$ both induce $K_{2,2}$, then either $\{uy, uz\} \subseteq E(G)$ or $\{vy, vz\} \subseteq E(G)$. In either case, $\{u, v, y, z\}$ induces $K_{1,3}$. Since F is triangle free, uv is in at most $\frac{(n-2)^2}{4}$ induced $K_{2,2}$s. Finally, summing over all edges uv counts each $K_{2,2}$ four times. ∎

If L is a good Type I in $\mathscr{M}$ where i, j, iux, and jux are the vertices of L in S, then $i, j \in A$, $u, x \in B$, but $ux \notin E(G_s)$. If $|E(G_s)| = e$, then the number of Type I Q_4s in $\mathscr{M}$ is at most $\left[\binom{b}{2} - e\right]\binom{a}{2}$ and of Type II, by Lemma 3.6.2, is at most

$\min\left\{\binom{\frac{b}{2}}{2}^2, \frac{e}{16}(b-2)^2\right\}$. If a and b are fixed, then one good candidate to maximize $|\mathscr{M}|$ is for $\mathscr{M}$ to have no Type II Q_4s. Then G_s has no edges and $|M| = \binom{a}{2}\binom{b}{2}$. Another good candidate is when $\mathscr{M}$ has the maximum possible number of Type II Q_4s. Then G_s is $K_{\frac{b}{2},\frac{b}{2}}$ (assuming b is even) and $|\mathscr{M}| = \left[\binom{b}{2} - \frac{b^2}{4}\right]\binom{a}{2} + \binom{\frac{b}{2}}{2}^2 = \frac{b(b-2)}{4}\binom{a}{2} + \binom{\frac{b}{2}}{2}^2$.

If $e = E(G_s)$, we have $|\mathscr{M}| \leq \left[\binom{b}{2} - e\right]\binom{a}{2} + \min\left\{\binom{\frac{b}{2}}{2}^2, \frac{e}{4}\frac{(b-2)^2}{4}\right\}$.

Case 1: If $e \geq \frac{b^2}{4}$, then

$$\begin{aligned} |\mathscr{M}| &\leq \left[\binom{b}{2} - \frac{b^2}{4}\right]\binom{a}{2} + \binom{\frac{b}{2}}{2}^2 \\ &= \frac{b(b-2)}{4}\binom{a}{2} + \binom{\frac{b}{2}}{2}^2 \end{aligned}$$

which is the size of $\mathscr{M}$ in the second good candidate above.

Case 2: If $e < \frac{b^2}{4}$, then

$$\begin{aligned} |\mathscr{M}| &\leq \left[\binom{b}{2} - e\right]\binom{a}{2} + \frac{e}{16}(b-2)^2 \\ &= \binom{a}{2}\binom{b}{2} + e\left[\frac{1}{16}(b-2)^2 - \binom{a}{2}\right]. \end{aligned}$$

If $\frac{1}{16}(b-2)^2 \leq \binom{a}{2}$, then $|\mathscr{M}| \leq \binom{a}{2}\binom{b}{2}$ which is the size of $\mathscr{M}$ in the first good candidate above.

If $\frac{1}{16}(b-2)^2 > \binom{a}{2}$, then

$$\begin{aligned} |\mathscr{M}| &< \binom{a}{2}\binom{b}{2} + \frac{b^2}{4}\left[\frac{1}{16}(b-2)^2 - \binom{a}{2}\right] \\ &= \binom{a}{2}\left[\binom{b}{2} - \frac{b^2}{4}\right] + \left(\frac{b(b-2)}{8}\right)^2 \\ &= \frac{b(b-2)}{4}\binom{a}{2} + \binom{\frac{b}{2}}{2}^2 \qquad (\star) \end{aligned}$$

the same upper bound as in Case 1.

Clearly the maximum value of $\binom{a}{2}\binom{b}{2}$ is $\binom{\frac{n}{2}}{2}\binom{\frac{n}{2}}{2} = \frac{n^2(n-2)^2}{64} = \frac{3}{8}\binom{n}{4}\frac{n(n-2)}{(n-1)(n-3)}$.

Lemma 3.6.3. *If $x, y,$ and z are nonnegative real numbers such that $x + y + z = n$,* *then the maximum value of*

$$\binom{x}{2}\binom{y}{2}+\binom{x}{2}\binom{z}{2}+\binom{y}{2}\binom{z}{2} \tag{$*$}$$

is $\binom{\frac{n}{2}}{2}^2$.

Proof. To simplify notation, let $n \equiv 0 \mod 6$.

This function counts the number of induced copies of $K_{2,2}$ in a complete tripartite graph with parts X, Y, and Z with part sizes x, y, and z, respectively, subject to the constraint $x + y + z = n$. In [], it was shown that for all $n \geq 4$ the maximum number of induced copies of $K_{2,2}$ in any graph is $\binom{\frac{n}{2}}{2}^2$. ∎

If $x = a$ and $y = z = \frac{b}{2}$, then $*$ reduces to $\star$, so the maximum of $\star$ occurs when $a = 0$ and $b = n$ and is equal to $\frac{3}{8}\binom{n}{4}\frac{n(n-2)}{(n-1)(n-3)}$. Hence, $\frac{3}{8}$ is an upper bound for $\lambda_{\text{local(out)}}(Y)$ and $\lambda_{\text{local(in)}}(Y)$, so $\frac{3}{8} \leq \lambda(Y, 4) \leq \lambda_{\text{local}}(Y, 4) \leq \frac{3}{8}$.

∎

Theorem 3.6.4. *If H is the configuration $\{0000, 1100, 1010, 0110\}$ in Q_4 (see Figure 3.7), then $\lambda(H, 4) = \frac{1}{2}$.*

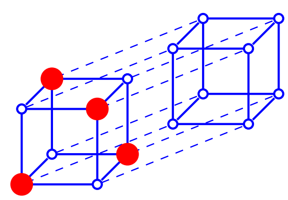


Figure 3.7: The configuration H for Theorem 3.6.4.

Proof. Explain the construction given by $A \cup B$ taking 00 mod 2 gives density $\frac{1}{2}$ (because 3-1 and 1-3 always give "good" Q_4s).

Suppose $\emptyset \in S$ and let $\mathscr{M}$ be the set of good Q_4s containing $\emptyset$. We define a graph G_S with $V(G_S) = [n]$ and $E(G_S) = \{xy : \emptyset, xy, yz, xz$ are the vertices in S of some $M \in \mathscr{M}\}$. If x, y, z, w are the coordinates of a good Q_4 where $\emptyset, xy, yz, xz$ are the vertices in S, then wx, wy, wz are not in $E(G_S)$, so $\{w, x, y, z\}$ induces K_3 plus an isolated vertex in G_S. Since this is the complement of $K_{1,3}$, $\lambda_{\text{local(in)}}(H, 4) = i(K_{1,3}) = \frac{1}{2}$.

Now suppose $\emptyset \notin S$. Let $A = \{i \in [n] : i \in S\}$, $B = [n] \setminus A$, $|A| = a$, $|B| = b$. Let $\mathscr{M}$ be the set of all good Q_4s containing $\emptyset$. If $M \in \mathscr{M}$ then the four vertices of M in S

have the structure of Type I, Type II, or Type III in Figure 3.8 (where $i, j, k \in A$ and $w, x, y, z \in B$).

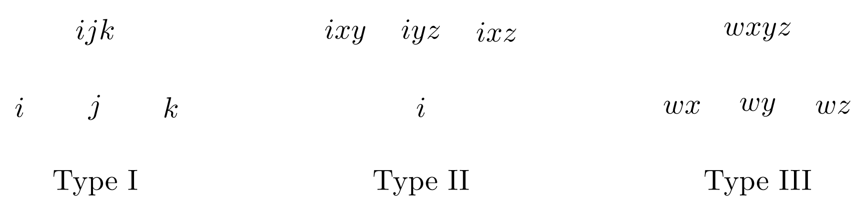


Figure 3.8: The three structures of vertices in S for $M \in \mathscr{M}$ where $\emptyset \notin S$.

Define a graph G by $V(G) = A \cup B$ and $E(G) = \{ix : i \in A \text{ and } x \in B\} \cup \{wx : wx, wy, wz, wxyz$ are the vertices in S of a Type III $M \in \mathscr{M}$ for some $y, z \in B\}$. If M is a Type I Q_4 with coordinates i, j, k, x, then $\{i, j, k, x\}$ induces $K_{1,3}$ in G. If M is a Type III Q_4 with vertices $wx, wy, wz, wxyz$ in S, then $\{w, x, y, z\}$ induces $K_{1,3}$ in G, since xy, yz and xz are not edges in G. That means $\{i, x, y, z\}$ induces $K_{1,3}$ in G since ix, iy, iz are all edges. It also means that the number of Type II Q_4s in $\mathscr{M}$ is at most the number of $K_{1,3}$s in G with one vertex in A and three vertices in B, since if i, x, y, z are the coordinates of a Type II M, then xy, yz, and xz are all non-edges. Thus $|\mathscr{M}|$ is at most the number of $K_{1,3}$s in G which have precisely 3,1, or 0 vertices in A, so is certainly at most the maximum number of $K_{1,3}$s in a graph with n vertices. Hence $\lambda_{\text{local(out)}}(H, 4) \le i(K_{1,3}) = \frac{1}{2}$, and $\frac{1}{2}] \le \lambda(H, 4) \le \lambda_{\text{local}}(H, 4) \le \frac{1}{2}$. ■

We remark that since the only optimizing host graph to maximize the number of induced $K_{1,3}$ subgraphs is complete bipartite, the graph G defined above can only be optimal if either there are no Type III $M \in \mathscr{M}$ (so both A and B are independent sets), or $A = \emptyset$, each $M \in \mathscr{M}$ is Type III, and B induces a complete bipartite graph (with parts not quite equal in size).

3.7 An Infinite Family

Theorem 3.3.1 can be generalized to apply to an infinite family of configurations containing V_3, W_6, and W_{13}. Let d and i be positive integers with $1 \le i < d$. We define the configuration $H(d, i)$ in Q_d by

$$H(d, i) = \left\{ (v_1, v_2, \ldots, v_d) \in V(Q_d) \;\middle|\; \sum_{j=1}^{i} v_j \text{ is even} \right\}.$$

Theorem 3.7.1. $\lambda(H(d,i)) = \binom{d}{i}\frac{i^i(d-i)^{d-i}}{d^d}$.

Proof. Each vertex in $H(d,i)$ has $d-i$ neighbors in $H(d,i)$ (change any one of the last $d-i$ coordinates). Since $H(d,i)$ is self-complementary in Q_d $\left(\sum_{j=1}^{i} v_j \text{ is odd}\right)$, $\lambda_{\text{local}}(H(d,i)) = \lambda_{\text{local(in)}}(H(d,i)) = \lambda_{\text{local(out)}}(H(d,i))$.

If $n \geq d$ and $v \in S \subseteq V(Q_n)$ and R is a sub-d-cube of Q_n containing v, then R can be good only if precisely $d-i$ neighbors of v in R are in S. If x is the fraction of neighbors of v in $V(Q_n)$ which are in S, then the fraction of sub-d-cubes of Q_n containing v which have precisely $d-i$ neighbors in S is $f(x) = \binom{d}{i}x^{d-i}(1-x)^i$. By simple calculus, $f(x)$ is maximized on $[0,1]$ when $x = \frac{d-i}{d}$, so $\lambda_{\text{local(in)}}(H(d,i)) \leq \binom{d}{i}\frac{(d-i)^{d-i}i^i}{d^d}$.

To show this upper bound is a lower bound as well, let $S = \big\{(v_1, v_2, \ldots, v_n) : \sum_{j=1}^{m} v_j$ is even, where $m = \lfloor \frac{in}{d} \rfloor\big\}$. Then any sub-$d$-cube of Q_n with precisely i flip-bits in $[1,m]$ is good, and this is a fraction

$$\binom{d}{i}\frac{m^i(n-m)^{d-i}}{n^d} = \binom{d}{i}\left(\frac{\lfloor \frac{in}{d} \rfloor}{n}\right)^i \left(\frac{\left\lceil \frac{(d-i)n}{d} \right\rceil}{n}\right)^{d-i}$$

of all sub-d-cubes, and the limit as n goes to infinity is $\binom{d}{i}\frac{i^i(d-i)^{d-i}}{d^d}$. ■

Note that the configuration W_6 in Q_3 is $H(3,1)$, W_{13} is $H(3,2)$, and V_3 in Q_2 is $H(2,1)$. Further note that $\lim_{d\to\infty} \lambda(H(d,i)) = \frac{i^i}{i!e^i}$. In particular, when $i = 1$, $\lim_{d\to\infty}(H(d,1)) = \frac{1}{e}$. ($H(d,1)$ is a copy of Q_{d-1} in Q_d)

3.8 Layered Configurations

Recall that we say a configuration H in Q_d is *layered* if it is an exact copy of a configuration K in Q_d such that $v \in K$ if and only if $\text{wt}(v) \in W$ for some set W of nonnegative integers. For example, $H = \{1001, 1110, 0010, 0100, 0111\}$ is layered because there is an automorphism of Q_4 (interchange 0 and 1 in the 2^{nd} and 3^{rd} coordinates) which maps H onto $K = \{1111, 1000, 0100, 0010, 0001\}$and $K = \{v \in Q_4 : \text{wt}(v) = 1 \text{ or } 4\}$. We call K a *canonical layered configuration.* The configurations $W_1, W_3, W_7, W_8, W_{12}$, and W_{14} (and their complements) are layered configurations in Q_3. One can get a good lower bound construction for any layered configuration in Q_d by using an appropriately layered set S in Q_n. If W_H is the set of weights for the vertices in a canonically layered configuration H in Q_d, and if we choose a layered set S in Q_n where W_S is the set of weights, and if $W_S \cap \{0, 1, \ldots, d\} = W_H$, then every sub-$d$-cube containing $\emptyset$ is "good", so $\lambda_{\text{local}}(H,d) = 1$. That means our usual procedure of using $\lambda_{\text{local}}(H,d)$ to get an upper bound for d-cube density cannot work and that is why we have not been able to obtain good upper bounds by hand for any layered configurations other than the three with d-cube density equal to 1 ($\emptyset$, $V(Q_d)$, and all even weight vertices in Q_d).

For example, if we represent the configuration W_8 by $H = \{110, 101, 011\}$ we define S by $S = \{v \in V_n : \text{wt}(v) \equiv 2 \mod 3\}$. Any sub-3-cube whose smallest weight vertex has weight congruent to 0 or 1 mod 3 is "good", showing that $\lambda(W_8, 3) \geq \frac{2}{3}$. Baber's flag algebra upper bound is .66666666675 so undoubtedly $\lambda(W_8, 3) = \frac{2}{3}$, but we have not proved it.

For each positive integer n let F^n denote the set of binary n-tuples. If $u \in F^n$ we let u^R denote the n-tuple obtained by reversing the order of the digits in u. If $k \leq n, u \in F^k$, and $v \in F^n$, we let $f_n(u, v)$ denote the fraction of the $n - k + 1$ strings of k consecutive digits of v which are equal to u or u^R and we define $f(u)$ by

$$f(u) = \lim_{n \to \infty} \max_{v \in F^n} f(u, v).$$

So $f(u)$ is the limit as n goes to infinity of the maximum fraction of strings of k consecutive digits in any n-tuple wihch are equal to u or u^R.

A beginning segment of $u = (u_1, u_2, \ldots, u_k)$ is the t-tuple $(u_1, u_2, \ldots, u_t)$ for some $t \in [1, k-1]$ and an ending segment is the m-tuple $(u_{k-m+1}, \ldots, u_k)$ for some $m \in [1, k-1]$. We let $s(u)$ be the maximum length of a beginning segment of u which is equal to an ending segment, and we let $p(u) = p = k - s(u)$. We construct $v = (v_1, v_2, \ldots, v_n) \in F^n$ by repeating the p-tuple $(u_1, u_2, \ldots, u_p)$, that is $v_i = u_j$ if $j = i \mod p$. For example, if $u = 1101001101$ then $k = 10, s(u) = 4$ (1101), $p = 6$, and we form v by repeating the string 110100. Each k consecutive digits of v whose first digit is in a position congruent to $1 \mod p$ is a copy of u, so $f(u) \geq \frac{1}{k - s(u)}$.

There is another way to get overlapping copies of u or u^R in v. We say $x \in F^n$ is a *palindrome* if $x = x^R$. If $u \in F^k$, let $b(u)$ and $e(u)$ be the lengths of the largest beginning segment and ending segment of u which are palindromes, so $1 \leq b(u)$ and $e(u) \leq k - 1$. We construct $v \in F^n$ for large n as follows. Take a copy of u for the first k digits. Then digits $k - e_u + l$ through $2k - e(u)$ are a copy of u^R, overlapping the initial copy of u in $e(u)$ digits, and these digits are a palindrome. Then digits $2k - e(u) - b(u) + 1$ through $3k - e(u) - b(u)$ are a copy of u, overlapping the previous u^R in $b(u)$ digits, and these are a palindrome. Then we repeat this process. Since the second copy of u begins in digit number $2k - e(u) - b(u) + 1$, we are generating an n-tuple with period $2k - e(u) - b(u)$, perhaps with something extra at the end. Hence $f(u) \geq \frac{2}{2k - e(u) - b(u)}$.

For example, if $u = 110101101101$ then $k = 12, b(u) = 7.e(u) = 9$, and the $2k - b(u) - e(u) = 8$ repeating digits are 11010110. A copy of u begins in digits 1,9,17,...

and a copy of u^R begins in digits 4,12,20,....

Since there is no other way to get overlap in v between two successive copies of u and/or u^R, we have proved the following:

Proposition 3.8.1. *If $u \in F^k$ let $s(u)$ be the longest beginning segment of u which is equal to an ending segment, and let $b(u)$ and $e(u)$ be the lengths of the largest beginning segment and ending segment, respectively, of u which are palindromes. Then*

$$f(u) = \max\left\{\frac{1}{k - s(u)}, \frac{2}{2k - b(u) - e(u)}\right\}.$$

If $u = (a_1, a_2, \ldots, a_k) \in F^k$ then clearly $b(u) + e(u) \geq 2$ with equality if and only if $u = 100\ldots001$ or the complement. It is easy to see that $b(u) + e(u) \leq 2k - 2$ with equality if and only if u is all 1's or all 0's or k is even and u is alternating 0's and 1's.

It is not hard to check that if $k \geq 4$ then for each j in $[2, 2k-2]$, except $j = 3$, there exists $u \in F^k$ such that $b(u) + e(u) = j$. Hence $f(u)$ can equal $\frac{2}{2k-j}$ for any integer j in $[2, 2k-2]$ except $j = 3$.

If K is a canonical layered configuration in Q_d then we define its weight vector $w_K = (a_0, a_1, \ldots, a_d) \in F^{d+1}$ by $a_i = 1$ if and only if the vertices of weight i are in K. If K is an exact copy of H then we define w_H to be equal to w_K. Given a vector $w_H \in F^{d+1}$, we can choose a layered configuration in Q_n just as we chose $v \in F^n$ to maximize $f_n(u, v)$. Hence we have the following

Proposition 3.8.2. *Let H be a layered configuration in Q_d. Then*

$$\lambda(H, d) \geq f(w_H) = \max\left\{\frac{1}{d + 1 - s(w_H)}, \frac{2}{2(d+1) - b(u_H) - e(w_H)}\right\}.$$

Proposition 3.8.3. *For each $d \geq 2$, there is a layered configuration H in Q_d with $\lambda(H, d) \geq \frac{2}{3}$.*

For $d = 7, 8, 9$ the layered configurations with weight vectors $10010010, 100100100$, and 0100100100 respectively have density at least $\frac{2}{3}$. But for 1001001001 we have $s = b = e = 7$ we can only say the density is at least $\frac{1}{10-7} = \frac{2}{20-7-7} = \frac{1}{3}$.

Example 4. Let $K_5 = \{v = (a_1, a_2, a_3, a_4, a_5) : a_1 + a_2 = a_3 + a_4 + a_5 = 0 \mod 2\}$. If A, B is a partition of the vertices $[n]$ of Q_n and S is the set of all vertices such that the number of 1s in A and in B is even, then any Q_5 with 2 or 3 vertices in A has an exact copy of K_5. Hence $\lambda(K_5, 5) \geq \frac{\binom{5}{2}+\binom{5}{3}}{2^5} = \frac{5}{8}$.

Conjecture 3.8.4. If H is a configuration in Q_d such that $\frac{5}{8} < \lambda(H, d) < 1$ then either $d = 3$ and $H = W_2$ ($\lambda(W_2, 3) = \frac{3}{4}$) or H is layered with a period 3 (possibly with remainder) weight vector.

There are 6 different layered configurations (counting each complementary pair once) in Q_3 and 10 of them in Q_4. For 13 of these 16 configurations the flag algebra upper bound that we have is very close to the lower bound provided by the layered construction. The exceptions are one vertex in Q_3 ($d(1000) = \frac{1}{2}$, flag algebra bound .6100), one vertex in Q_4 ($d(10000) = \frac{2}{5}$, flag algebra bound .6025), and all even weight vertices except one in Q_4 ($d(10100) = \frac{2}{5}$, flag algebra bound .6123). We have no idea if the layered construction is optimal for these configurations, while it probably is for the other 13.

Finally, we have remarked that if H is a layered configuration in Q_d then $\lambda_{\text{local}}(H, d) = 1$. We suspect the converse is true.

Conjecture 3.8.5. If H is a configuration in Q_d such that $\lambda_{\text{local}}(H, d) = 1$ then H is layered.

Bibliography

[1] N. Alon, D. Hefetz, M. Krivelevich, and M. Tyomkyn, *Edge-statistics on large graphs*, Combinatorics, Probability and Computing **29** (2020), no. 2, 163–189.

[2] N. Alon, A. Krech, and T. Szabó, *Turán's theorem in the hypercube*, SIAM Journal on Discrete Mathematics **21** (2007), 66–72.

[3] N. Alon, R. Radoičić, B. Sudakov, and J. Vondrák, *A Ramsey-type result for the hypercube*, Journal of Graph Theory **23** (2006), 196–208.

[4] M. Axenovich, J. Goldwasser, R. Hansen, B. Lidický, R. R. Martin, D. Offner, J. Talbot, and M. Young, *Polychromatic colorings of complete graphs with respect to 1-, 2-factors and hamiltonian cycles*, Journal of Graph Theory **87** (2018), 660–671.

[5] M. Axenovich and R. Martin, *A note on short cycles in a hupercube*, Discrete Mathematics **306** (2006), 2212–2218.

[6] R. Baber, private communication, 2014.

[7] A. Bialostocki, *Some ramsey type results regarding the graph of the n-cube*, Ars Combinatorica (1983).

[8] B. Bollobás, C. Nara, and S. Tachibana, *The maximal number of induced complete bipartite graphs*, Discrete Mathematics **62** (1986), no. 3, 271–275.

[9] B. Bollobas, C. Nara, and S. Tachibana, *The maximal number of induced complete bipartite graphs*, Discrete Mathematics **62** (1986), no. 3, 271–275.

[10] B. Bollobás, D. Pritchard, T. Rothvoss, and A. Scott, *Cover-decomposition and polychromatic numbers*, SIAM Journal on Discrete Mathematics **27** (2013), 240–256.

[11] J. A. Bondy and P. Erdős, *Ramsey numbers for cycles in graphs*, Journal of Combinatorial Theory (B) **14** (1973), 46–54.

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

[13] G. Chartrand and S. Schuster, *On the existence of specified cycles in complementary graphs*, Bulletin of the American Mathematical Society **77** (1971), 995–998.

[14] Z. Chen, private communication, 2016.

[15] I. Choi, B. Lickický, and F. Pfender, *Inducibility of directed paths*, Discrete Mathematics **343** (2020).

[16] F. R. K. Chung, *Subgraphs of the hypercube containing no small even cycles*, Journal of Graph Theory **16** (1992), no. 3, 273–286.

[17] M. Conder, *Hexagon-free subgraphs of hypercubes*, Journal of Graph Theory **17** (1993), 477–479.

[18] D. Conlon, *An extremal theorem in the hypercube*, Electronic Journal of Combinatorics **17** (2010), R111.

[19] P. Erdős and A. Gyárfás, *A variant of the classical ramsey problem*, Combinatorica **17** (1997), 459–467.

[20] C. Even-Zohar and N. Linial, *A note on the inducibility of 4-vertex graphs*, Graphs and Combinatorics **31** (2015), 1367–1380.

[21] G. Exoo, *Dense packings of induced subgraphs*, Ars Combinatoria **22** (1986), 5–10.

[22] V. Falgas-Ravry and E. R. Vaughan, *Turán H-densities for 3-graphs*, Electronic Journal of Combinatorics **19** (2012), no. 3, 1–26.

[23] R. J. Faudree and R. H. Schelp, *All ramsey numbers for cycles in graphs*, Discrete Mathematics **8** (1974), 313–329.

[24] Z. Füredi and L. Ozkahya, *On 14-cycle-free subgraphs of the hypercube*, Combinatorics, Probability, and Computing **18** (2009), 725–729.

[25] W. Goddard and M. Henning, *Thoroughly dispersed colorings*, Journal of Graph Theory **88** (2018), 174–191.

[26] J. Goldwasser and R. Hansen, *Maximum density of vertex-induced perfect cycles and paths in the hypercube*, Discrete Mathematics **344** (2021).

[27] ______, *Polychromatic colorings of 1-regular and 2-regular subgraphs of complete graphs*, Discrete Mathematics **345** (2022).

[28] J. Goldwasser, B. Lidický, R. R. Martin, D. Offner, J. Talbot, and M. Young, *Polychromatic colorings on the hypercube*, Journal of Combinatorics **9** (2018), 631–657.

[29] J. Goldwasser and J. Talbot, *Vertex Ramsey problems in the hypercube*, SIAM Journal on Discrete Mathematics **26** (2012), 838–853.

[30] A. W. Goodman, *Triangles in a complete chromatic graph with three colors*, Discrete Mathematics **57** (1985), no. 3, 225–235.

[31] A. Grzesik, *On the maximum number of five-cycles in a triangle-free graph*, Journal of Combinatorial Theory B **102** (2012), 1061–1066.

[32] H. Hatami, J. Hladký, D. Král, S. Norine, and A. Razborov, *On the number of pentagons in triangle-free graphs*, Journal of Combinatorial Theory A **120** (2013), 722–732.

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

[34] B. Jackson, *Long cycles in bipartite graphs*, Journal of Combinatorial Theory Series B **38** (1985), 118–131.

[35] J. R. Johnson and J. Talbot, *Vertex turan problems in the hypercube*, Journal of Combinatorial Theory Series A **117** (2010), 454–465.

[36] K. A. Johnson and R. Entringer, *Largest induced subgraphs of the n-cube that contain no 4-cycles*, Journal of Combinatorial Theory Series B **46** (1989), no. 3, 346–355.

[37] E. A. Kostochka, *Piercing the edges of the n-dimensional unit cube*, Diskret. Analiz Vyp **28** (1976), no. 55-64, 223.

[38] I. Leader and E. Long, *Long geodesics in subgraphs of the cube*, Discrete Mathematics **326** (2014), 29–33.

[39] D. Offner, *Polychromatic colorings of subcubes of the hypercube*, SIAM Journal on Discrete Mathematics **22** (2008), 450–454.

[40] N. Pippenger and M. C. Golumbic, *The inducibility of graphs*, Journal of Combinatorial Theory Series B **19** (1975), 189–203.

[41] M. S. Rahman, M. Kaykobad, and M. T. Kaykobad, *Bipartite graphs, hamiltonicity and Z graphs*, Electronic Notes in Discrete Mahtematics **44** (2013), 307–312.

[42] A. Thomason and P. Wagner, *Bounding the size of square-free subgraphs of the hypercube*, Journal Discrete Mathematics **309** (2009), 1730–1735.

[43] E. Vaughan, *Flagmatic: a tool for researchers in extremal graph theory (version 2.0)*, 2013, http://flagmatic.org/graph.html.