Extremal Graph Theoretical Questions for q-ary Vectors

Thesis

Koppány István Encz

MSc in Applied Mathematics

Supervisor:

Balázs Patkós, senior researcher
Alfréd Rényi Institute of Mathematics

Eötvös Loránd University, Budapest, Hungary
Department of Computer Science

2023

Acknowledgements

First and foremost, I would like to express my gratitude towards my supervisor, Balázs Patkós, whose priceless insights proved to be vital for understanding the mathematical background of my research topic. He never missed an opportunity to shed light on some remarkable connections with other fields of mathematics; and he always connected the dots for me when I was in desperate need of professional support. I firmly believe that this thesis could not have been successfully completed, had it not been for his thorough explanations and endless patience.

I also owe special thanks to Kristóf Bérczi, who has always been at my disposal to solve administrative problems; and as my former tutor in the Matroid Theory course, he introduced me to amazing the topic of matroids.

Contents

1 Introduction

In the last century, the broad area of graph theory became one of the most popular and prospective fields in mathematics. Many scientists devote their lives to studying graphs, the mathematical objects that formally describe our notion of networks. As networks appear in many forms in various disciplines, such as social networks or nervous systems; new questions emerge from other fields of science that can be best interpreted by using graph structures, and thus present a colourful palette of challenging problems for the researchers of graph theory.

A fundamental part of these questions fall under the umbrella of *extremal graph theory*, which focuses on maximizing certain properties of graphs subject to a set of fixed constraints. It mainly encompasses questions in the form "At most how many edges may a graph on n vertices have without containing a forbidden subgraph H?". Since the original proposal of the topic by Pál Turán ([12] and [13]), it has quickly gained popularity among mathematicians, and several results have been established over the course of a few decades. Erdős, Stone and Simonovits achieved a major breakthrough in [3] and [4], where they linked extremal numbers to another fundamental concept in graph theory, the chromatic number of graphs. Their famous result states that

$$\mathrm{ex}(n, H) = \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right) \cdot \binom{n}{2}, \tag{1}$$

whenever H is a simple, bipartite graph on n vertices. Here, $\mathrm{ex}(n, H)$ denotes the maximal number of edges an n-vertex graph may have without containing H as a subgraph, and $\chi(H)$ is the chromatic number of H: the minimal number of colors needed for coloring the vertices of H in a way that no two adjacent vertices get the same color. This formula, however, does not give a satisfying answer when $\chi(H) = 2$, in which case H is called *bipartite*. For bipartite graphs, many special cases have been sufficiently discussed, yet it remains open to come up with a general answer such as 1. A relatively fresh collection of relevant statements is listed in [5].

Since their original definition, many alternative versions of graphs (such as directed-, weighted- or hypergraphs) have emerged; and while they all incorporate the key properties of ordinary graphs, the slightly modified variants allow for completely different applications. Many of the generalisations stem from the *incidence* *matrix* of graphs; a useful tool for storing them and representing the edge-node relations, as the name suggests. The rows of the matrix pertain to the nodes of the graph, and the columns encode the edges in the following manner: the intersection of a row and a column has 1 as an entry if and only if the edge pertaining to the column contains the vertex pertaining to the row, and 0 otherwise.

Perhaps the most widely studied of the above categories is hypergraphs, where hyperedges are subsets of arbitrary size on a fixed vertex set. In terms of incidence matrices, the definition naturally translates to hypergraphs, with the $0-1$ characteristic vectors of hyperedges replacing columns of exactly two 1-entries. Similar to the case of ordinary graphs, the question of extremality [1] is of utter importance, yet there has not been a major breakthrough comparable to the result of Erdős, Stone and Simonovits. In fact, even the case of K_r^n (the complete r-uniform graph on n vertices with each r distinct element forming a hyperedge) is still open for $n > r > 2$, in spite of being in the spotlight since Turán's proposal of the problem. According to the survey of Keevash [6], many of the recent results revolve around hypergraphs of a special structure, with a small number of vertices end hyperedges. For instance, it was only recently that the Turán-number of the Fano plane (considered as a 3-graph on 7 vertices) has been determined asymptotically in [8].

Theorem 1.1 (Keevash, Sudakov). $ex(n, Fano) \sim \frac{3}{4}\binom{n}{3}$.

Another important aspect of extremal graph theory is the question of stability. Even for ordinary graphs, it is useful to know under what circumstances F-free graphs with $\mathrm{ex}(n, F)$ edges are unique. For instance, the most popular proof of determining $\mathrm{ex}(n, K_k)$ terminates with pointing out that it is realised exclusively by the complete $k-1$-partite graph with (almost) equal class sizes. A strengthened version of the above requirement is the following: if G is F-free with almost $\mathrm{ex}(n, F)$ edges, then it is structurally close to the unique F-free graph with $\mathrm{ex}(n, F)$ edges. The theorem of Erdős and Simonovits ([11]) captures the essence of this approach:

Theorem 1.2 (Erdős-Simonovits Stability Theorem). *For any $\varepsilon > 0$ there exists $\delta > 0$ such that if G is a K_t-free graph with at least $(1-\delta)ex(n, K_t)$ edges, then there is a partition of $V(G)$ as $V_1 \cup \ldots \cup V_{t-1}$ with $\sum_i |E[V_i]| \leq \varepsilon n^2$.*

[1] The extremal number is only defined for r-uniform graphs, where each hyperedge contains exactly r vertices. For an r-uniform F, $\mathrm{ex}(n, F)$ denotes the maximal size of an r-uniform F-free graph on n vertices.

Here, $\sum_i |E[V_i]|$ denotes the number of edges inside the partition classes, a common measure of resemblance to the complete $t-1$-partite graph. A variant of the stability requirement is depicted by the theorem of Andrásfai, Erdős and T. Sós ([1]):

Theorem 1.3 (Andrásfai, Erdős, T. Sós)**.** *Any triangle-free graph G on n vertices with minimum degree $\delta(G) > 2n/5$ is bipartite.*

This improved version of Mantel's theorem, which was chronologically first of its kind, uses a different indicator of similarity than Theorem 1.2: it simply counts the difference between the number of edges, since the optimal construction is the complete bipartite graph on $\lfloor \frac{n}{2} \rfloor$ and $\lceil \frac{n}{2} \rceil$ vertices, and the theorem guarantees that any C_3-free graph sufficiently close to the optimal one with respect to the degrees is also bipartite.

Ultimately, let us mention that for practically every concept of extremality, we can define coloured variants of the problem. The following example originates from Keevash, Mubayi, Sudakov and Verstraëte. In a simple graph, we call a colouring of the edges *proper* if every pair of adjacent edges have different colours. We can forbid rainbow copies of a graph H (meaning that all subgraphs isomorphic to H must have some edges of the same colour), and maximize the number of edges on n vertices subject to this property. This maximum is denoted by $\mathrm{ex}^*(n, H)$, and referred to as the *rainbow Turán number*. Keevash et. al. showed in [7] that for any non-bipartite H, $\mathrm{ex}^*(n, H) \sim \mathrm{ex}(n, H)$. Just like for the original Turán problem, attempts to solve the bipartite case have been met with insurmountable obstacles. Even the $H = C_{2k}$ case is unknown, with a current conjecture of $\mathrm{ex}^*(n, C_{2k}) = O(n^{1+1/k})$.

In their recent paper [10], Patkós, Tuza and Vizer presented an alternative version of incidence matrices. Their definition of *q-graphs* replaces the 1-entries of an incidence matrix with arbitrary positive integers between 1 and a fixed integer q. Subsequently, a q-graph H on n vertices is determined by a matrix $A \in \{0, 1, \ldots, q\}^{n \times |H|}$, where each column representing a *q-edge* has exactly two non-zero entries. It is easy to see that q-graphs indeed generalise ordinary graphs: substituting $q = 1$ to the definition leads back to incidence matrices.

The topic of this thesis is to give a detailed summary of the main questions regarding q-graphs, revise the important results in the topic, and highlight open problems that could potentially be in the focus of future research. The structure of the thesis is as follows:

In Section 2, we give formal definitions for basic concepts in graph theory; and extend them to q-ary graphs as well. A brief discussion about the relationship of ordinary and q-graphs will follow, and the section concludes with a few elementary observations about the extremal number of q-graphs. In Section 3, we turn to examine the extremal number of special graph classes, and introduce useful methods commonly used in the proofs. Section 4 will be dedicated to $\mathrm{ex}(n, F, q)$ for even values of q, and we will show that determining the q-extremal number for an even q is equivalent with finding $\mathrm{ex}(n, F, 2)$. Finally, Section 5 presents a brief highlight on the main characteristics of the thesis.

In Section 4, we list theorems that are the result of a joint work between the author of the thesis, and fellow researchers Márton Marits, Máté Weisz and Benedek Váli. The research was supported by the Hungarian Research Experience for Undergraduates 2022 program, and the results were published in the Proceedings of the 12th Japanese-Hungarian Symposium in Discrete Mathematics and Its Applications.

2 Definitions, preliminaries

2.1 Definitions

Throughout this paper, $G = (V, E)$ will denote a finite, simple, connected graph; and when not stated otherwise, the vertex set of a graph on n vertices will always be $\{1, 2, \ldots, n\} = [n]$. Let us note that most theorems and definitions can be extended, with little to no modification, to graphs with more than one connected components, or graphs with parallel edges and loops; yet for the sake of simplicity, G will almost always be assumed to have the above properties.

A common way of defining graphs is by giving their *incidence matrix*, which represents a node-edge inclusion by a row-column intersection.

Definition 2.1. *Let $G = (V, E)$ be a simple graph on n vertices, with $V = \{1, 2, \ldots, n\}$. The incidence matrix of G is the matrix $A \in \{0, 1\}^{V \times E}$ where $A_{i,e} = 1 \iff i$ is an endpoint of e.*

Apart from providing a concise way to store and represent graphs, the incidence matrix introduces linear algebraic tools in graph theory, which can be exploited to derive a wide array of theorems in domains such as *spectral graph theory.*

For every edge e, the column pertaining to e has exactly 2 non-zero entries: it has 1 as an entry in the rows that belong to the endpoints of e (Note that this property does not hold when G is not simple and e is a loop, in which case the column of e has only one non-zero element). This idea leads to the following generalisation of incidence matrices: let us fix an arbitrary integer $q \in \mathbb{N}^+$, and construct a matrix $A \in \{0, 1, \ldots, q\}^{V \times E}$ such that each column has exactly two non-zero elements. In unison with ordinary graphs, we say that a column of A represents a q-edge on the vertex set $\{1, 2, \ldots, n\}$, and the two coordinates for which the column vector is non-zero form the *support* of the q-edge. The collection of q-edges defined by the columns of A will form a *q-graph* on V.

In the following few definitions, we outline the basic concepts we will rely on in the upcoming sections. Before that, let us briefly summarise commonly accepted notations; we try to adhere to using them when possible.

Notation 2.2. *When not stated otherwise, a graph $G = (V, E)$ will always have* *a vertex set of size n. For a graph $G = (V, E)$ on n vertices, we will assume that $V = \{1, 2, \dots, n\}$.*

Notation 2.3. *For a vector x, the i-th coordinate of the vector is either x_i or $x(i)$.*

Notation 2.4. *For $n \in \mathbb{N}$, let $[n] = \{1, 2, \dots, n\}$.*

Our first definition gives an alternative name of the endpoints of an edge for q-edges. As q-edges are formally introduced through the use of vectors, it is reasonable to abstain from using the same name. Therefore, the non-zero coordinates of a vector x will be called the *support* of x.

Definition 2.5. *Let $q \in \mathbb{N}^+$ an arbitrary integer, and let $x \in \{0, 1, \dots q\}^n$ be a vector of length n. The support of x is $S_x = \{i \in [n] : x_i \neq 0\}$.*

Patkós, Tuza and Vizer first introduced their definition of q-graphs in [10]. Their definition stems from an extremal set theoretical approach, and considers vectors with integer elements as weighted characteristic vectors of sets.

Definition 2.6 (Patkós, Tuza, Vizer)**.** $\mathcal{Q}(n, r) := \{\boldsymbol{x} \in \{0, 1, \dots, q\}^n : |S_{\boldsymbol{x}}| = r\}$. *A q-graph H on n vertices is $H \subseteq \mathcal{Q}(n, 2)$. The vertex set of H is $\bigcup_{\boldsymbol{x} \in H} S_{\boldsymbol{x}}$. A q-edge of H is $\boldsymbol{x} \in H$. The size of H is the total number of q-edges in H, and will be denoted by $|H|$.*

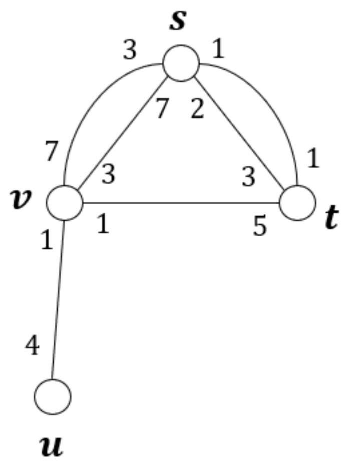
From this point of view, a q-graph is a collection of weighted incidence vectors of a 2-uniform set family, with each set having weights from $\{0, 1, \dots, q\}$. As a 2-uniform set family on the base set $[n]$ can be viewed as a graph on the vertices $\{1, 2, \dots, n\}$, we can envision a q-graph as a graph on the vertex set $[n]$, where each edge has two non-zero integer weights between 1 and q assigned to its endpoints.

Similarly to the case of ordinary graphs and their incidence matrices, we do not distinguish between the formal weighted incidence vector representation of a q-graph, and its visual form given by a vertex set and double-weighted edges. Most proofs will rely on the incidence vector representation of the q-edges, whereas intuitive arguments are more comprehensible if they are visually aided by standard graph theoretical sketches.

Example 2.7. *Consider the following matrix for $q = 8$:*

$$\begin{pmatrix} 1 & 1 & 3 & 7 & 0 & 0 \\ 0 & 5 & 0 & 0 & 1 & 3 \\ 0 & 0 & 7 & 3 & 1 & 2 \\ 4 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

This matrix defines a q-graph that may be depicted the following way:



It is worth pointing out a few things we can observe. First, a pair of vertices might be the support of more then one q-edges. Second, a q-edge might have the same weight for both of its support vertices. Third, unlike for ordinary graphs, the two non-zero entries of an incidence vector for a q-edge are not interchangeable; that is, the q-edges x and x' with $S_x = \{v, s\}$, $x_v = 7$, $x_s = 3$ and $S_{x'} = \{v, s\}$, $x'_s = 7$, $x'_v = 3$ are different. Lastly, the q-edges of a q-graph have entries from the set $\{1, \ldots, q\}$, but it does not entail that any of the entries will be exactly q. To put it another way, a q-graph H is a $(q+1)$-graph at the same time, on the same vertex set.

Another important note to make is that substituting $q = 1$ in the formulae yields the definition of ordinary graphs. The weighted incidence vector of a q-edge simplifies into the incidence vector of an ordinary edge. Therefore, q-graphs indeed generalise the notion of ordinary graphs.

For the purpose of defining the extremal number $\text{ex}(n, F)$ for the case of q-graphs, the authors of [10] first determined when a q-graph Q contains an ordinary graph F.

Two q-edges $e, f \in Q$, are said to s-intersect at the vertex v, if $v \in S_e$, $v \in S_f$ and the sum of the entries of the incidence matrix at (v, e) and (v, f) is at least s. The q-graph Q is thus said to contain an s-copy of the ordinary graph F, if there is a set of q-edges in Q which, when deprived of their weights, form a graph isomorphic to F; and each pair of incident q-edges s-intersect. Formally,

Definition 2.8 (Patkós, Tuza, Vizer). *Let $F = (v(F), E(F))$ be an ordinary graph without isolated vertices, and $H \subseteq \mathcal{Q}(n, 2)$ be a q-graph on n vertices. Then H is an* s-copy *of F if $(\bigcup_{\mathbf{x}\in\mathbf{H}} S_{\mathbf{x}}, \{S_{\mathbf{x}} : \mathbf{x} \in \mathbf{H}\})$ is isomorphic* [2] *to F, and there exists an isomorphism $\iota : F \to (\bigcup_{\mathbf{x}\in\mathbf{H}} S_{\mathbf{x}}, \{S_{\mathbf{x}} : \mathbf{x} \in \mathbf{H}\})$ such that for every $uv, wv \in E(F), u \neq w$, it holds that the q-edges x, x' in H with $S_x = \{\iota(u), \iota(v)\}, S_{x'} = \{\iota(w), \iota(v)\}$ satisfy the condition $x_{\iota(v)} + x'_{\iota(v)} \geq s$.*

If F contains isolated vertices, then $H \subseteq \mathcal{Q}(n, 2)$ is said to be an s-copy of F if $n \geq |V(F)|$ and H is an s-copy of $F[U]$, where U is the set of non-isolated vertices in F.

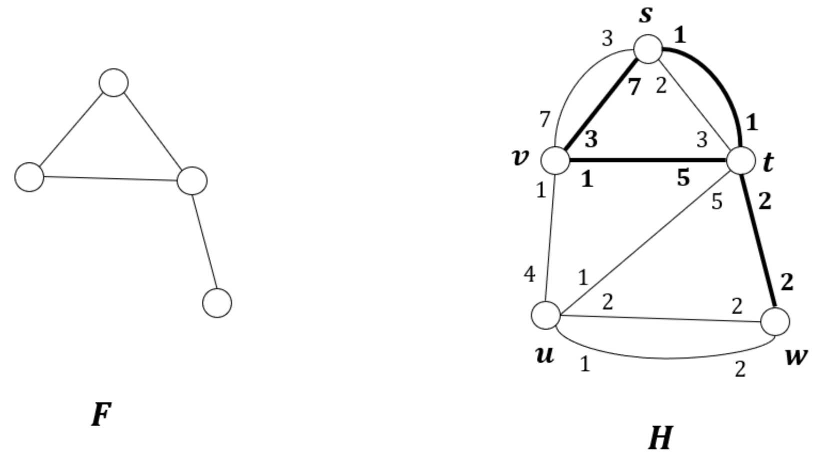


Figure 1: An ordinary graph F and an 8-graph H with a 3-copy of F on the support $\{v, s, t, w\}$.

In Figure 1, we provide an example for an 8-graph containing a 3-copy of a given ordinary graph F. A natural question one may formulate is how big s can be such

[2]We call an $\iota : F \to (V, E)$ an isomorphism if $\iota : V(F) \to V$ is a bijection that induces $\iota : E(F) \to E$ such that $\iota(e) = \{\iota(v) : v \in e\} \forall e \in E(F)$

that H contains an s-copy of F. That leads us to the concept of Turán numbers for q-graphs:

Definition 2.9 (Patkós, Tuza, Vizer)**.** *For a graph F and integers $n, q, s \geq 1$,*

$$ex(n, F, q, s) := max\{|H| : H \subseteq \mathcal{Q}(n, 2),\ H \text{ does not contain an } s\text{-copy of } F\}.$$

Furthermore,

$$EX(n, F, q, s) :=$$

$$:= \{H : H \subseteq \mathcal{Q}(n, 2), H \text{ does not contain an } s\text{-copy of } F,\ |H| = ex(n, F, q, s)\}.$$

Again, we emphasize that this definition includes the extremal number for ordinary graphs: substituting $q = 1, s = 2$ in the above formula yields $\text{ex}(n, F)$, as 2-intersecting 1-edges are just ordinary graph-edges with a common endpoint.

Very much like in [10], we only address the case $s = q+1$, for which we introduce a special notation:

Notation 2.10 (Patkós, Tuza, Vizer)**.** *$ex(n, F, q, q+1) = ex(n, F, q)$, $EX(n, F, q, q+1) = EX(n, F, q)$*

The reason behind this is that most of the other cases are redundant, or can be retraced to $s = q + 1$. Take, for instance, an ordinary graph F without vertices of degree one, and $s = q + 2$. Then, in a q-graph in $\text{EX}(n, F, q, q + 2)$, we may include every q-edge where at least one of the labels is 1, as these edges cannot be present in a $q + 2$-copy of G, since $1 + x_v \leq 1 + q < q + 2$ for every q-edge x. Therefore, we only need to pay attention to the remaining q-edges with labels from $\{2, 3, \ldots, q\}$, which (after identifying $i \in \{2, \ldots, q\}$ with $i - 1 \in \{1, \ldots, q - 1\}$) is equivalent to having a $q-1$-graph H' without a q-copy of F; thus reducing the problem to finding $\text{ex}(n, F, q - 1, q) = \text{ex}(n, F, q - 1)$.

To wrap this chapter up, we list yet another set of definitions that will appear in most of the proofs.

Definition 2.11 (Patkós, Tuza, Vizer)**.** *For $H \in \mathcal{Q}(n, 2)$, and $(a, b) \in [q]^2$, let $\overrightarrow{H}_{a,b}$ be the directed graph on $[n]$ with edges (i, j) for which the q-edge $\mathbf{x}$ with $S_{\mathbf{x}} = \{i, j\}$, $x_i = a$, $x_j = b$ appears in H. For $a, b, c, d \in [q]$, let $\overrightarrow{H}_{(a,b),(c,d)} = \overrightarrow{H}_{a,b} \cap \overrightarrow{H}_{c,d}$. Finally, let $H_{a,b}$ and $H_{(a,b),(c,d)}$ the graphs obtained by first removing orientations, and then the multiple edges from $\overrightarrow{H}_{a,b}$ and $\overrightarrow{H}_{(a,b),(c,d)}$, respectively.*

A fundamental part of the proofs rely on a partition of the q-edges with respect to the size of their two labels. Intuitively, one may wish to separately study the "large" q-edges and the others, as it turns out to be easier to handle them that way. The next definition formalizes the notion of these "large" q-edges of a q-graph.

Definition 2.12 (Patkós, Tuza, Vizer)**.** *For a q-graph $H \subseteq \mathcal{Q}(n,2)$, let*

$$H^L = \{\boldsymbol{x} \in H : S_{\boldsymbol{x}} = \{i,j\},\, x_i, x_j \geq \frac{q+1}{2}\}.$$

A label a, on many occasions, will be paired with $b = q+1-a$ to form a pair for which $\overrightarrow{H}_{a,b}$ will be examined. To simplify matters, let us introduce the following notation:

Notation 2.13. *For $a \in [q]$, let $\overline{a}$ be $q+1-a$.*

Our current way of defining a q-edge is somewhat complicated, therefore we introduce a concise notation for a q-edge with support $\{u,v\}$ and labels a and b:

Notation 2.14. *Let $x \in \mathcal{Q}(n,2)$ be a q-edge with $S_x = \{u,v\}$ and $x_u = a, x_v = b$, $a, b \in [q]$. Then (u,v,a,b) will be another notation for x.*

2.2 Basic properties of ex(n, F, q)

A trivial observation regarding the monotonicity of the extremal number $\mathrm{ex}(n,F,q)$ is the following:

Proposition 2.15. *Let F be an ordinary graph and F' its subgraph, and let $n' \leq n$, $q' \leq q$ be arbitrary integers. Then*

1. *$ex(n,F,q') \leq ex(n,F,q)$*
2. *$ex(n',F,q) \leq ex(n,F,q)$*
3. *$ex(n,F',q) \leq ex(n,F,q)$*

Proof. We only prove the first inequality, the other two can be proven using a similar reasoning. Let H be a q'-graph on n vertices without a $(q'+1)$-copy of F with a maximal number of q'-edges; that is, $H \in \mathrm{EX}(n,F,q')$. Then $q' \leq q$ implies that H is a q-graph without a $(q+1)$-copy of F, so $\mathrm{ex}(n,F,q') = |H| \leq \mathrm{ex}(n,F,q)$. $\square$

In relation to the monotone properties of $\mathrm{ex}(n, F, q)$, one can naturally ask if there was a connection between the extremal number and the total number of q-edges on n vertices, which is $q^2 \cdot \binom{n}{2}$. The result of Erdős, Stone and Simonovits in 1 stated that the Turán number of a regular, non-bipartite graph can be expressed in terms of the total number of possible edges. Ideally, we could draw a similar conclusion for q-graphs. As for now, we only know that the limit of their ratio exists.

Proposition 2.16. *For $m \leq n$,*

$$\frac{\mathrm{ex}(n, F, q)}{n(n-1)} \leq \frac{\mathrm{ex}(m, F, q)}{m(m-1)}.$$

Proof. Let $H \in \mathrm{EX}(n, F, q)$ and count the number of pairs (V', x) where $V' \subset V(H)$ is of size m, $x \in H$ and $S_x \subset V'$. Note that each q-edge in H gets counted $\binom{n-2}{m-2}$ times and for any V' there are at most $\mathrm{ex}(m, F, q)$ many q-edges in H with support in V'. The statement follows. □

We obtain the following corollary.

Corollary 2.17. *For any graph F and $q \geq 1$, the limit $\lim\limits_{n\to\infty} \frac{\mathrm{ex}(n,F,q)}{q^2\binom{n}{2}}$ exists.*

In order to strengthen the relationship between ordinary graphs and q-graphs, we present an inequality between extremal numbers and their q-ary versions.

Proposition 2.18. *For any $n \in \mathbb{N}$ and ordinary graph F, we have*

$$\mathrm{ex}(n, F, q) \geq q^2 \cdot \mathrm{ex}(n, F).$$

Proof. Let us consider an ordinary graph G with a maximal number of edges on n vertices without containing F as a subgraph. Let H be the q-graph obtained from G by putting every possible pair of labels on the edges of G. Formally, let

$$H = \{(u, v, a, b) : a, b \in [q], uv \in E(G)\}.$$

It is clear that $|H| = q^2 \cdot |E(G)| = q^2 \cdot \mathrm{ex}(n, F, q)$, and H does not admit a $(q+1)$-copy of F, since the definition of a $(q+1)$-copy requires a subset of the supports in the q-graph to be isomorphic to F. □

This proposition establishes a trivial lower bound for the q-ary Turán number, which hardly ever turns out to be sharp. The few exceptions from this rule include the case when $F = C_3$, which we will address later. In a more general case, when $F = C_{2k+1}$, $k \in \mathbb{N}^+$, this lower bound will also be the asymptotically best that one may achieve.

3 The q-extremal number of some graphs

In this section, we will closely examine three graph classes, and determine the q-extremal number for them (although for most cases, we have to settle for $q = 2$ for an exact answer). These classes include trees, the cycle on $2k + 1$ vertices where $k \in \mathbb{N}$; and bipartite graphs where all components are unicyclic or trees, with at least one component containing a cycle. Here a unicyclic graph means that exactly one cycle is contained in the graph. It is easy to see that unicyclic graphs can be constructed by taking a forest, and adding at most one edge to one of the connected components. We introduce a definition for these type of graphs.

3.1 Bipartite pseudoforests

Definition 3.1. *A simple graph F is a pseudo-forest if each of its connected components contain at most one cycle.*

With this definition, we can formulate the result of Patkós et al. the following way:

Theorem 3.2 (Patkós, Tuza, Vizer)**.** *Suppose F is a bipartite pseudo-forest. Then, for every $q, n \in \mathbb{N}$,*

$$ex(n, F, q) \leq \left(\left\lfloor \frac{q^2}{2} \right\rfloor + o(1) \right) \binom{n}{2}.$$

and if at least one of its connected components contains a cycle,

$$ex(n, F, q) = \left(\left\lfloor \frac{q^2}{2} \right\rfloor + o(1) \right) \binom{n}{2}.$$

The proof of this theorem relies on the following lemma, which gives an upper bound for $H_{(a,b),(\overline{a},\overline{b})}$ for a pair of labels (a, b). Let us give a reminder that $\overline{a} = q+1-a$, and that $H_{(a,b),(\overline{a},\overline{b})}$ contains an edge (u, v) if either (u, v, a, b) and $(u, v, \overline{a}, \overline{b})$, or (u, v, b, a) and $(u, v, \overline{b}, \overline{a})$ are present in H.

Lemma 3.3 (Patkós, Tuza, Vizer)**.** *Let G be a bipartite pseudoforest, and suppose that the q-graph H does not contain a $(q + 1)$-copy of G. Then for $(a, b) \in [q]^2$, $|H_{(a,b),(\overline{a},\overline{b})}| = o(n^2)$.*

In order to prove the lemma, we first define the bipartite Ramsey number.

Definition 3.4. *Let n and k be positive integers. The bipartite Ramsey number $R(n,k)$ is the smallest positive integer r such that in every 2-coloring of the complete bipartite graph $K_{r,r}$ with red and blue colors, there exists either a monochromatic copy of $K_{n,n}$ with red color, or a monochromatic copy of $K_{k,k}$ with blue color.*

Proof of Lemma 3.3. We claim that $H_{(a,b),(\overline{a},\overline{b})}$ does not contain $K_{R,R}$, where $R = R(|V(G)|, |V(G)|)$ is the bipartite Ramsey-number. Suppose for contradiction that it is not the case, and $H_{(a,b),(\overline{a},\overline{b})}$ contains a $(q+1)$-copy of $K_{R,R}$.

If $a, b \geq \frac{q+1}{2}$ or $\overline{a}, \overline{b} \geq \frac{q+1}{2}$, then $G \subseteq K_{|V(G)|,|V(G)|}$ implies that $H_{(a,b),(\overline{a},\overline{b})}$, along with H, contains a $(q+1)$-copy of G. This contradicts the assumption of the lemma, so it holds that (possibly after swapping the value of a with b) $a \leq \overline{b} \leq \frac{q+1}{2} \leq b \leq \overline{a}$.

Let $U \dot{\cup} V$ be a partition of the vertex set $V(H_{(a,b),(\overline{a},\overline{b})})$ obtained from a proper 2-coloring, and define the following coloring of the edges in $H_{(a,b),(\overline{a},\overline{b})}$: color an edge uv with $u \in U$ and $v \in V$ blue if $(u,v,a,b), (u,v,\overline{a},\overline{b}) \in H$, and red if $(u,v,b,a), (u,v,\overline{b},\overline{a}) \in H$ (if both conditions hold, color the edge arbitrarily).

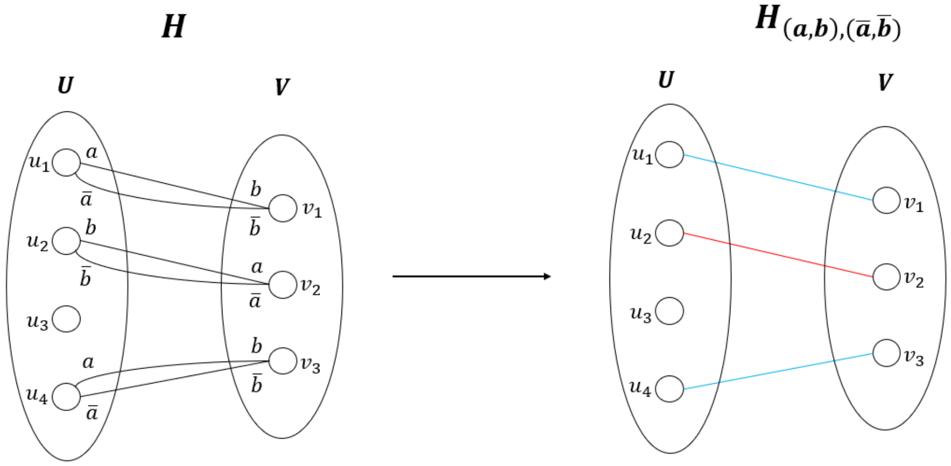


Figure 2: A 2-coloring of the edges of $H_{(a,b),(\overline{a},\overline{b})}$ based on the label pairs in H

The definition of bipartite Ramsey numbers gives a monochromatic complete bipartite subgraph on $U' \subseteq U, V' \subseteq V$ with $|U'| = |V'| = |V(G)|$; we may assume that its color is blue. Consequently, $(u,v,a,b), (u,v,\overline{a},\overline{b}) \in H$ for every $u \in U', v \in V'$. We will show that a $(q+1)$-copy of G is present in H with support $U' \cup V'$, which again would give a contradiction with the assumption of the lemma.

Using induction on the number of vertices in G, it is enough to prove this claim when G is connected. In that case, G is either a tree, or unicyclic. Let us first assume that G is unicyclic, and its only cycle is of length $2k$ (G was bipartite, so the cycle has even length). Fix the distinct vertices $u_1, v_1, u_2, v_2, \dots u_k, v_k$ such that $u_i \in U', v_i \in V'$ $\forall i$. Then the q-edges (u_i, v_i, a, b) and $(u_j, v_{(j-1 \bmod 2k)}, \overline{a}, \overline{b})$ form a $(q+1)$-copy of C_{2k}; as for every adjacent pair of q-edges, their labels at the common endpoint are either a and $\overline{a}$, or b and $\overline{b}$, and the sum of these labels is $q+1$ in both cases.

For the other branches in G rooted in C_{2k} (or in the other case when G is a tree), we can prove the claim by induction on the distance from the cycle: suppose that we have already embedded a path $w_1, w_2, \dots, w_{l-1}$ in $H|_{U' \cup V'}$ such that w_1 is the vertex of C_{2k} (when G is a tree, any vertex can be fixed to represent a cycle of length 0), and $w_{l-1} \in U'$; the other case, when $w_{l-1} \in V'$ can be treated similarly. We now want to find a vertex $w_l \in V'$ such that it has not been used yet, and there is a q-edge x in H with $S_x = \{w_{l-1}, w_l\}$ that $(q+1)$-intersects with the previously embedded q-edge y with $S_y = \{w_{l-2}, w_{l-1}\}$.

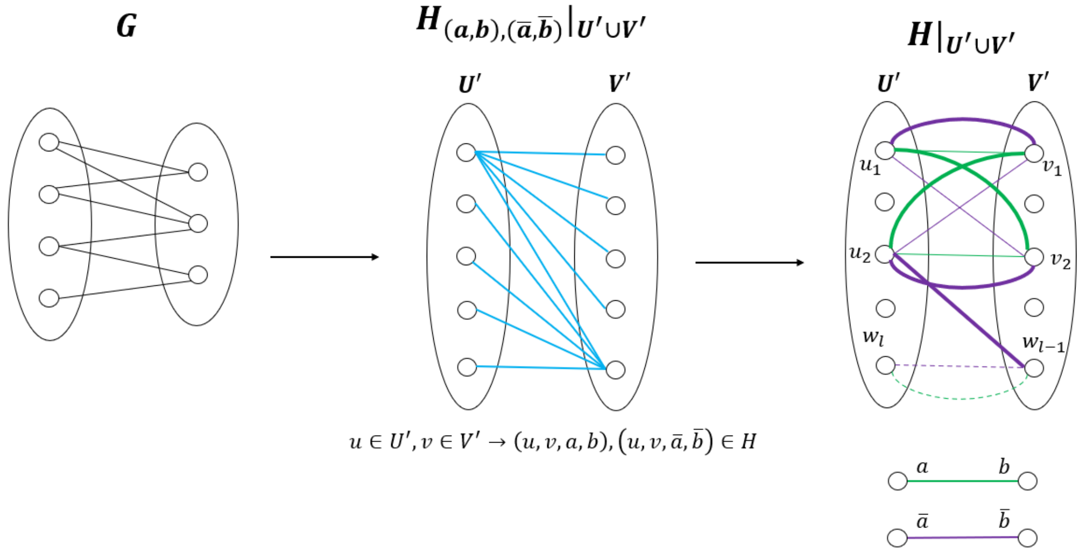


Figure 3: Embedding a $(q+1)$-copy of G to H. The next q-edge on $\{w_{l-1}, w_l\}$ can always be chosen to $(q+1)$-intersect the previous one.

Since $|V'| \leq |V| = |V(G)|$ and there is a vertex in G which we have not embedded (the last vertex of the path we are taking care of now), there is a vertex in V' which we have not used for embedding G; let this vertex be denoted by w_l. The graph $H_{(a,b),(\overline{a},\overline{b})}|_{U' \cup V'}$ was a complete bipartite graph, and $w_{l-1} \in U', w_l \in V'$, so by the choice of U' and V', the q-edges (w_{l-1}, w_l, a, b) and $(w_{l-1}, w_l, \overline{a}, \overline{b})$ are in H. In the first step of the inductive proof, when $l-1 = 1$, there are 2 q-edges of the embedded C_{2k} sharing the vertex w_1 in their support: $y_1, y_2 \in H$ with $w_1 \in S_{y_1} \cap S_{y_2}$ and $(y_1)_{w_1} = a, (y_2)_{w_1} = \overline{a}$. In this case, let x be the q-edge $(w_1, w_2, \overline{a}, \overline{b})$. Since $\overline{a} \geq \frac{q+1}{2}$, the q-edge x $(q+1)$-intersects both y_1 and y_2. Now suppose that $l-1 > 1$. If the previously embedded q-edge on the support $\{w_{l-2}, w_{l-1}\}$ was $y = (w_{l-2}, w_{l-1}, b, a)$, then let x be $(w_{l-1}, w_l, \overline{a}, \overline{b})$; if $y = (w_{l-2}, w_{l-1}, \overline{b}, \overline{a})$ then let $x = (w_{l-1}, w_l, a, b)$. The choice of x guarantees that x and y will $(q+1)$-intersect at the vertex w_{l-1}, which concludes the inductive proof.

The above method gave an embedding of G to $H|_{U' \cup V'} \subseteq H$ such that the image of intersecting edges in G is a $(q+1)$-intersecting pair of q-edges; subsequently, H contains a $(q+1)$-copy of G. This yields a contradiction with the assumption of the lemma, so the contradictory hypothesis does not stand; therefore, $H_{(a,b),(\overline{a},\overline{b})}$ does not contain $K_{R,R}$ where R is the bipartite Ramsey-number with respect to G.

From here, the lemma is the immediate consequence of the famous result of Kővári, Sós and Turán regarding the ordinary Turán-number for complete bipartite graphs:

□

Theorem 3.5 (Kővári, Sós, Turán)**.** *For any fixed $s \leq t$,*

$$ex(n, K_{s,t}) = O(n^{2-1/s})$$

Proof of Theorem 3.2. For the upper bound on $\mathrm{ex}(n, F, q)$, consider a q-graph $H \subseteq \mathcal{Q}(n, 2)$ without a $(q+1)$-copy of F, where F is a bipartite pseudoforest, and partition the q-edges of H by the label pairs. For every $(a, b) \in [q]^2$, it holds by Lemma 3.3 that $|H_{(a,b),(\overline{a},\overline{b})}| = o(n^2)$. Subsequently, for a fixed pair $(a, b) \in [q]^2$,

$$|\{(u, v, a, b) \in H : 1 \leq u < v \leq n\}| + |\{(u, v, \overline{a}, \overline{b}) \in H : 1 \leq u < v \leq n\}| \leq$$

$$\leq \binom{n}{2} + |H_{(a,b),(\overline{a},\overline{b})}| = (1 + o(1)) \cdot \binom{n}{2}.$$

Summing it up for the $\left\lfloor \frac{q^2}{2} \right\rfloor$ different pairs of $(a, b), (\overline{a}, \overline{b}) \in [q]^2$, we conclude that

$$|H| \le \left(\left\lfloor \frac{q^2}{2} \right\rfloor + o(1) \right) \cdot \binom{n}{2}.$$

When F is a bipartite pseudoforest with at least one cycle (of even length), we can come up with a construction that does not contain a $(q+1)$-copy of F, and the number of its q-edges coincides with the upper bound from the first part, proving that equality holds in this special case. As a matter of fact, the construction of universal q-trees will be a valid q-graph for every graph G that contains any cycle, not just bipartite pseudoforests. Simply put, a universal q-tree is the q-graph with vertex set $[n]$ such that it contains every "small" q-edge (whose two labels add up to at most $q+1$), but the q-edges for which the label sum is exactly $q+1$ are oriented towards the bigger vertex.

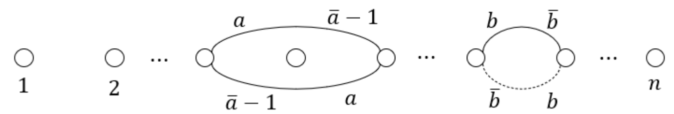
Definition 3.6 (Universal q-tree)**.** *For $n, q \ge 1$, let $U_{q,n} = U^{<} \cup \bigcup_{1 \le u < v \le n} U^{u,v}$, where*

$$U^{<} = \{(u, v, a, b) : u, v \in [n], a, b \in [q], a + b < q + 1\},$$

and

$$U^{u,v} = \{(u, v, a, \overline{a}) : u, v \in [n], a \in [q], u < v, a < \frac{q+1}{2}\}.$$

As we mentioned before, the universal q-graph does not contain a $(q+1)$-copy of any cycle, and therefore it does not contain a $(q+1)$-copy of any graph with at least one cycle.



$\boldsymbol{U_{q,n}}$

Figure 4: The universal q-tree. If $b < \overline{b}$, the label pair $(b, \overline{b})$ only appears once at every support pair.

Lemma 3.7. *The universal q-tree does not contain a $(q+1)$-copy of any cycle.*

Proof. Suppose $C \subseteq U_{q,n}$ is a $(q+1)$-copy of a cycle. As every pair of incident q-edges in C $(q+1)$-intersect, it follows that

$$|C| \cdot (q+1) \geq \sum_{x \in C} \sum_{h=1}^{n} x_h = \sum_{v \in S_C} \sum_{x \in C} x_v \geq |C| \cdot (q+1),$$

which means that $C \subseteq \bigcup_{1 \leq u < v \leq n} U^{u,v}$. But this gives a contradiction: if $u = \min\{v \in [n] : v \in S_C\}$, then the two q-edges in C adjacent to v, (v, u, a, b) and (v, u', a', b') satisfy that $a + a' < 2 \cdot \frac{q+1}{2}$, hence they do not $(q+1)$-intersect. $\square$

To end the proof of Theorem 3.2, let us point out that

$$|U_{q,n}| = \left\lfloor \frac{q^2}{2} \right\rfloor \cdot \binom{n}{2},$$

which asymptotically equals to the upper bound of the theorem, $\left(\left\lfloor \frac{q^2}{2} \right\rfloor + o(1)\right) \cdot \binom{n}{2}$. Furthermore, Lemma 3.7 showed that $U_{q,n}$ does not admit a $(q+1)$-copy of a graph with at least one cycle, which means that $U_{q,n} \in \mathrm{EX}(n, F, q)$.

$\square$

The upper bound of Theorem 3.2 is valid for any bipartite pseudoforest, but we only prove that the construction of the universal q-tree is $(q+1) - F$-free for graphs that contain at least one cycle. This begs the question whether the universal q-tree is a good example of a q-graph without a $(q+1)$-copy of a tree, but the answer is negative: one can easily show that $U_{q,n}$ contains a $(q+1)$-copy of any given forest on n vertices.

This means that in case F is a forest, two options remain: we can either present another construction of the same size without $(q+1)$-copies of the forest; or we give a better upper bound on $\mathrm{ex}(n, F, q)$ in this special case. We will explore the latter approach in the following part.

3.2 Trees

As we will soon see, a central concept in the study of trees plays a vital role in determining the 2-extremal number. The center of a tree, as it can be expected, denotes a vertex in the tree which is as closely located to every other vertex as possible.

Definition 3.8. *If T is a tree, then the center of T is*

$$c(T) = \arg \min_{v \in V(T)} \max_{u \in V(T)-v} d(u,v),$$

and the radius of T is the maximal distance from the center to another vertex:

$$r(T) = \min_{v \in V(T)} \max_{u \in V(T)-v} d(u,v),$$

where $d(u,v)$ denotes the distance of u and v in T, i.e. the length of the unique path between u and v in T.

The radius of a tree is always unique, however the center, which determines the radius, is not necessarily. Take, for instance, a path of odd length; for which two distinct nodes minimize the maximal distance from the other vertices.

Another useful definition, the diameter of a tree, shows the maximal distance between two vertices.

Definition 3.9. *The diameter of a tree T is*

$$d(T) = \max_{u,v \in V(T)} d(u,v).$$

It is easy to see that the diameter and the radius are closely related: it holds that $d(T)$ is either $2r(T)$ or $2r(T)+1$. Indeed, in the unique path between u and v in T for which $d(u,v) = d(T)$, the middle vertex (or vertices) w satisfy that $\lfloor \frac{d(T)+1}{2} \rfloor = \max\{d(v,w), d(u,w)\} \leq r(T)$; and on the other hand, $d(u,v) \leq d(u,c(T)) + d(v,c(T)) \leq 2 \cdot r(T)$ for every pair (u,v) in T.

With the help of these prerequisites, we can formulate the result for the 2-extremal number of trees the following way:

Theorem 3.10 (Patkós, Tuza, Vizer)**.** *Suppose T is a tree of radius r.*

(1) If the diameter of T is $2r$, then $ex(n,T,2) = (1+o(1)) \cdot \left(\binom{n}{2} + t_{n,r}\right)$.

(2) If the diameter of T is $2r-1$, then $ex(n,T,2) = (1+o(1)) \cdot \left(\binom{n}{2} + t'_{n,r} - \binom{\lfloor \frac{n}{2r-1} \rfloor}{2}\right)$.

Here, $t_{n,r}$ denotes the number of edges in the r-partite Turán graph on n vertices, and $t'_{n,r}$ is the number of edges in the complete r-partite graph on n vertices, where one class has size $\lfloor \frac{n}{2r-1} \rfloor$, and the other class sizes differ by at most one.

Definition 3.11. *The Turán graph $T_{n,r}$ on n vertices with r classes is the (unique) complete r-partite graph on n vertices , where class sizes are either $\lfloor \frac{n}{r} \rfloor$ or $\lceil \frac{n}{r} \rceil$ and are selected in a way that the r class sizes add up to exactly n.*

The Turán-number $t_{n,r}$ is the size of the Turán-graph, that is,

$$t_{n,r} = |E(T_{n,r})|.$$

Proof of Theorem 3.10. Let H be a 2-graph on n vertices without a 3-copy of T. To begin with, let us focus on the first statement. As in the previous proofs, the partition of the q-edges according to the label pairs will be fruitful for setting the upper bound.

For the q-edges in $H_{1,1}$, we only need the trivial upper bound $\binom{n}{2}$. Moreover, $|H_{(1,2),(2,1)}|$ is bounded by $o(n^2)$ according to Lemma 3.3, since the tree T is a bipartite pseudoforest, and $q = 3$ gives $\overline{1} = 2, \overline{2} = 1$. Furthermore, observe that $|H_{2,2}| \leq \mathrm{ex}(n, T) = o(n^2)$, since we can apply Theorem 3.5 to the bipartite graph T. Now it only remains to deal with edges in $\overrightarrow{H}_{1,2}$ which we did not count in $H_{(1,2),(2,1)}$. For that purpose, let H' denote the graph obtained by removing directed cycles of length 2 from $\overrightarrow{H}_{1,2}$; the directed graph H' represents the q-edges we have not accounted for so far.

Let $V_1, \dots, V_t$ be a partition of the vertices as follows: V_1 is the set of vertices v for which $\varrho_{H'}(v) < |T|$, and then inductively

$$V_{i+1} := \{v \in V \setminus \left(\cup_{j=1}^{i} V_j\right) : \varrho(v) < |T| \text{ in } H'[V \setminus \left(\cup_{j=1}^{i} V_j\right)]\}.$$

Observe that if $V \setminus \left(\cup_{j=1}^{i} V_j\right) \neq \emptyset$, then $V_{i+1} \neq \emptyset$: if every in-degree in $H'[V \setminus \left(\cup_{j=1}^{i} V_j\right)]$ is at least $|T|$, then a 3-copy of T with support $V \setminus \left(\cup_{j=1}^{i} V_j\right)$ can be greedily embedded in H. Indeed, start the embedding from an arbitrary vertex in $V \setminus \left(\cup_{j=1}^{i} V_j\right)$ (this node will play the role of the center), and while there is a vertex in T that we have not used yet, we can choose the next vertex from the (at least) $|T|$ neighbours of the previous vertex such that the new vertex has not been used before; this will guarantee that adjacent 2-edges will 3-intersect.

We may similarly deduce that $t \leq r$: if $v \in V_{r+1}$, then we can again greedily embed a 3-copy of T to H with support $\{v\} \cup (\cup_{j=1}^{r} V_j)$, where the vertex v plays the role of $c(T)$. Note that here we needed T to have radius r, as we might not be able to construct a counterexample for a tree of radius $r+1$: it might happen that after r consecutive steps, we arrive at V_1, for which we may not find an appropriate vertex in the subsequent step (provided that each of its at most $|T|-1$ neighbours have been used before).

Consequently, the edges in H' spanned by the partition classes add up to at most $(|T|-1)n$, and the remaining edges form a t-partite graph for $t \leq r$. Combining these two observations, we may conclude that $|H'| \leq (|T|-1) \cdot n + t_{n,r}$. Altogether, it holds that

$$|H| = |H_{1,1}| + |H_{2,2}| + |H_{(1,2),(2,1)}| + |H'| \leq$$
$$\leq \binom{n}{2} + o(n^2) + o(n^2) + t_{n,r} = (1+o(1)) \cdot \left(\binom{n}{2} + t_{n,r} \right).$$

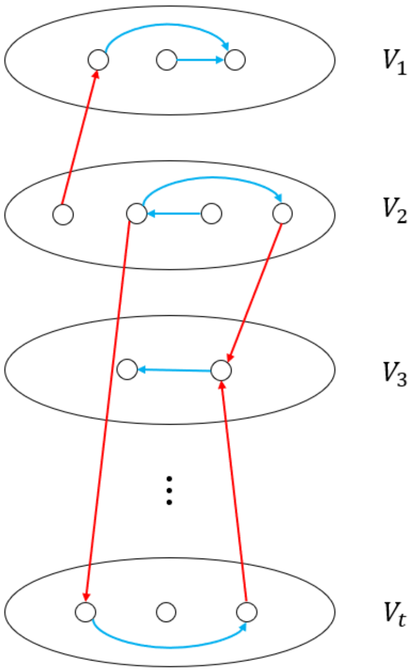


Figure 5: The digraph H'. For every $i = 1, \dots, t$ and $v \in V_i$, $\varrho(v) \leq |T|-1$ in $H'[V_i]$. The remaining edges form a t-partite graph on n vertices.

For the upper bound in the second case, we only need to slightly adjust the previous method. Suppose that T has diameter $2r-1$, and consider the same partition $V_1, \dots, V_t$ as before. It can be easily shown that when $r(T) = r$ and $d(T) = 2 \cdot r(T) - 1$, the center of the tree consist of two separate nodes. Using this observation, we can prove that $H_{1,1}$ does not span any edge in V_t in this case. If there is $u, v \in V_t$ such that $uv \in H_{1,1}$, then we can start the greedy embedding from u and v as the center of T, and then in each step, embedding a neighbour of the current vertex $v \in V_i$ in $V_{i-1} \cup V_i \cup \ldots \cup V_t$. As long as $v \in V_i$ for $i > 1$, we are able to choose such a vertex, since v has in-degree at least $|T|$ in $V_{i-1} \cup V_i \cup \ldots \cup V_t$. The condition that $r(T) = r, d(T) = 2r - 1$ and $c(T) = \{u, v\} \subseteq V_t$ (we might assume $t = r$) ensures that when $v \in V_1$, the corresponding vertex in T was a leaf, and therefore we do not need to find a neighbour of v that has not been used already.

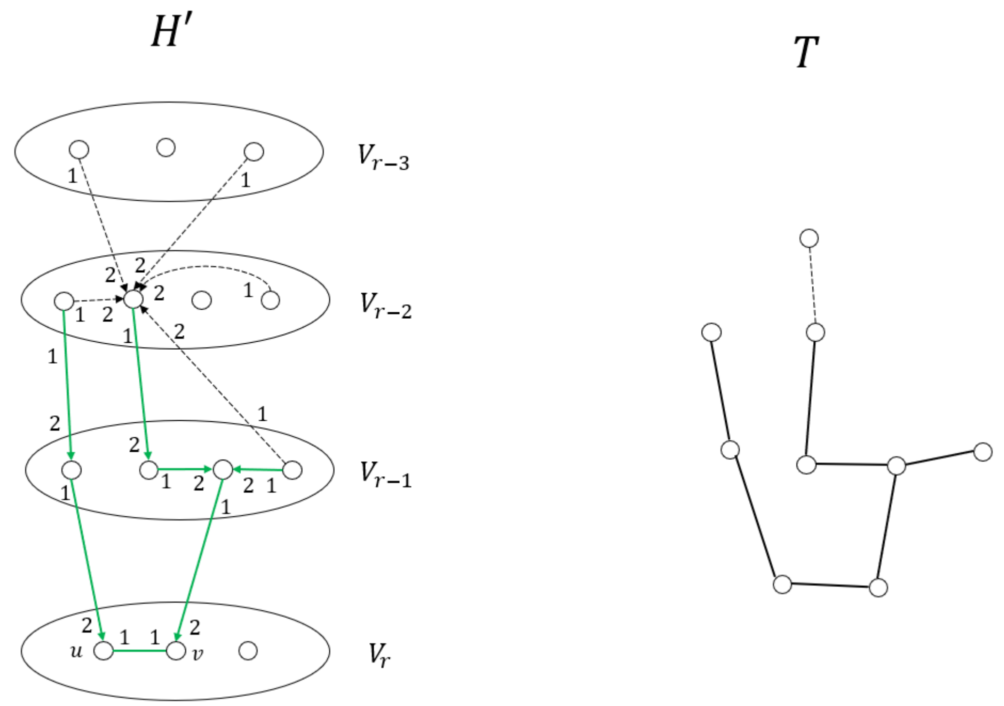


Figure 6: The greedy algorithm for embedding T in H if $H_{1,1}$ has an edge in V_r.

The updated upper bounds we got from the above reasoning are as follows: the edges in $H_{(1,2),(2,1)}$ and $H_{2,2}$ are bounded by the same limit as in the first case, $|H'|$ is at most the edge number of a complete r-partite graph on $V_1, \ldots, V_r$, and $|H_{1,1}|$ is bounded by the number of vertex pairs not completely inside V_r. We are looking to maximize the latter two expressions; and if we fix the size of V_r, the maximum is achieved if $V_1, \ldots, V_{r-1}$ have almost equal sizes. In this case, we have $r - 1$ classes of size αn, and one class of size $(1 - (r - 1)\alpha)n$. Then, apart from a sub-quadratic error, we need to bound the expression

$$\binom{n}{2} - \binom{(1-(r-1)\alpha)n}{2} + \binom{n}{2} - (r-1)\binom{\alpha n}{2} - \binom{(1-(r-1)\alpha)n}{2},$$

which takes its maximum at $\alpha = \frac{2n}{2r-1}$, proving the upper bound of the second claim.

In order to show that these upper bounds are asymptotically accurate, let us consider the following constructions: let $V_1, \ldots, V_r$ be a partition of the vertex set $[n]$ where $|V_i| = \lfloor \frac{n}{r} \rfloor$ or $|V_i| = \lceil \frac{n}{r} \rceil$, $\forall i$; and $V_1', \ldots, V_r'$ another partition where $|V_r'| = \lfloor \frac{n}{2r-1} \rfloor$, and the other class sizes differ by at most one. The 2-graph H^1 will consist of every $(1,1)$-edge on $[n]$, and every $(1,2)$-edge between the partition classes $V_1, \ldots, V_r$ such that the label 2 will belong to the class with the larger index:

$$H^1 = \{(u, v, 1, 1) : u, v \in [n]\} \cup \{(u, v, 1, 2) : u \in V_i, v \in V_j, i < j\}.$$

The 2-graph H^2 has a similar structure: it has every $(1,2)$-edge between the classes $V_1', \ldots, V_r'$ with the 2-labels at the larger index, and almost every $(1,1)$-edge, with the exception of $(1,1)$-edges spanned by V_r':

$$H^2 = \{(u, v, 1, 1) : u, v \in [n], \{u, v\} \not\subseteq V_r'\} \cup \{(u, v, 1, 2) : u \in V_i', v \in V_j', i < j\}.$$

Clearly, $|H^1| = t_{n,r} + \binom{n}{2}$, $|H^2| = \binom{n}{2} - \binom{\lfloor \frac{n}{2r-1} \rfloor}{2} + t'_{n,r}$, so it suffices to prove that the longest 3-copy of a path in H^1 has length at most $2r-1$, and the longest 3-copy of a path in H^2 has length at most $2r-2$. It will then follow that H^1 does not have a 3-copy of a tree with diameter $2r$, and H^2 does not have a 3-copy of a tree with diameter $2r-1$, which concludes the proof of Theorem 3.10.

Let us focus on the claim for H^1; the one for H^2 has an almost identical proof. Notice that a 3-copy of any path cannot have an inner vertex in V_1, since any 2-edge x with a support vertex $v \in V_1$ has $x_v = 1$. Consequently, the 3-copy of the path can only have at most 2 2-edges with a support vertex in V_1: the first and last edges of the path. After removing V_1 from the vertex set, and the 2 (possible) 2-edges adjacent to it, it follows from induction that the remaining part of the 3-copy of the path can contain at most $2r-3$ 2-edges. Adding back the (at most) 2 removed 2-edges, we finally get that any path with a 3-copy in H^1 must be of length at most $2r-1$.

□

As the final part of this section, we mention that in the case of forests, we can give a similar bound for the 2-extremal number as in Theorem 3.10. This observation is based on the following: suppose that the forest F consists of the tree components $T_1, \ldots, T_k$, with the radii in decreasing order. Let us choose a vertex from the center in each component: $v_1 \in C(V_1), \ldots, v_k \in C(V_k)$, and construct a tree T from the trees $T_1, \ldots, T_k$ by adding the edges $\{v_1 v_i : i = 2, \ldots, k\}$. It will hold that $v_1 \in c(T)$, $r(T) \leq r(T_1) + 1$. Indeed, for a vertex v in T_1, $d_T(v, v_1) \leq r(T_1)$ by the definition of the center; and for $v \in T_i, i \neq 1$, $d_T(v, v_1) = d_T(v_1, v_i) + d_T(v_i, v) \leq 1 + r(T_i) \leq 1 + r(T_1)$, as T_1 had the largest radius.

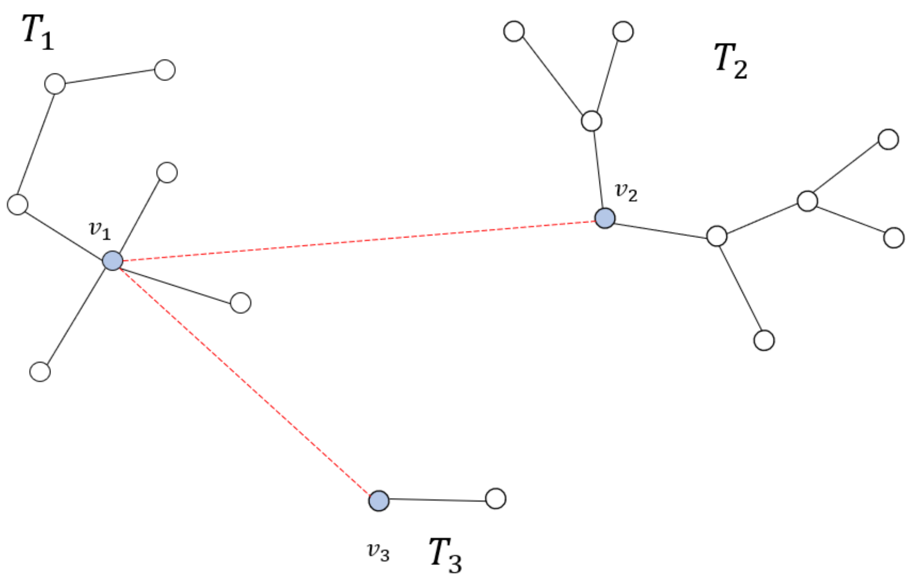


Figure 7: The tree T with $r(T) \leq \max\{r(T_i) : i = 1, \ldots, k\} + 1$

Since a 2-graph without a 3-copy of $\cup_{i=1}^{k} T_k$ cannot contain a 3-copy of T, it follows that the same general upper bound from Theorem 3.10 is valid for F. Unfortunately, $d(T)$ cannot be expressed by the radii and the diameters of $T_1, \ldots, T_k$, so we cannot specify the cases when the stronger upper bound is applicable (i.e. when $d(T) = 2r(T) - 1$). Consequently, in the general case, the best upper bound we can achieve is $\mathrm{ex}(n, F, 2) \leq (1 + o(1)) \cdot \left(\binom{n}{2} + t_{n, r(T_1)+1}\right)$.

As for a lower bound, we can say that a 2-graph without a 3-copy of T_1 is a

2-graph without a 3-copy of F, and in conclusion,

$$\mathrm{ex}(n, T_1, 2) \leq \mathrm{ex}(n, F, 2) \leq \mathrm{ex}(n, T, 2),$$

or more precisely,

Corollary 3.12. *Suppose F is a forest with components $T_1, \ldots, T_k$ such that $r = r(T_1) \geq \ldots \geq r(T_k)$. Then*

$$(1+o(1)) \cdot \left(\binom{n}{2} + t'_{n,r} - \binom{\lfloor \frac{n}{2r-1} \rfloor}{2} \right) \leq ex(n, F, 2) \leq (1+o(1)) \cdot \left(\binom{n}{2} + t_{n,r+1} \right).$$

In a special setup, we can strengthen the lower and upper bounds to coincide with each other, and therefore determine the exact value of $\mathrm{ex}(n, F, 2)$. When $d(T_1) = 2 \cdot r(T_1)$, the lower bound can be replaced with $(1+o(1)) \cdot \left(\binom{n}{2} + t_{n,r(T_1)}\right)$; and when $r(T_2) < r(T_1)$, the radius of T will simply become $r(T_1)$: $d_T(v, v_1) \leq d_T(v, v_i) + d_T(v_i, v_1) \leq r(T_1) \quad \forall v \in [n]$.

Corollary 3.13. *Suppose F is a forest with components $T_1, \ldots, T_k$ such that $r = r(T_1) > r(T_2) \geq \ldots \geq r(T_k)$, and $d(T_1) = 2 \cdot r(T_1)$. Then*

$$ex(n, F, 2) = (1+o(1)) \cdot \left(\binom{n}{2} + t_{n,r} \right).$$

3.3 Circles of odd length

In the previous two subsections, our focus was on determining the 2-extremal number for trees, and bipartite pseudoforests with at least one cycle. A bipartite pseudoforest without a cycle is just a forest, for which we gave an upper bound on the 2-extremal number. A reasonable next step would be to examine a well-known graph class with a simple structure, that does not fall under the previous definitions. The best candidate for this role is a cycle of odd length, which we will turn to discuss in the upcoming subsection.

When $F = C_3$, the usual partition of the label pairs prove to be sufficient for a simple inductive proof. For k greater than 1, the 2-Turán number of C_{2k+1} will be determined by studying C_{4k-2}; an idea which we will use again later in the thesis.

As we remarked in Proposition 2.18, C_{2k+1} is one of the few exceptions when the trivial lower bound yields the exact value of the q-Turán number:

Theorem 3.14 (Patkós, Tuza, Vizer)**.** *For $n \geq 2$,*

$$ex(n, C_3, 2) = 4 \cdot ex(n, C_3) = 4 \left\lfloor \frac{n^2}{4} \right\rfloor.$$

Proof. The proof follows the same pattern as many times before. For a 2-graph H on n vertices without a 3-copy of C_3, $|H_{1,1} \cup H_{2,2}|$ and $|\overrightarrow{H}_{1,2}|$ will be independently bounded, and will jointly give the desired value.

We use induction on n to prove that $|H_{1,1} \cup H_{2,2}| \leq 2 \cdot \lfloor \frac{n^2}{4} \rfloor$, and $|\overrightarrow{H}_{1,2}| \leq 2 \cdot \lfloor \frac{n^2}{4} \rfloor$. We begin with the first statement. Let $\{u, v\}$ be a vertex pair for which $(u, v, 2, 2) \in H$. If there are no such vertices, then $|H_{1,1} \cup H_{2,2}| = |H_{1,1}| \leq \binom{n}{2} \leq 2 \cdot \lfloor \frac{n^2}{4} \rfloor$. As the number of 2-edges is maximized, we may assume that $(u, v, 1, 1) \in H$. By induction, $|H_{1,1}[[n] - \{u, v\}] \cup H_{2,2}[[n] - \{u, v\}]| \leq 2 \cdot \lfloor \frac{(n-2)^2}{4} \rfloor$. Let us group the remaining potential 2-edges in $H_{1,1} \cup H_{2,2}$ between $\{u, v\}$ and $[n] - \{u, v\}$ into classes of size 2 in the following way: for $w \in [n] - \{u, v\}$, the pair of 2-edges $(u, w, 1, 1), (v, w, 2, 2)$ will be a partition class, and the pair of 2-edges $(u, w, 2, 2), (v, w, 1, 1)$ will be another class. After preparing the classes for every $w \in [n] - \{u, v\}$ we may observe that H cannot contain both edges of a class, or else the triple $(u, v, 2, 2), (u, w, 1, 1), (v, w, 2, 2)$ (or the triple $(u, v, 2, 2), (u, w, 2, 2), (v, w, 1, 1)$) would form a 3-copy of C_3. Hence, H can contain at most one 2-edge from each of the $2(n-1)$ partition classes; in total, $|H_{1,1} \cup H_{2,2}| \leq 2 + 2 \cdot \lfloor \frac{(n-2)^2}{4} \rfloor + 2(n-1) = 2 \cdot \lfloor \frac{n^2}{4} \rfloor$.

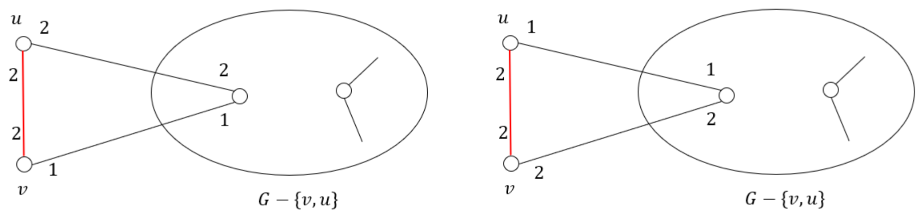


Figure 8: The partition classes in $H_{1,1} \cup H_{2,2}$

The proof of the second inequality is essentially the same. Consider a pair of vertices $\{u, v\} \in [n]^2$ such that $(u, v, 1, 2) \in H$ and $(u, v, 2, 1) \in H$. If there is no such vertex pair, $|\overrightarrow{H}_{1,2}| \leq \binom{n}{2} \leq 2 \cdot \lfloor \frac{n^2}{4} \rfloor$. By induction, $|\overrightarrow{H}_{1,2}[[n] - \{u, v\}]| \leq 2 \cdot \lfloor \frac{(n-2)^2}{4} \rfloor$.

Consider the following parititon of the 2-edges in $\overrightarrow{H}_{1,2}$ between $\{u,v\}$ and $[n]-\{u,v\}$:

$$\{(u,w,1,2),(v,w,2,1) : w \in [n]-\{u,v\}\} \cup \{(u,w,2,1),(v,w,1,2) : w \in [n]-\{u,v\}\}.$$

H can contain at most one 2-edge from each of the $2(n-2)$ partition classes, or else the triple $(u,v,1,2)$, $(u,w,2,1)$, $(v,w,1,2)$, or the triple $(u,v,2,1)$, $(u,w,1,2)$, $(v,w,2,1)$ would form a 3-copy of C_3 in H. Consequently,

$$|\overrightarrow{H}_{1,2}| \leq 2 + 2 \cdot \lfloor \frac{(n-2)^2}{4} \rfloor + 2(n-2) = 2 \cdot \left\lfloor \frac{n^2}{2} \right\rfloor.$$

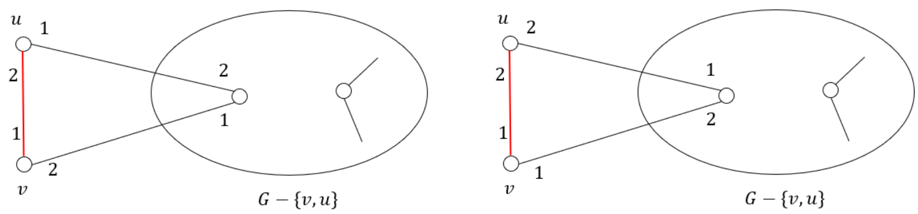


Figure 9: The partition classes in $\overrightarrow{H}_{1,2}$

It follows that

$$|H| = |\overrightarrow{H}_{1,2}| + |H_{1,1} \cup H_{2,2}| \leq 4 \cdot \left\lfloor \frac{n^2}{4} \right\rfloor.$$

To see that this upper bound is achievable, consider the complete bipartite graph on n vertices with class sizes $\lfloor \frac{n}{2} \rfloor$ and $\lceil \frac{n}{2} \rceil$, and put each 4 potential pair of labels from $\{1,2\}$ on every edge. The resulting 2-graph will not contain any copy of a triangle, and has exactly $4 \cdot \lfloor \frac{n^4}{4} \rfloor$ 2-edges.

$\square$

For $k > 1$, it is not that trivial to prove the same statement, and we can only manage to asymptotically determine the 2-extremal number.

Proposition 3.15 (Encz, Marits, Váli, Weisz)**.** $\mathrm{ex}(n, C_{2k+1}, 2) = \left(\lfloor \frac{2^2}{2} \rfloor + o(1)\right) \cdot \binom{n}{2} = n^2 + o(n^2)$.

The core of the proof stems from the same idea as in Proposition 4.6; namely, to use the already established results for $2(2k+1)$ (or in this case, $2(2k-1)$). Let us recall from Definition 2.11 that $\overrightarrow{H}_{a,b} = \{(u,v) \in [n]^2 : (u,v,a,b) \in H\}$ with an edge (u,v) directed from u to v, $H_{a,b} = \{(u,v) \in [n]^2 : (u,v,a,b) \in H \text{ or } (v,u,a,b) \in H\}$, and $H_{(a,b),(c,d)}$ is $\overrightarrow{H}_{a,b} \cap \overrightarrow{H}_{c,d}$ without orientations and multiple edges. We will use a result of Zhou and Li (Theorem 1.8. from [9]), and Lemma 3.3 from a previous section, which gave an upper bound of $o(n^2)$ on $|H_{(a,b),(\overline{a},\overline{b})}|$ when H does not contain a given bipartite pseudoforest.

Theorem 3.16 (Li, Zhou)**.** *Let $k, n \in \mathbb{N}^*, n = qk + r, 0 \leq r < k$, and let $\overrightarrow{C_{k+1}}$ be the directed cycle on $k+1$ vertices. Then*

$$ex(n, \overrightarrow{C_{k+1}}) = \frac{1}{2}n^2 + \frac{k-2}{2}n - \frac{r(k-r)}{2}.$$

Proof of Proposition 3.15. Let H be an optimal 2-graph without a 3-copy of C_{2k+1}. To bound the number of 2-edges in $\overrightarrow{H}_{1,2}$, we rely on Theorem 3.16. It implies that

$$|\overrightarrow{H}_{1,2}| = \frac{1}{2}n^2 + \frac{2k-1}{2}n + O(1) = \frac{1}{2}n^2 + o(n^2).$$

Now we turn to examine $H_{1,1} \cup H_{2,2}$, where an edge is included twice if it is both in $H_{1,1}$ and $H_{2,2}$. We prove by induction on n that $|H_{1,1} \cup H_{2,2}| = \frac{1}{2}n^2 + o(n^2)$. Suppose for contradiction that $H_{1,1} \cup H_{2,2}$ has more edges. In that case, a 3-copy of C_{4k-2} is present in $H_{1,1} \cup H_{2,2}$: substituting $q = 2, a = b = 1$ into Lemma 3.3 gives a 3-copy of a C_{4k-2} (otherwise $|H_{1,1} \cap H_{2,2}| = o(n^2)$ would hold, and since the optimal property of H entails $H_{1,1} \cap H_{2,2} = H_{2,2}$, $|H_{1,1} \cup H_{2,2}| = |H_{1,1}| + |H_{2,2}| = |H_{1,1}| + |H_{1,1} \cap H_{2,2}| \leq \binom{n}{2} + o(n^2)$ would stand).

Let us denote the support of this 3-copy by S_{4k-2}, and the vertices in S_{4k-2} by $v_1, v_2, \ldots, v_{4k-2}$. As the pair of 2-edges attached to v_1 in the cycle 3-intersect, at least one of them must have labels $(2,2)$. The same stands for the 2-edges attached to v_{2k}. Using the symmetry of C_{4k-2}, this can happen in one of two ways: either $(v_1, v_2, 2, 2)$ and $(v_{2k}, v_{2k-1}, 2, 2)$ are in H, or $(v_1, v_2, 2, 2)$ and $(v_{2k}, v_{2k+1}, 2, 2)$ are in H. Consider an arbitrary vertex v in $S_H \setminus S_{4k-2}$.

First suppose that the 2-edges $(v_1, v_2, 2, 2)$ and $(v_{2k}, v_{2k-1}, 2, 2)$ are in H. Then we need to omit at least one 2-edge from both $\{(v, v_1, 2, 2), (v, v_{2k}, 1, 1)\}$ and $\{(v, v_1, 1, 1), (v, v_{2k}, 2, 2)\}$, or else a 3-copy of C_{2k+1} would be formed by $vv_1v_2 \ldots v_{2k}v$.

If $(v_{2k}, v_{2k-1}, 2, 2) \notin H$, then the 2-edges $(v_1, v_2, 2, 2)$ and $(v_{2k}, v_{2k+1}, 2, 2)$ are in the 3-copy of C_{4k-2} in H. Then H can only contain at most one 2-edge from both

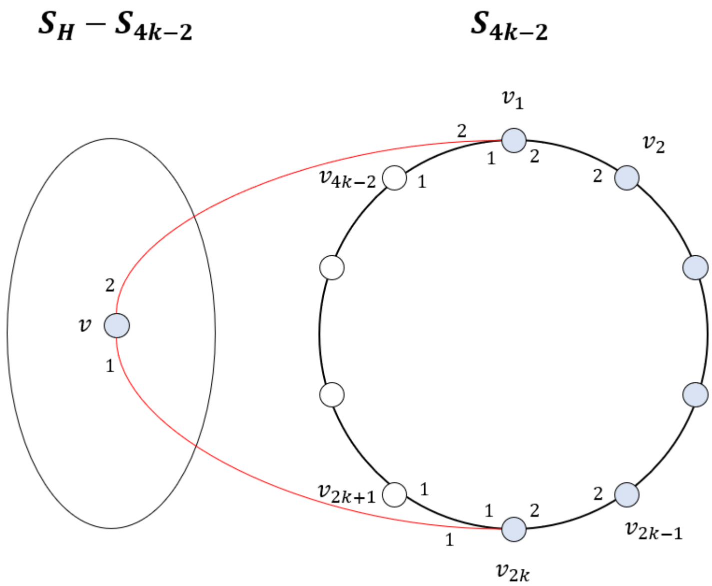


Figure 10: If $(v, v_1, 2, 2), (v, v_{2k}, 1, 1) \in H$, then $vv_1v_2 \ldots v_{2k}v$ is a 3-copy of C_{2k+1}. The same stands for $(v, v_1, 1, 1), (v, v_{2k}, 2, 2)$.

of the pairs $\{(v, v_1, 2, 2), (v, v_{2k}, 1, 1)\}$ (or else a 3-copy of C_{2k+1} would be formed by $vv_{2k}v_{2k+1} \ldots v_{4k-2}v_1v$) and $\{(v, v_1, 1, 1), (v, v_{2k}, 2, 2)\}$ (or else a 3-copy of C_{2k+1} would be formed by $vv_1v_2 \ldots v_{2k}v$). In both cases, we may conclude that out of the 4 possible 2-edges in $H_{1,1} \cup H_{2,2}$ between v and $\{v_{2k}, v_1\}$, at most 2 may be included in H.

The same reasoning can be repeated for v_i and $v_{i+2k-1 (\text{mod } 4k-2)}$ instead of v_1 and v_{2k}. By summing it up for $\{v_i, v_{i+2k-1(\text{mod } 4k-2)} : i = 1, 2, \ldots, 2k-1\}$, and for every $v \in S_H - S_{4k-2}$, we gain that at most half of the potential 2-edges between S_{4k-2} and $S_H - S_{4k-2}$ can be present in $H_{1,1} \cup H_{2,2}$. Applying the induction hypothesis shows that $\frac{1}{2}|S_H - S_{4k-2}|^2 + o(n^2)$ 2-edges can be spanned by $S_H - S_{4k-2}$ in $S_{H_{1,1}} \cup S_{H_{2,2}}$. The number of 2-edges spanned by H in S_{4k-2} can be bounded by a constant which is independent from n. Consequently, the total number of 2-edges in $H_{1,1} \cup H_{2,2}$ amounts to

$$2 \cdot |S_{4k-2}| \cdot |S_H - S_{4k-2}| + \frac{1}{2}|S_H - S_{4k-2}|^2 + o(n^2) =$$
$$= \frac{1}{2}|S_H \cup S_{4k-2}|^2 + o(n^2) = \frac{1}{2}n^2 + o(n^2),$$

which concludes our inductive proof. Finally, $|H| = |H_{1,1} \cup H_{2,2}| + |\overrightarrow{H}_{1,2}| = n^2 + o(n^2)$.

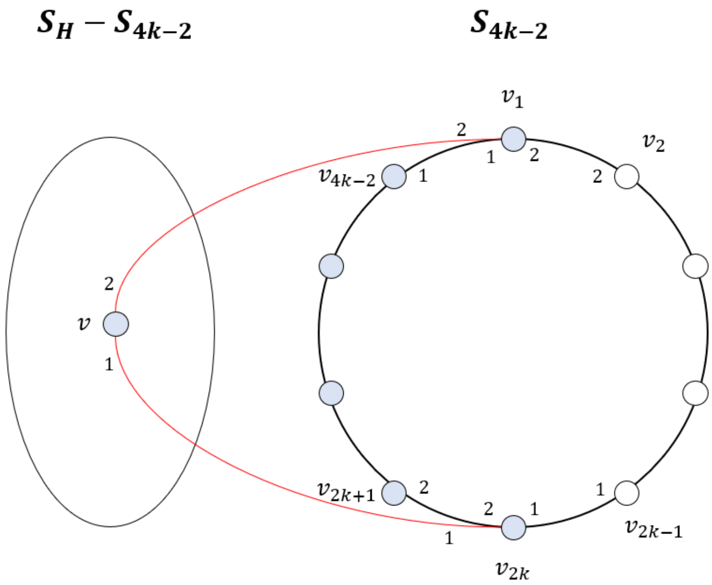


Figure 11: If $(v, v_1, 2, 2), (v, v_{2k}, 1, 1) \in H$, then $vv_{2k}v_{2k+1} \dots v_{4k-2}v_1v$ is a 3-copy of C_{2k+1}. The same stands for $(v, v_1, 1, 1), (v, v_{2k}, 2, 2)$ and $vv_1v_2 \dots v_{2k}v$.

$\square$

4 Reduction to $q = 2$ from even values of q

So far, apart from Theorem 3.2, we only presented results regarding the 2-extremal number of some graph classes. The reason behind this is that the methods used in the $q = 2$ case are not sufficient for $q > 2$. Almost every proof used the partition of the 2-edges into $\overrightarrow{H}_{1,2}$ and $H_{1,1} \cup H_{2,2}$. Unfortunately, we cannot extend this partition for $q > 2$ in a sensible way. A natural idea would be to just focus on partition classes such as $H_{a,b} \cup H_{\overline{a},\overline{b}}$ or $\overrightarrow{H}_{a,\overline{a}}$. Unfortunately, this approach does not have the desired effect, even though in some instances we can find similar upper bounds as in the previous sections.

Instead, we will show a way to connect the $q = 2$ case with $q = 2k$ for arbitrary values of k, and express $\mathrm{ex}(n, F, q)$ in terms of $\mathrm{ex}(n, F, 2)$ for any graph F and even $q \in \mathbb{N}$. The fundamental part of the proof will be establishing a connection between the problem of determining $\mathrm{ex}(n, F, q)$ and a problem for ordinary graphs which is closely related to the fractional vertex covering problem. In fact, the traditional way to show that there always exists a half-integral minimal vertex cover can be applied to our case with little to no change. Let us outline the technical background for this in the next few lemmas.

Lemma 4.1. *Let G be an ordinary graph and $q \in \mathbb{N}$. Then there is a function y taking values in $\{0, \lfloor \frac{q}{2} \rfloor, \lceil \frac{q}{2} \rceil, q\}$ maximizing $\sum\limits_{u \in V(G)} x(u)$ on the set $L(G) = \Big\{x : V(G) \to \{0\} \cup [q] \;\mid\; x(u) + x(v) \leq q \;\forall uv \in E(G)\Big\}$.*

Proof. At first we assume that there is a set of independent vertices $A \subset V(G)$ so that $|A| > |N(A)|$. We may pick a minimal such A, that is, $|B| \leq |N(B)|\ \forall B \subsetneq A$. By assumption A is nonempty.

Let $B = A \setminus \{v\}$ for some arbitrary $v \in A$. Then $|A| - 1 = |B| \leq |N(B)| \leq |N(A)| < |A|$ so $N(B) = N(A)$. By Hall's theorem there is a matching $M \subset E(G)$ from B to $N(A)$. For $x \in L(G)$ we get $\sum_{u \in A \cup N(A)} x(u) = x(v) + \sum_{u \in B \cup N(B)} x(u) \leq q + |B| \cdot q = q|A|$, since $x(u) + x(v) \leq q$ for $uv \in M$.

As $v \in A$ was arbitrary, equality holds above if and only if $x|_A \equiv q$ and $x|_{N(A)} \equiv 0$. Since there are no edges from A to $V(G) \setminus (A \cup N(A))$, any maximal $x \in L(G)$ is the union of $(A \times \{q\}) \cup (N(A) \times \{0\})$ and some maximal $x' \in L(G')$ where $G' = G \setminus (A \cup N(A))$.

By repeating the argument if G' has an independent $A' \subset V(G')$ with $|A'| > |N(A')|$, any maximal $x \in L(G)$ is the union of $(V_1 \times \{q\}) \cup (V_2 \times \{0\})$ and some maximal $x' \in L(G_1)$, where V_1 and V_2 are the disjoint subsets of $V(G)$ that we obtain by the argument, $G_1 = G \setminus (V_1 \cup V_2)$ and $|H| \leq |N(H)|$ for all independent $H \subset V(G_1)$.

Observe that $S = \{u \in V(G_1) \mid x(u) > \lfloor \frac{q}{2} \rfloor\}$ is an independent set of vertices in G_1 for $x \in L(G_1)$. By Hall's theorem there is a matching from S to some $T \subset N(S)$ in G_1, so we calculate

$$\sum_{u \in V(G_1)} x(u) \;=\; \sum_{u \in S \cup T} x(u) \;+\; \sum_{u \notin S \cup T} x(u) \leq q|S| + \left\lfloor \frac{q}{2} \right\rfloor \big(|V(G_1)| - 2|S|\big).$$

Note that this maximum is achieved by $x' \equiv \frac{q}{2}$ if q is even, and

$$x'(u) = \begin{cases} \lceil \frac{q}{2} \rceil & \text{if } u \in S \\ \lfloor \frac{q}{2} \rfloor & \text{if } u \notin S, \end{cases}$$

if q is odd; and this x' is in $L(G_1)$. So there is a maximal $x' \in L(G_1)$ taking values in $\big\{0, \lfloor \frac{q}{2} \rfloor, \lceil \frac{q}{2} \rceil, q\big\}$. The statement follows.

□

Although this lemma is valid for arbitrary values of q, we will only use this lemma when q is even, for which case it simplifies to the following:

Lemma 4.2. *Let $G = (V, E)$ be a simple graph, and let $x : V \to [0,1]$ be a function on the vertices of G such that it satisfies the linear program* $\max\{\boldsymbol{1} \cdot x : x(u) + x(v) \leq 1 \;\; \forall (u,v) \in E(G)\}$. *Then x can be chosen to have values from the set* $\{0, \frac{1}{2}, 1\}$.

Notice that this version of the problem was derived from the original by dividing every inequality with q, combined with the fact that for even values of q, $\lfloor \frac{q}{2} \rfloor = \lceil \frac{q}{2} \rceil = \frac{q}{2}$. Let us point out that this problem bears a significant resemblance to the vertex cover problem, for which there also exists a half-integer optimal solution. In fact, we give another proof of the lemma which is practically the same as the one for the vertex cover case.

Proof of Lemma 4.2. If G is bipartite, then the matrix describing the linear program is totally unimodular, and the right hand side of the inequality has integer entries, hence it has an optimal integer solution, which can only have values from $\{0, 1\}$.

If G is not bipartite, then let us consider the bipartite graph $H = (A, B, E)$ where $A = B = V(G)$, and $v_A \in A$ and $v_B \in B$ are the two vertices corresponding to $v \in V$; and $E = \{u_A v_B, u_B v_A : uv \in E(G)\}$. If x is a permitted vector for G with respect to the linear program, then it can be naturally expanded to a solution for H: let $y(u_A) = y(u_B) = x(u)$. This y is clearly permitted for H, and $\sum_{v \in A \cup B} y(v) = 2 \cdot \sum_{v \in V(G)} x(v)$, which yields $2 \cdot \mathrm{opt}_G \leq \mathrm{opt}_H$.

The first part of the proof gives a y for which $\sum_{v \in A \cup B} y(v) = \mathrm{opt}_H$ and $y(v) = 0$ or 1 for every $v \in V(H)$. We define the following x: if $y(u_A) = y(u_B) = 1$, then $x(u) = 1$. If either $y(u_A) = 1, y(u_B) = 0$, or $y(u_A) = 0, y(u_B) = 1$, then let $x(u) = \frac{1}{2}$; and let $x(u) = 0$ if both $y(u_A)$ and $y(u_B)$ are 0. It is easy to check that this x satisfies the conditions for G, and gives equality in $2 \cdot \mathrm{opt}_G \leq \mathrm{opt}_H$. Moreover, it has entries from $\{0, \frac{1}{2}, 1\}$, which concludes our proof. $\square$

We now generalize Lemma 4.1 to hypergraphs. There are many ways to do this, the one we discuss here is the case that will be useful for us in the setting of q-graphs.

Lemma 4.3. *Let $\mathcal{H} = (V, H)$ be a hypergraph, and let $x : V \to \{0\} \cup [q]$ be a function on the vertices of $\mathcal{H}$ such that it satisfies the following condition: $\forall h \in H\ \exists u, v \in V(h) : x(u) + x(v) \leq q$. Then an x that maximizes the expression $\mathbf{1} \cdot \mathbf{x}$ can be chosen to have values from the set $\{0, \lfloor \frac{q}{2} \rfloor, \lceil \frac{q}{2} \rceil, q\}$.*

Proof. Construct an ordinary graph G on the vertex set $V(\mathcal{H})$ as follows: choose a pair of vertices $\{u, v\}$ from each hyperedge $h \in H$ (this pair will guarantee the sum condition of x for h), and add the edge (u, v) to G. By applying Lemma 4.1, we can set $x_G = \arg\min\{\mathbf{1} \cdot x : x(u) + x(v) \leq q, \forall uv \in E(G)\}$ to have values from $\{0, \lfloor \frac{q}{2} \rfloor, \lceil \frac{q}{2} \rceil, q\}$. It is easy to see that if we take the solution x for which $\mathbf{1} \cdot x = \min_G\{\mathbf{1} \cdot x_G\}$ over every possible choice of G, we get an optimal solution for the original problem for $\mathcal{H}$. $\square$

Again, an important note to make is that the interesting case for us will be when q is even, for which the optimal solution will have values from $\{0, \frac{q}{2}, q\}$.

The key feature of this thesis simply states that for every graph F and even q, it suffices to examine the $q = 2$ case. This has a long-reaching impact, as combining it with the results of Patkós et al. significantly narrows down the unknown values of $\mathrm{ex}(n, F, q)$, at least when q is even.

Theorem 4.4 (Encz, Marits, Váli, Weisz)**.** *For every even q and ordinary graph F, $ex(n, F, q) = \frac{q^2}{4} \cdot \mathrm{ex}(n, F, 2)$.*

The proof is comprised of a somewhat technical part where we explain how Lemma 4.3 provides an optimal q-graph with a special structure; and a part where we exploit that structure to connect the general setup to the $q = 2$ case.

Proof of Theorem 4.4. Consider $H \subseteq \mathcal{Q}(n, 2)$ without a $(q + 1)$-copy of F with $\mathrm{ex}(n, F, q)$ q-edges. Let v be an arbitrary vertex of H. For another vertex $u \neq v$ and $i \in [q]$, let $m(u, i) = \max_{r\in[q]}\{(u, v, i, r) \in H\}$ (if $\{r \in [q] : (u, v, i, r) \in H\} = \emptyset$, then let $m_{u,i} = 0$). We intend to alter the q-edges adjacent to v in a way that every $m(u, i)$ will become $m'(u, i) \in \{0, \frac{q}{2}, q\}$, a $q + 1$-copy of H does not appear; and in the meantime, the total number of q-edges does not decrease.

For that purpose, let $x_{u,i}$ be a variable reflecting the current value of $m(u, i)$, and let $x = \{x_{u,i} : i \in [q], u \in [n]\}$ be the set of all such variables. For $t = \{t_{u,i} : t_{u,i} \in [q] \cup \{0\}\ \forall u \in [n],\ \forall i \in [q]\}$, let $x(t)$ be the evaluation x according to t; that is, setting $x_{u,i} = t_{u,i}\ \forall u \in [n],\ \forall i \in [q]$. For an evaluation of x according to t, let $H^{x(t)}$ be the q-graph we obtain from H by changing the q-edge labels at node v according to the values of $x_{u,i} = t_{u,i}$. More precisely, H and $H^{x(t)}$ are identical on $[n] - \{v\}$, and $\max_{r\in[q]\cup\{0\}}\{(u, v, i, r) \in H^{x(t)}\} = t_{u,i}$. We call an evaluation $x(t)$ of x admissible, if the corresponding q-graph $H^{x(t)}$ does not admit a $(q+1)$-copy of F.

Take an arbitrary evaluation of x with values from $\{0, 1, \ldots, q\}$, and consider the q-graph $H^{x(t)}$. Suppose that $x(t)$ is not admissible, let $F^{x(t)} \subseteq H^{x(t)}$ be a $(q+1)$-copy of F. Since H did not contain a $(q + 1)$-copy of F, and we only changed labels at vertex v, $F^{x(t)}$ must contain q-edges attached to v; let those q-edges be

$$L^{x(t)} = \{(u_k, v, r_k, x_{u_k,r_k} = t_{u_k,r_k}) : u_k \in S_{H^{x(t)}} - \{v\},\ r_k \in [q],\ k = 1, 2, \ldots\},$$

and let

$$U^{x(t)} = \{\{u_k, r_k\} : (u_k, v, r_k, x_{u_k,r_k} = t_{u_k,r_k}) \in L^{x(t)}\}.$$

We gather that if an evaluation of x according to t' is admissible, at least one of the following inequalities must hold:

$$\{t'_{u_i,r_i} + t'_{u_j,r_j} \leq q : \{u_i, r_i\} \in U^{x(t)}, \{u_j, r_j\} \in U^{x(t)}, i \neq j\},$$

or else the q-graph

$$\{(u_k, v, r_k, x_{u_k,r_k} = t'_{u_k,r_k}) : (u_k, v, r_k, x_{u_k,r_k} = t_{u_k,r_k}) \in L^{x(t)}\}$$

would be a $(q+1)$-copy of F. Consequently, if we take the union of these conditions for every non-admissible evaluation $x(t)$, we can characterize when $x(t')$ is an admissible evaluation:

Claim 4.5. *$x(t')$ is admissible if and only if satisfies at least one inequality from*

$$\{t'_{u_i,r_i} + t'_{u_j,r_j} \leq q : \{u_i, r_i\} \in U^{x(t)}, \{u_j, r_j\} \in U^{x(t)}, i \neq j\}$$

for every non-admissible evaluation $x(t)$.

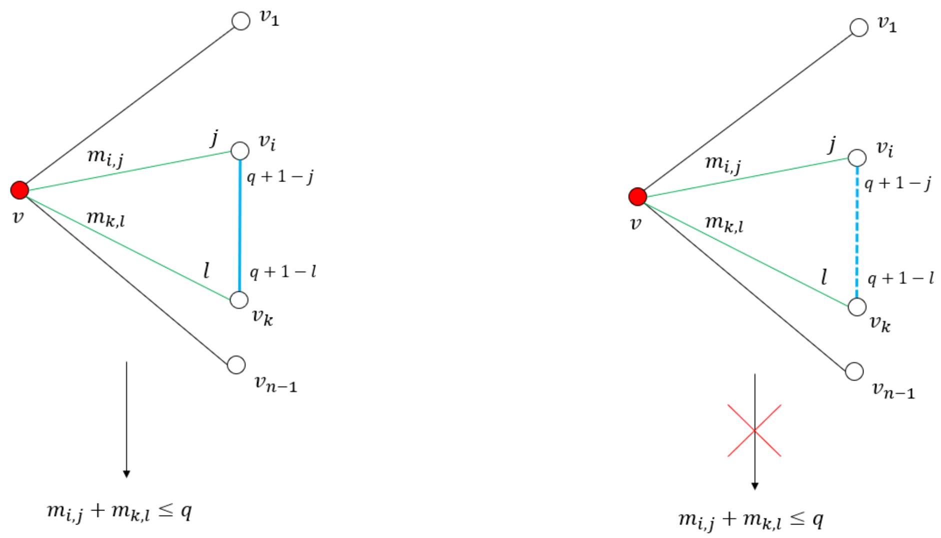


Figure 12: $F = C_3$ case. The presence of $(v_i, v_j, q+1-j, q+1-l)$ in H implies $m_{i,j} + m_{k,l} \leq q$ for an admissible evaluation $x(m)$.

In the following part of the proof, we construct a hypergraph $\mathcal{H}$, in which we encode our previously acquired knowledge of the q-edges with common endpoint v. The nodes of the hypergraph will pertain to the variables $x_{u,i}$, and each hyperedge represents a variable set obtained from $U^{x(t)}$ for a non-admissible evaluation $x(t)$. Claim 4.5 indicates that for each such hyperedge, at least two variables must have a sum at most q in order to represent an admissible evaluation, which demonstrates why Lemma 4.3 is applicable.

Let $V(\mathcal{H}) = \{w_{u,i} : u \in [n], i \in [q]\}$, and

$$E(\mathcal{H}) = \bigcup_{\{\{u_1,r_1\},\ldots,\{u_j,r_j\}\}=U^{x(t)}, x(t) \text{ is non-admissable}} \{w_{u_1,r_1}, w_{u_2,r_2}, \ldots, w_{u_j,r_j}\}.$$

Now $x = \{x_{u,i} : u \in [n],\ i \in [q]\}$ is a function on $V(\mathcal{H})$, and an evaluation $x(t')$ of x gives an admissible evaluation if and only if $\forall h \in E(\mathcal{H})\ \exists w_{u_i,r_i}, w_{u_j,r_j} \in V(h) : t'_{u_i,r_i} + t'_{u_j,r_j} \leq q$ by Claim 4.5. Thus, the hypergraph $\mathcal{H}$ and the function x satisfy the conditions of Lemma 4.3; so, bearing in mind that now $\lfloor \frac{q}{2} \rfloor = \lceil \frac{q}{2} \rceil = \frac{q}{2}$, we can find an admissible evaluation of x with values $0, \frac{q}{2}$ or q which maximizes the target function. Let us fix this evaluation to be $x(t')$. In the meantime, we can alter the q-edges adjacent to v according to the evaluation of x so that $\max_{r \in [q] \cup \{0\}} \{(u, v, i, r) \in H^{x(t')}\}$ becomes $m'(u, i) \in \{0, \frac{q}{2}, q\}$ for every $u \in [n]$ and $i \in [q]$. Meanwhile, the number of q-edges attached to v does not decrease:

$$\sum_{u \in [n]} \sum_{i \in [q]} m(u, i) \leq \max_{x(t) \text{ is admissible}} \mathbf{1} \cdot x(t) = \mathbf{1} \cdot x(t') = \sum_{u \in [n]} \sum_{i \in [q]} m'(u, i),$$

since initially H did not contain a $(q + 1)$-copy of, meaning that $x_{u,i} = m(u, i)$ was an admissible evaluation.

By iterating the above modification for every vertex v in $[n]$, we end up with a q-graph H' that has the following property: $\forall (u, v, a, b) \in H'$,

$$\max_{r \in [q] \cup \{0\}} \{(u, v, r, b) \in H'\} \in \left\{0, \frac{q}{2}, q\right\}, \max_{r \in [q] \cup \{0\}} \{(u, v, a, r) \in H'\} \in \left\{0, \frac{q}{2}, q\right\}.$$

Indeed, suppose that we have already processed the q-edges adjacent to v, meaning that for the current q-graph H, it holds that for every $u \in [n]$ and $a \in [q]$: $\max_{r \in [q] \cup \{0\}} \{(u, v, a, r)\} \in \{0, \frac{q}{2}, q\}$. When we arrive at processing the node u, we may replace the label a of a q-edge (u, v, a, b), but the label b at node v remains the same. This implies that $\max_{r \in [q] \cup \{0\}} \{(u, v, a, r)\} \in \{0, \frac{q}{2}, q\}$ remains true for every $u \in [n]$ and $a \in [q]$.

Note that the modified H' is still optimal, so if a q-edge (u, v, a, b) is in H', then so is every other (u, v, a', b') with $a' \leq a$, $b' \leq b$. With this remark, the special structure of H' can be rephrased in the following way: for every support $\{u, v\}$, consider the following partition of potential q-edges:

- $E_{s,s} = \{(u, v, a, b) : 1 \leq a \leq \frac{q}{2}, 1 \leq b \leq \frac{q}{2}\}$, $|E_{s,s}| = \frac{q^2}{4}$
- $E_{b,s} = \{(u, v, a, b) : \frac{q}{2} < a \leq q, 1 \leq b \leq \frac{q}{2}\}$, $|E_{b,s}| = \frac{q^2}{4}$
- $E_{s,b} = \{(u, v, a, b) : 1 \leq a \leq \frac{q}{2}, \frac{q}{2} < b \leq q\}$, $|E_{s,b}| = \frac{q^2}{4}$
- $E_{b,b} = \{(u, v, a, b) : \frac{q}{2} < a \leq q, \frac{q}{2} < b \leq q\}$, $|E_{b,b}| = \frac{q^2}{4}$.

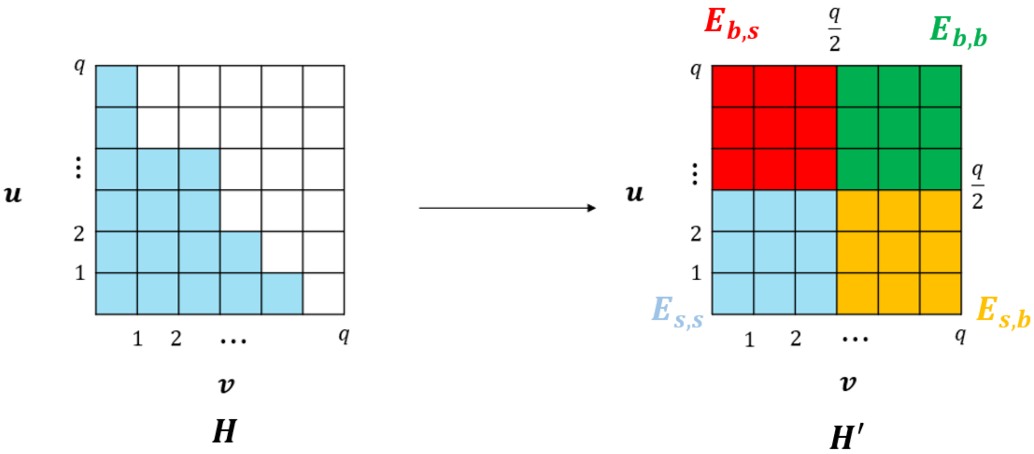


Figure 13: The intersection of the i-th row and j-th column represents the q-edge $(u, v, i, j) \in H$. q-edges of the same color are either all in H', or all not in H'.

We may observe that if there is a q-edge (u, v, a, b) from $E_{s,s}$ in H', then $E_{s,s} \subseteq H'$ must hold; and a similar statement is true for $E_{s,b}, E_{b,s}$ and $E_{b,b}$. For each support pair $\{u, v\}$, let us identify these four q-edge sets with the $(u, v, 1, 1)$, $(u, v, 2, 1)$, $(u, v, 1, 2)$ and $(u, v, 2, 2)$ 2-edges of a 2-graph H'', and consider the usual partition of the 2-edges in H'' by the label pairs:

- $e_{s,s} = |(u, v) \in [n]^2 : (u, v, 1, 1) \in H''|$
- $e_{s,b} = |(u, v) \in [n]^2 : (u, v, 1, 2) \in H''|$
- $e_{b,s} = |(u, v) \in [n]^2 : (u, v, 2, 1) \in H''|$

- $e_{b,b} = |(u,v) \in [n]^2 : (u,v,2,2) \in H''|$

Then

$$\mathrm{ex}(n,F,q) = |H'| = |E_{s,s}| \cdot e_{s,s} + |E_{s,b}| \cdot e_{s,b} + |E_{b,s}| \cdot e_{b,s} + |E_{b,b}| \cdot e_{b,b} = \frac{q^2}{4} \cdot |H''|.$$

H'' cannot contain a 3-copy of F, because H' did not contain a $(q+1)$-copy of F, so $|H''| \leq \mathrm{ex}(n,F,2)$, and

$$\mathrm{ex}(n,F,q) \leq \frac{q^2}{4} \cdot \mathrm{ex}(n,F,2).$$

For the other direction, consider a 2-graph $H \in \mathrm{EX}(n,F,2)$. One only needs to reverse the above construction: substituting the edges $(u,v,1,1)$, $(u,v,1,2)$, $(u,v,2,1)$, and, $(u,v,2,2)$ in H with the edge sets $E_{s,s}, E_{s,b}, E_{b,s}, E_{b,b}$ respectively gives a q-graph H' with $|H'| = \frac{q^2}{4} \cdot |H| = \frac{q^2}{4} \cdot$ ex$(n,F,2)$. H' does not contain a $(q+1)$-copy of F, so $\mathrm{ex}(n,F,q) \geq |H'| = \frac{q^2}{4} \cdot$ ex$(n,F,2)$.

□

For odd values of q, there is no easy way to interpret a mapping of $\{0, \lfloor \frac{q}{2} \rfloor, \lceil \frac{q}{2} \rceil, q\}$ to the values $\{0,1,2\}$ the same way as in the previous reasoning. The proof of Theorem 4.4 strongly relies on the parity of q, so in the general case, when q is not necessarily even, the same method will not suffice. As for now, we must settle for an upper bound when q is odd:

Proposition 4.6. *$ex(n,F,q) \leq \frac{q^2}{4} \cdot ex(n,F,2)$, $\forall q \in \mathbb{N}$*

Proof. Consider a q-graph $H \in \mathrm{EX}(n,F,q)$ and define H' as

$$H' = \{(u,v,2a,2b), (u,v,2a-1,2b), (u,v,2a,2b-1), (u,v,2a-1,2b-1) : (u,v,a,b) \in H\}.$$

In the obtained $2q$-graph H', a $(2q+1)$-copy of F does not appear: the largest value of s for which an s-copy of F is present in H is at most q, so the largest value of s for which an s-copy of F is present in H' is at most $2q$. Hence, by Theorem 4.4,

$$|H'| = 4 \cdot |H| = 4 \cdot \mathrm{ex}(n,F,q) \leq \mathrm{ex}(n,F,2q) = q^2 \cdot \mathrm{ex}(n,F,2).$$

□

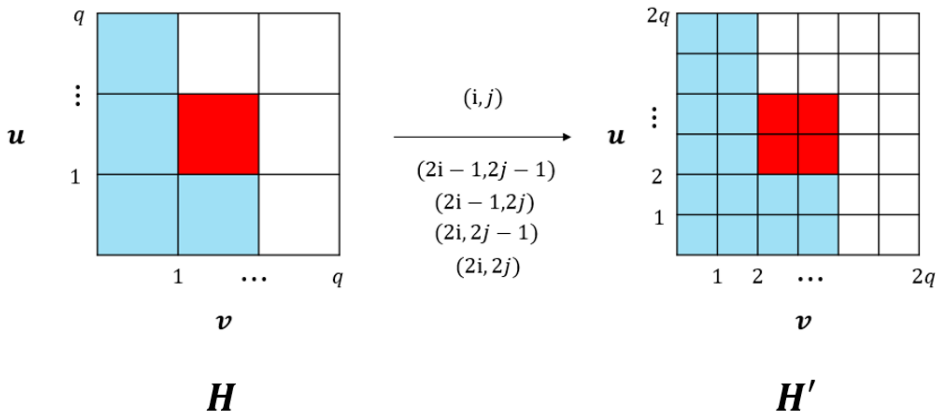


Figure 14: The reduction to $q' = 2q$ from an odd q.

The thorough examination of the q-edges $\{(\lceil \frac{q}{2} \rceil, a) : a \in [q]\}$ and $\{(\lfloor \frac{q}{2} \rfloor, a) : a \in [q]\}$ might provide better answers than Proposition 4.6, as the proof consists of a simple reduction from q to $2q$, and does not use the underlying structure of the q-graph.

By itself, Proposition 4.6 does not carry a huge significance, as it only provides an upper bound for $\mathrm{ex}(n, F, q)$, but in some special cases it coincides with the trivial lower bound $q^2 \cdot \mathrm{ex}(n, F)$, hence giving the exact value of $\mathrm{ex}(n, F, q)$; as is the case with C_3. We present a generalization of Theorem 3.14.

Proposition 4.7. $\mathrm{ex}(n, C_3, q) = q^2 \cdot \lfloor \frac{n^2}{4} \rfloor$, $\forall q \in \mathbb{N}$.

Proof. The statement easily follows from the combination of two previous propositions: On one hand, we know from Proposition 2.18 that $\mathrm{ex}(n, C_3, q) \geq q^2 \cdot \mathrm{ex}(n, C_3) = q^2 \lfloor \frac{n^2}{4} \rfloor$; and on the other hand, $\mathrm{ex}(n, C_3, q) \leq \frac{q^2}{4} \cdot \mathrm{ex}(n, C_3, 2) = q^2 \cdot \lfloor \frac{n^2}{4} \rfloor$ comes from Proposition 4.6. □

Finally, we exploit Proposition 4.4 in order to strengthen previous theorems for $q = 2k$.

Theorem 4.8. *Suppose T is a tree of radius r, and q is even.*

(1) If the diameter of T is $2r$, then $ex(n, T, q) = \left(\frac{q^2}{4} + o(1)\right) \cdot \left(\binom{n}{2} + t_{n,r}\right)$.

(2) If the diameter of T is $2r-1$, then $ex(n,T,q) = \left(\frac{q^2}{4}+o(1)\right)\cdot\left(\binom{n}{2}+t'_{n,r}-\binom{\lfloor\frac{n}{2r-1}\rfloor}{2}\right)$.

Since we determined the asymptotic value of $\mathrm{ex}(n, C_{2k+1}, 2)$, we can assert the conjecture of Patkós et al. that $\mathrm{ex}(n, C_{2k+1}, q)$ is asymptotically $q^2 \cdot \mathrm{ex}(n, C_{2k+1})$.

Proposition 4.9. *For every $q \geq 2$, $\mathrm{ex}(n, C_{2k+1}, q) = \frac{n^2}{4} \cdot q^2 + o(n^2)$.*

Proof. The proof is the same as for C_3; we only need to compare the upper bound provided by Proposition 4.6 with the trivial lower bound in Proposition 2.18. It follows that

$$q^2 \cdot \left\lfloor \frac{n^2}{4} \right\rfloor = q^2 \cdot \mathrm{ex}(n, C_{2k+1}) \leq \mathrm{ex}(n, C_{2k+1}, q) \leq \frac{q^2}{4} \cdot (n^2 + o(n^2)) = q^2 \cdot \frac{n^2}{4} + o(n^2).$$

□

As the concluding part of this chapter, let us combine Proposition 4.6 with the monotonicity of $\mathrm{ex}(n, F, q)$ in q from Proposition 2.15:

Proposition 4.10. *For every graph F and $q \geq 2$,*

$$\frac{(q-1)^2}{4} \cdot ex(n, F, 2) \leq ex(n, F, q) \leq \frac{q^2}{4} \cdot ex(n, F, 2)$$

Proof. The second inequality is simply Proposition 4.6. If q is even, then $\mathrm{ex}(n, F, q)$ equals to the right hand side, and if q is odd, then $q-1$ is even, so Proposition 4.4 is applicable: $\frac{(q-1)^2}{4} \cdot \mathrm{ex}(n, F, 2) = \mathrm{ex}(n, F, q-1) \leq \mathrm{ex}(n, F, q)$. □

The above proposition limits $\mathrm{ex}(n, F, q)$ to an interval of size $\frac{2q-1}{4} \cdot \mathrm{ex}(n, F, 2)$. When q is odd, we feel that this boundary can be improved both ways, as the proof does not take into account the specific attributes of the q-graph. An equality in either side would entail that an optimal construction for $q-1$ or $2q$ is simultaneously the best one can achieve for q. We suspect this is not the case, and there is some room for improvement.

5 Concluding remarks

We conclude this paper by highlighting some questions that can be articulated in relation to q-graphs. As a consequence of our results, one may immediately transfer every statement for $q = 2$ to every even q. Moreover, an upper and a lower bound is established for the general case as well. However, we need to emphasize that these constraints may not provide a precise value of the Turán number for every graph F, as we suspect is the case for trees. Consequently, the exact (or asymptotic) value of $\mathrm{ex}(n, F, q)$ remains unknown for odd values of q. A potential next step could be to improve our trivial bounds, or prove that one of the bounds coincides with $\mathrm{ex}(n, F, q)$.

On a general note, let us highlight that our main achievement was the reduction of even q values. Apart from providing an asymptotical answer for $\mathrm{ex}(n, C_{2k+1}, 2)$, we did not contribute to solving any other questions imposed by Patkós, Tuza and Vizer. Their conjectures and questions remain open, the seemingly most attainable among them being the case of forests.

Bibliography

[1] Andrásfai, B., Erdős, P., T. Sós, V., On the connection between chromatic number, maximal clique and minimal degree of a graph. *Disc. Math.* **8** (1974), 205-218.

[2] Encz, K., Marits, M., Váli, B., Weisz, M., Results on extremal graph theoretic questions for q-ary vectors. *Proceedings of 12th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications* (2023), 155-164.

[3] Erdos, P., Simonovits, M., A limit theorem in graph theory. *Studia Sci. Math. Hungar.* **1** (1966), 51-57.

[4] Erdős, P., Stone, A. H., On the structure of linear graphs. *Bulletin of the American Mathematical Society* **52** (1946), 1087–1091.

[5] Füredi, Z., Simonovits, M., The history of degenerate (bipartite) extremal graph problems. Erdös centennial, 169-224, *Bolyai Soc. Math. Stud.*, **25** (2013), János Bolyai Math. Soc., Budapest.

[6] Keevash, P., Hypergraph Turán Problems. *Surveys in combinatorics* **392** (2011), 83-140.

[7] Keevash, P., Mubayi, D., Sudakov, B., Verstraëte, J., Rainbow Turán problems. *Combin. Probab. Comput.* **16** (2007), 109-126.

[8] Keevash, P., Sudakov, B., The Turán number of the Fano plane. *Combinatorica* **25** (2005), 561-574.

[9] Li, B., Zhou, W., The Turán problems of directed paths and cycles in digraphs, arXiv:2102.10529v1 [math.CO].

[10] Patkós, B., Tuza, Zs., Vizer, M. (2022), Extremal graph theoretic questions for q-ary vectors. arXiv:2305.01919v1 [math.CO].

[11] Simonovits, M., A method for solving extremal problems in graph theory, stability problems. *Theory of Graphs* (Proc. Colloq. Tihany, 1966), Academic Press, New York, and Akad. Kiadó, Budapest, 1968, 279–319.

[12] Turán, P., On an extremal problem in graph theory. *Matematikai és Fizikai Lapok* (in Hungarian). **48** (1941), 436-452.

[13] Turán, P. On the theory of graphs. *Colloq. Math.* **3** (1954), 19-30.