Generalizations of Turán-type problems

Bartłomiej Kielak

A thesis presented for the degree of
Doctor of Philosophy

Advisor
dr hab. Bartosz Walczak

Co-advisor
dr Andrzej Grzesik

Jagiellonian University
Kraków, Poland
2023

Abstract

In 1941, Turán determined the maximum number of edges which a graph on a fixed number of vertices can contain without forcing a complete graph of a given order. Since then, many so-called Turán-type questions have been raised about maximizing the value of some graph parameter, like the number of subgraphs isomorphic to a particular graph, in a graph with some forbidden substructure. We contribute to the study by proving several results on different Turán-type problems.

First, we contribute to the problem of maximizing the number of k-cycles in a graph without ℓ-cycles by finding the exact asymptotics of this number when $k \geq 7$ is odd and $\ell = k - 2$. This is an extension of a long-standing problem of Erdős about the maximum number of pentagons in a graph without triangles, which was resolved only recently.

Next, we study the problem of maximizing the number of arcs in an oriented graph which does not contain a specific oriented graph as a subgraph. The exact asymptotics is linked to a certain graph parameter called compressibility. We prove several results regarding the growth rate of compressibility with respect to the length of a longest path. In particular, we prove that if the maximum out-degree of an acyclic oriented graph is at most two, then its compressibility is bounded from above by a polynomial of degree four.

We also investigate the problem of maximizing the number of directed k-cycles in an oriented graph without directed ℓ-cycles. We determine the order of magnitude for all choices of k and ℓ, and the exact asymptotics for a number of cases.

Finally, we consider the problem of maximizing the number of induced subgraphs isomorphic to some fixed graph, both in the undirected and the oriented settings. This leads to a graph parameter called inducibility. We contribute to the study by improving the lower bound on the inducibility of a 4-vertex path, and by determining or approximating the inducibility for all oriented graphs on four vertices.

Acknowledgements

First of all, I would like to express my deepest gratitute to my co-advisor Andrzej Grzesik, who introduced me to the field of Extremal Graph Theory and offered me constant support throughout my whole PhD studies. Working with you was a real pleasure and it kept me motivated and in high spirits despite all my struggles and doubts.

I would also like to thank all other people who directly influenced the contents of this thesis — my advisor Bartosz Walczak, who suggested a great deal of impactful improvements to the text, and my other co-authors: Łukasz Bożyk, Justyna Jaworska, Piotr Kuc, Aliaksandra Novik, Tomasz Ślusarczyk, and Radosław Żak, for the countless hours spent together on solving extremal problems and on inspiring discussions.

Special thanks to my spiritual guardians, Dan Kral' and Michał Wojciechowski — for your kindness, words of wisdom, and sharing with me your taste for good mathematics.

Finally, I would have never made it this far without my family and friends. Thank you for all the time spent together on holiday or escape room trips, rpg sessions, board games meetings, bouldering, and many more activities that kept me sane and happy, and gave me a reason to keep going.

Contents

Chapter 1

Introduction

Extremal Graph Theory is a field of combinatorics that studies how various graph properties behave with respect to different global graph parameters. A typical problem structure is the following — for a given family $\mathcal{F}$ of graphs and a graph parameter f, what is the maximum value α of $f(G)$ over all graphs $G \in \mathcal{F}$? It is also of interest to determine all $G \in \mathcal{F}$ such that $f(G) = \alpha$ — they are usually referred to as *extremal graphs.*

One of the most basic and fundamental graph parameters is the number of edges in a graph. It was already considered by Euler [30] in 1758, who proved that a planar graph on n vertices can have at most $3n - 6$ edges. Much later, in 1907, Mantel [64] observed that any graph on n vertices without triangles can contain at most $n^2/4$ edges, and Erdős [24] in 1938 proved an upper bound on the number of edges in a graph without 4-cycles. However, the systematic study of this parameter started with the work of Turán [75] in 1941 who determined the maximum number of edges in a graph without a complete graph of a fixed order.

For a graph H and a natural number n, the *Turán number* of H, denoted by $\mathrm{ex}(n, H)$, is the maximum number of edges in a graph on n vertices which is H-free, i.e. which does not have a subgraph isomorphic to H. In this context, Mantel determined the Turán number of a triangle, and Turán determined the Turán number of a general complete graph. But it was Erdős and Stone who proved in 1946 an asymptotic formula for any graph H:

Theorem 1.1 (Erdős, Stone [29]). *For any graph H with $\chi(H) \geq 2$,*

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

For bipartite graphs H, from Theorem 1.1 we can only conclude that $\mathrm{ex}(n, H) = o(n^2)$. Finding a more precise asymptotics is a challenging task and it was resolved only for a narrow class of bipartite graphs. There are many partial results regarding the Turán number of complete bipartite graphs, even cycles or d-degenerate graphs; for more details on the topic see for instance the survey [37]. We shall instead focus on another directions of research and try to investigate related problems.

There are many possible variations of the Turán problem, which are usually referred to as the Turán-type problems. Let us list those few of them which will be considered in this thesis. Basic notation and tools are introduced in Chapter 2.

- Instead of counting edges, one can fix a graph T and count the number of subgraphs isomorphic to T; this leads to the concept of a *generalized Turán number* $\mathrm{ex}(n, T, H)$, which we shall discuss in more detail in Chapter 3. Of our particular interest is when both T and H are cycles. This problem was considered already by Erdős [25], who conjectured that $\mathrm{ex}(n, C_5, C_3) \leq (n/5)^5$ and the equality holds only for a balanced blow-up of C_5; this conjecture was proved with a sophisticated use of the flag algebra method and stability arguments [41, 49, 60]. Up to now, the exact asymptotics of $\mathrm{ex}(n, C_k, C_l)$ was determined only for specific pairs of k and l. We contribute to the study by determining asymptotically $\mathrm{ex}(n, C_{2k+1}, C_{2k-1})$ for $k \geq 3$ (Theorem 3.1). In contrast to the proofs of Erdős' conjecture, the proof of Theorem 3.1 is not computer assisted. Instead, it relies on the method developed by Kráľ, Norin, and Volec [59] and on Regularity Lemma.

- One can consider other classes of discrete structures. In Chapter 4, we shall discuss Turán-type problems for directed and oriented graphs. For the latter, there exists a parameter $\tau(H)$ called ***compressibility***, which plays the same role as the chromatic number in Theorem 1.1. Therefore, to determine the exact asymptotics of $\mathrm{ex}_o(n, H)$ for an acyclic oriented graph H, it is enough to compute $\tau(H)$. (If H is not acyclic, then the problem is trivial.) We prove several results regarding the growth rate of compressibility with respect to the order $p(H)$ of a longest path in H. In particular, we show that $\tau(H) = O(p(H)^4)$ if the maximum out-degree of H is two (Theorem 4.14). There is a general lower bound $p(H) \leq \tau(H)$, and we provide a large family of acyclic oriented graphs for which this bound is tight (Theorem 4.27).

- One can mix both concepts. In Chapter 5, we shall consider generalized Turán numbers $\mathrm{ex}_o(n, T, H)$ for oriented graphs, with the focus on the case when both $T = \overrightarrow{C_k}$ and $H = \overrightarrow{C_\ell}$ are directed cycles. We determine the order of magnitude of $\mathrm{ex}_o(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$ for all choices of k and ℓ (Theorem 5.2). We also find the exact asymptotics for $k \in \{3, 4, 5\}$ and $\ell > k$ not divisible by k (Theorems 5.5, 5.6, and 5.7), for odd $k > 6$ and sufficiently large ℓ not divisible by k (Theorem 5.4), and for $(k, \ell) = (3, 6)$ (Theorem 5.3). The proofs combine many different tools and methods, including Regularity Lemma, the flag algebra method, and Frobenius coin problem.

- What if we try to maximize not the number of subgraphs isomorphic to some graph, but the number of *induced* subgraphs? This results in a graph parameter called ***inducibility***, which measures approximately how many induced subgraphs isomorphic to a fixed graph H can appear in a graph on n vertices. Our understanding of this parameter is currently rather low, as it was determined only for a very narrow class of graphs. The smallest graph whose inducibility is still unknown is P_4, a path on 4 vertices. Over the years, there were many attempts to bound the inducibility of P_4 [32, 51, 77, 31]. We found a construction, which is presented in Chapter 6, yielding an improvement to the best previously known lower bound for the inducibility of P_4 obtained in [31].

- As in the case of the original Turán problem, inducibility can also be considered for other classes of combinatorial structures. In Chapter 7, we shall discuss the results on inducibility of oriented graphs. We contribute to the problem by determining inducibility exactly or finding very good estimates for all oriented graphs on four vertices. Most of the proofs of the upper bounds rely on the flag algebra method and are therefore computer assisted. We also show that for almost every oriented graph on 4 vertices, there exists a construction yielding a strictly better lower bound on inducibility than the general iterated blow-up construction, which is almost always optimal in the case of larger graphs, as shown by Fox, Huang, and Lee [35].

Chapter 2

Notation and basic tools

In this chapter, we shall introduce the basic notation and some of the tools used in the proofs. Section 2.5.6 is Appendix A from [14] with no substantial changes. The remaining content of this chapter was written by me for the purpose of this thesis.

If A is a set and k is a nonnegative integer, we write $|A|$ for the cardinality of A and $\binom{A}{k}$ for the family of all k-subsets of A. For an integer $k \geq 0$, let $[k] = \{i \in \mathbb{Z} : 1 \leq i \leq k\}$. For functions, we write

- $f(n) = O(g(n))$ if $|f(n)| \leq Cg(n)$ for some constant $C > 0$,
- $f(n) = o(g(n))$ if $\lim\limits_{n\to\infty} f(n)/g(n) = 0$,
- $f(n) = \Theta(g(n))$ if $cg(n) \leq f(n) \leq Cg(n)$ for some constants $c, C > 0$.

2.1 Undirected graphs

A ***graph*** is a pair $G = (V(G), E(G))$, where $V(G)$ is a set of vertices and $E(G) \subseteq \binom{V(G)}{2}$ is a set of edges; if $v, w \in V(G)$ are vertices, we shall usually write vw for an edge $\{v, w\}$ and say that vw is ***incident*** to v and w. A graph is of ***order*** n if its vertex set is of cardinality n. We say that H is a ***subgraph*** of G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. A subgraph H of G is ***induced*** if for every $vw \in E(G)$ with $v, w \in V(H)$ we have $vw \in E(H)$. For a subset $X \subseteq V(G)$, let $G[X]$ denote the subgraph of G induced by the vertex set X, i.e. $V(G[X]) = X$ and $E(G[X]) = \{vw \in E(G) : v, w \in X\}$. For $v \in V(G)$, we shall also write $G - v$ for a subgraph of G induced by $V(G) \setminus \{v\}$.

For $v \in V(G)$, the ***neighborhood*** of v in G is defined as $N(v) = \{w \in V(G) : vw \in E(G)\}$ and the ***degree*** of v in G is $d(v) = |N(v)|$. The ***maximum degree*** $\Delta(G)$ of a graph G is the maximum of $d(v)$ taken over all vertices of G, and the ***minimum degree*** $\delta(G)$ is the minimum of $d(v)$ taken over all vertices of G. A vertex $v \in V(G)$ is ***isolated*** if its degree in G is zero.

For any graphs G and H, a map $f : V(G) \to V(H)$ is a ***homomorphism*** if it preserves edges, i.e. for any $vw \in E(G)$ we have $f(v)f(w) \in E(H)$; in this case, we usually write $G \to H$ to indicate that there exists some homomorphism between them. We say that a homomorphism $f : G \to H$ is ***surjective*** (***injective***) if the map $f : V(G) \to V(H)$ is surjective (injective). A graph H is a ***homomorphic image*** of G if there exists a surjective homomorphism $f : G \to H$ such that for each edge $vw \in E(H)$ there exists an edge $v'w' \in E(G)$ such that $vw = f(v')f(w')$. Two graphs G and H are ***isomorphic***, which we write shortly as $G \approx H$, if there exists a bijective map $f : V(G) \to V(H)$ such that both f and its inverse are homomorphisms. If H' is a subgraph of G and H is isomorphic to H', we shall refer to H' as to a ***copy*** of H in G. If G does not have a copy of H, we say that G is H-***free***. For a family $\mathcal{F}$ of graphs, we say that G is $\mathcal{F}$-free if it is H-free for every $H \in \mathcal{F}$.

We define the following special graphs:

- a *complete graph* K_n on $n \geq 1$ vertices with edge set $E(K_n) = \binom{V(K_n)}{2}$;
- a *path* P_n, also called an *n-path*, on $n \geq 1$ vertices $v_1, \dots, v_n$ with edge set $E(P_n) = \{v_i v_{i+1} : 1 \leq i < n\}$. Vertices v_1 and v_n are *endpoints* of P_n, while vertices $v_2, \dots, v_{n-1}$ are *internal vertices* of P_n. We say that a path P *joins* vertices v and w (or *goes* from v to w) if these vertices are the endpoints of P. We also define the *length* of a path as the number of its edges;
- a *cycle* C_n, also called an *n-cycle*, on $n \geq 3$ vertices $v_1, \dots, v_n$ with edges $v_1 v_n$ and $v_i v_{i+1}$ for $1 \leq i < n$; We define the *length* of a cycle as the number of its edges;
- a *complete bipartite graph* $K_{n,m}$ on $n+m$ vertices, where $m, n \geq 1$, with vertex set $V(K_{n,m}) = A \cup B$, where $|A| = n$ and $|B| = m$, and edge set $E(K_{n,m}) = \{vw : v \in A \text{ and } w \in B\}$. For $n \geq 2$, we shall also write $K_{1,n-1}$ as S_n and call it a *star* on n vertices;
- an *empty graph* I_n on $n \geq 1$ vertices with empty edge set.

Let v and w be two vertices of a graph G. We define the *distance* $d_G(v, w)$ from v to w in G as the minimum length of a path in G from v to w; if no such path exists, then we put $d_G(v, w) = \infty$. A graph G is *connected* if $d_G(v, w) < \infty$ for every two vertices $v, w \in V(G)$. A *connected component* of G is a connected subgraph of G which is not contained in any larger connected subgraph of H.

2.2 Directed and oriented graphs

A *directed graph* (also called a *digraph*) is a pair $G = (V(G), E(G))$, where $V(G)$ is a set of vertices and $E(G) \subseteq V(G) \times V(G)$ is a set of arcs; if $v, w \in V(G)$ are vertices, we shall usually write vw for an arc (v, w), and say that this arc is *(directed) from v to w*. A directed graph is of *order* n if its vertex set is of cardinality n. In this thesis, we consider only directed graphs without loops, i.e. arcs with both endpoints at the same vertex. An *oriented graph* is a directed graph such that for every pair of vertices there is at most one arc between them. For a vertex $v \in V(G)$, the *out-neighborhood* of v in G is the set $N^+(v) = \{w \in V(G) : vw \in E(G)\}$ and the *in-neighborhood* of v in G is the set $N^-(v) = \{w \in V(G) : wv \in E(G)\}$. We define also the *out-degree* of v in G as $d^+(v) = |N^+(v)|$, *in-degree* of v in G as $d^-(v) = |N^-(v)|$, *degree* of v as $d(v) = d^+(v) + d^-(v)$, and *non-degree* of v in G as $d'(v) = |V(G) \setminus (N^+(v) \cup N^-(v) \cup \{v\})|$. The *maximum degree* $\Delta(G)$, *maximum out-degree* $\Delta^+(G)$, and *maximum in-degree* $\Delta^-(G)$ of a directed graph G are the maximum of $d(v)$, $d^+(v)$, or $d^-(v)$, respectively, taken over all vertices of G. Other degree-related notions can be defined analogously.

If H is an oriented graph, then its *underlying graph* is an undirected graph G with $V(G) = V(H)$ and the edge set $E(G) = \{\{v, w\} : (v, w) \in E(H)\}$. We refer to H as an *orientation* of G. Any orientation of K_n is called a *tournament* on n vertices.

The notions of subgraphs, induced subgraphs, homomorphisms, and isomorphisms are essentially the same as for undirected graphs.

We define the following special graphs:

- a *complete digraph* $\overrightarrow{K_n}$ on $n \geq 1$ vertices with arcs vw and wv for every pair $v, w \in V(\overrightarrow{K_n})$ of distinct vertices;
- a *transitive tournament* $\overrightarrow{T_n}$ on $n \geq 1$ vertices $v_1, \dots, v_n$ and arc set $\{v_i v_j : 1 \leq i < j \leq n\}$;
- a *directed path* $\overrightarrow{P_n}$, also called a *directed n-path*, on $n \geq 1$ vertices $v_1, \dots, v_n$ with arc set $E(\overrightarrow{P_n}) = \{v_i v_{i+1} : 1 \leq i < n\}$. Vertices v_1 and v_n are *endpoints* of $\overrightarrow{P_n}$, while vertices $v_2, \dots, v_{n-1}$ are *internal vertices* of $\overrightarrow{P_n}$. We say that a directed path P *joins* vertices v and w if these vertices are the endpoints of P, and it *goes from v to w* if the arcs of P are directed towards w. We also define the *length* of a directed path as the number of its arcs;

- a *directed cycle* $\overrightarrow{C_n}$, also called a *directed n-cycle*, on $n \geq 3$ vertices $v_1, \ldots, v_n$ with arcs $v_n v_1$ and $v_i v_{i+1}$ for $1 \leq i < n$. We define the *length* of a directed cycle as the number of its arcs;
- an orientation $\overrightarrow{K_{n,m}}$ of $K_{n,m}$ with vertex set $V(\overrightarrow{K_{n,m}}) = A \cup B$, where $|A| = n$ and $|B| = m$, and arc set $E(\overrightarrow{K_{n,m}}) = \{vw : v \in A \text{ and } w \in B\}$. For $n \geq 2$, we shall write $\overrightarrow{K_{1,n-1}}$ also as $\overrightarrow{S_n}$ and call it a *directed star* on n vertices;
- an *empty graph* I_n on $n \geq 1$ vertices with empty arc set.

Let v and w be two vertices of a directed graph G. We define the *distance* $d_G(v, w)$ from v to w in G as the minimum length of a directed path in G from v to w; if no such path exists, then we put $d_G(v, w) = \infty$. A directed graph G is *strongly connected* if $d_G(v, w) < \infty$ for every two vertices $v, w \in V(G)$. A *strongly connected component* of G is a strongly connected subgraph of G which is not contained in any larger strongly connected subgraph of H. A directed graph G is *connected* if its underlying graph is connected, and a subgraph of G is a *connected component* if the respective subgraph of the underlying graph of G is its connected component. A *Hamiltonian path* in G is a subgraph of G which is isomorphic to $\overrightarrow{P_{|V(G)|}}$. We say that a subgraph M of G is a *matching* if every two arcs of M are vertex disjoint.

2.3 Blow-ups and other graph operations

Let G and H be graphs (both undirected, directed, or oriented). We define the following graph operations:

- the *union* $G \cup H$ of graphs G and H as the graph with $V(G \cup H) = V(G) \cup V(H)$ and $E(G \cup H) = E(G) \cup E(H)$;
- a *disjoint union* $G \sqcup H$ of graphs G and H as a graph isomorphic to $G' \cup H'$, where G is isomorphic to G', H is isomorphic to H', and $V(G') \cap V(H') = \emptyset$;
- if G and H are oriented graphs, then $G \Rightarrow H$ denotes a graph being a disjoint union of G and H with all possible arcs from vertices of G to vertices of H;
- the *tensor product* $G \otimes H$ of graphs G and H as the graph with $V(G \otimes H) = V(G) \times V(H)$; a vertex (g_1, h_1) is adjacent to (g_2, h_2) if and only if $g_1 g_2 \in E(G)$ and $h_1 h_2 \in E(H)$;
- the *composition* $G \odot H$ of graphs G and H as the graph with $V(G \odot H) = V(G) \times V(H)$; a vertex (g_1, h_1) is adjacent to (g_2, h_2) if either $g_1 = g_2$ and $h_1 h_2 \in E(H)$ or $g_1 g_2 \in E(G)$.

Let G_1 and G_2 be vertex disjoint graphs and $v \in V(G_1)$. We say that a graph G is obtained by *substituting* G_2 for v in G_1 if the following holds:

- $V(G) = V(G_1 - v) \cup V(G_2)$,
- $G[V(G_1 - v)] = G_1 - v$,
- $G[V(G_2)] = G_2$,
- for every $u \in V(G_1 - v)$ and $w \in V(G_2)$, we have $uw \in E(G)$ if and only if $uv \in E(G_1)$; if G_1 and G_2 are directed, then also $wu \in E(G)$ if and only if $vu \in E(G_1)$.

We say that a graph G' is a *blow-up* of a graph G if there exists a sequence of vertex disjoint graphs $(H_v)_{v \in V(G)}$ such that G' was obtained by sequentially substituting H_v for v in G for every $v \in V(G)$. Operation $G \odot H$ is an example of a blow-up of a graph G with $H_v = H$ for all $v \in V(G)$, see Figure 2.1. Usually, graphs H_v are taken to be isomorphic to $I_{k(v)}$ for some integers $k(v) > 0$, and we shall assume the same for the rest of the thesis if not specified otherwise. Moreover, if $|k(v) - k(w)| \leq 1$ for every choice of vertices $v, w \in V(G)$, then we shall refer to G' as a *balanced blow-up* of G.

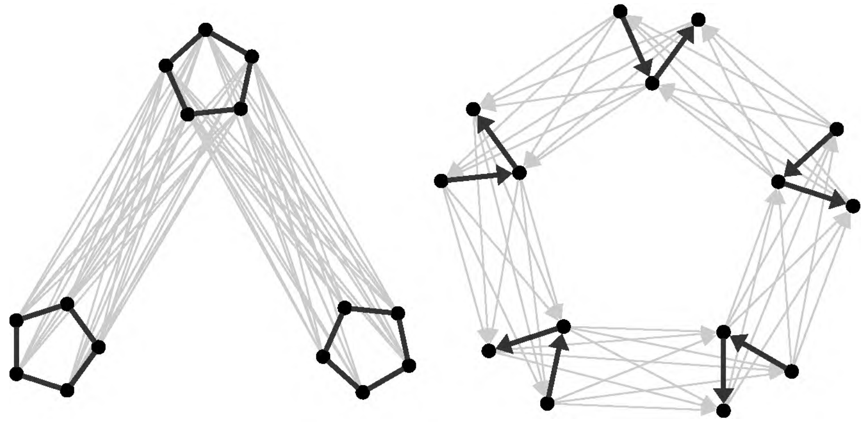


Figure 2.1: Compositions $P_3 \odot C_5$ and $\overrightarrow{C_5} \odot \overrightarrow{P_3}$. Composition of graphs as a binary operation on graphs is in general non-commutative.

For arbitrary graphs G_1, G_2, G_3 we have an isomorphism $G_1 \odot (G_2 \odot G_3) \approx (G_1 \odot G_2) \odot G_3$. Therefore, it makes sense to write $G_1 \odot G_2 \odot G_3$ and define inductively $G^{\odot n}$ as $G \odot G^{\odot(n-1)}$ for $n > 0$ and I_1 for $n = 0$. We shall call graphs of the form $G^{\odot n}$ *iterated blow-ups* of G.

For an integer $k \geq 1$, a graph G' is a *k-th power* of a graph G if $V(G') = V(G)$ and $E(G') = \{vw : 1 \leq d_G(v, w) \leq k\}$. If $k = 2$, then we also say that G' is a *square* of G.

2.4 Regularity Lemma

In this section, we shall discuss one of the most powerful tools in Extremal Graph Theory, known as Szemerédi's Regularity Lemma, in both undirected and directed settings.

2.4.1 Undirected case

Let G be a graph and $A, B \subseteq V(G)$ be disjoint vertex subsets. Let $e(A, B)$ denote the number of edges between A and B, and define the *density* of a pair (A, B) as

$$\rho(A, B) = \frac{e(A, B)}{|A| \cdot |B|}.$$

A pair (A, B) is ε*-regular* if for any $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \varepsilon |A|$ and $|Y| \geq \varepsilon |B|$, we have

$$|\rho(X, Y) - \rho(A, B)| < \varepsilon.$$

Szemerédi's Regularity Lemma asserts that every large graph can be partitioned into a small number of parts such that the edges between most of those parts behave almost like in a random bipartite graph.

Theorem 2.1 (Szemerédi's Regularity Lemma [74]). *For any $\varepsilon > 0$ and $m > 0$, there exists $M > m$ such that every graph on $n \geq m$ vertices has a partition of its vertex set into $k + 1$ parts $V_0, \ldots, V_k$ for some k with $m \leq k \leq M$ such that:*

- $|V_0| \leq \varepsilon n$,
- $|V_1| = \ldots = |V_k|$,
- *all but at most εk^2 pairs (V_i, V_j) with $1 \leq i < j \leq k$ are ε-regular.*

A partition $V_0, \ldots, V_k$ is called *ε-regular* if it satisfies all the properties in Theorem 2.1.

For an ε-regular partition $V_0, \ldots, V_k$ of some graph G and for $d \in [0,1]$, we define the *reduced graph* $R(\varepsilon, d)$ as the graph with vertex set $\{1, 2, \ldots, k\}$ and edge set $\{ij :$ a pair (V_i, V_j) is ε-regular with density at least $d\}$.

One of the most useful properties of ε-regular partitions is known as Embedding Lemma, which roughly says that we can imply the existence of a copy of a graph H in G from the existence of a copy of H in a blow-up of $R(\varepsilon, d)$. Since the latter implies the existence of a homomorphism $H \to R(\varepsilon, d)$, we state Embedding Lemma in the following simplified form.

Theorem 2.2 (Embedding Lemma [57]). *For any $d \in (0,1]$ and $\Delta \geq 1$, there exists ε_0 such that for any $\varepsilon \leq \varepsilon_0$ and any $s \geq 1$ the following holds. Let H be a graph with $\Delta(H) \leq \Delta$ and $|V(H)| \leq s$, and let $R = R(\varepsilon, d)$ be the reduced graph of an ε-regular partition of some graph G with parts of size at least $C = C(s, d, \Delta)$. If there exists a homomorphism $H \to R$, then G contains a copy of H.*

The following corollary, which we formulate as a lemma, and its directed analogue will be used in many proofs regarding Turán numbers.

Lemma 2.3. *For any graph H and $\varepsilon > 0$, there exists n_0 such that from any H-free graph G on $n \geq n_0$ vertices we can remove at most εn^2 edges and obtain a graph G' for which there exists no homomorphism $H \to G'$.*

Proof. Since the proof is mostly about choosing the right value of constants, we shall emphasize the crucial dependencies. Let $\Delta = \Delta(H)$, $s = |V(H)|$, and $d = \varepsilon/2$. Let $\varepsilon_0 = \varepsilon_0(d, \Delta) > 0$ and $C = C(s, d, \Delta)$ be as in Theorem 2.2. Take $m > \frac{1}{4\varepsilon}$ and $\bar{\varepsilon} > 0$ such that $\bar{\varepsilon} < \min(\varepsilon_0, \varepsilon/4)$, and let $M = M(m, \bar{\varepsilon})$ be as in Theorem 2.1. Suppose that G is an H-free graph on n vertices for some $n \geq n_0 := CM/(1 - \bar{\varepsilon})$. Apply Theorem 2.4 to obtain an $\bar{\varepsilon}$-regular partition $V_0, \ldots, V_k$ of G for some integer k satisfying $m \leq k \leq M$, and let $R = R(\bar{\varepsilon}, d)$ denote the reduced graph of this partition. Since each V_i for $i \geq 1$ has at least C vertices, we conclude from Theorem 2.2 that there exists no homomorphism $H \to R$, as otherwise G would contain a copy of H. Remove from G all edges:

- between parts of pairs which are not $\bar{\varepsilon}$-regular (at most $\bar{\varepsilon} n^2$ removed edges),
- between parts of pairs which are $\bar{\varepsilon}$-regular with density at most d (at most $\frac{d}{2} n^2$ removed edges),
- incident to V_0 (at most $\bar{\varepsilon} n^2$ removed edges),
- inside each part V_i for $i \geq 1$ (at most $\frac{n^2}{m}$ removed edges),

and let G' denote the obtained graph. Then, G' is H-free, as any copy of H in G' would induce a homomorphism $H \to R$, and it was obtained from G by removing at most εn^2 edges. $\square$

2.4.2 Directed case

In order to make concepts from the unoriented case work in the directed setting, one needs to slightly modify the definition of an ε-regular pair.

Let G be a directed graph and let $A, B \subseteq V(G)$ be disjoint vertex subsets. Denote by $\vec{e}(A, B)$ the number of arcs from A to B and by $\overline{e}(A, B)$ the number of pairs $(v, w) \in A \times B$ such that $vw, wv \in E(G)$. Define the *directed densities* of a pair (A, B) as

$$\vec{\rho}(A, B) = \frac{\vec{e}(A, B)}{|A| \cdot |B|}, \quad \overline{\rho}(A, B) = \frac{\overline{e}(A, B)}{|A| \cdot |B|}.$$

We say that a pair (A, B) is *ε-regular* if for any $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \varepsilon |A|$ and $|Y| \geq \varepsilon |B|$ we have

$$|\vec{\rho}(A, B) - \vec{\rho}(X, Y)| < \varepsilon, \qquad |\vec{\rho}(B, A) - \vec{\rho}(Y, X)| < \varepsilon, \qquad |\overline{\rho}(A, B) - \overline{\rho}(X, Y)| < \varepsilon.$$

Theorem 2.4 (Regularity Lemma for directed graphs [2]). *For any $\varepsilon > 0$ and integer $m > 0$, there exists $M > m$ such that every graph on $n \geq m$ vertices has a partition of its vertex set into $k+1$ parts $V_0, \ldots, V_k$ for some $m \leq k \leq M$ such that:*

- $|V_0| \leq \varepsilon n$,
- $|V_1| = \ldots = |V_k|$,
- *all but at most εk^2 pairs (V_i, V_j) with $1 \leq i < j \leq k$ are ε-regular.*

A partition $V_0, \ldots, V_k$ is called ε-regular if it satisfies all the properties in Theorem 2.4.

For an ε-regular partition $V_0, \ldots, V_k$ of some directed graph G and for $d \in [0,1]$, we define the *reduced digraph* $\overrightarrow{R}(\varepsilon, d)$ as the directed graph with vertex set $\{1, 2, \ldots, k\}$ and arc set $\{ij :$ a pair (V_i, V_j) is ε-regular and $\vec{\rho}(V_i, V_j) \geq d\}$.

One may prove the following analogue of Theorem 2.2 using the same approach as in [57].

Theorem 2.5 (Embedding Lemma for directed graphs). *For any $d \in (0, 1]$ and $\Delta \geq 1$, there exists ε_0 such that for any $\varepsilon \leq \varepsilon_0$ the following holds. Let H be a digraph with $\Delta(H) \leq \Delta$ and $\overrightarrow{R} = \overrightarrow{R}(\varepsilon, d)$ be the reduced digraph of an ε-regular partition of some digraph G with parts of size at least $C = C(d, \Delta)$. If there exists a homomorphism $H \to R$, then G contains a copy of H.*

It is also immediate that we have an analogue of Lemma 2.3.

Lemma 2.6. *For any digraph H and $\varepsilon_0 > 0$, there exists n_0 such that from any H-free digraph G on $n \geq n_0$ vertices we can remove at most $\varepsilon_0 n^2$ arcs and obtain a digraph G' for which there exists no homomorphism $H \to G'$.*

2.5 Flag algebra method

The flag algebra method was developed by Razborov [69] and quickly became recognized as a very powerful tool for tackling numerous open problems in Extremal Graph Theory and other branches of Extremal Combinatorics. Some of the results presented in this thesis were obtained using the Flagmatic software written by Vaughan [77], which implements the flag algebra method. The goal of this section is to outline the most important mathematical concepts behind this method. In the last subsection, we also briefly explain the main structure of our codes that use Flagmatic.

For simplicity, we shall restrict our attention to undirected or directed graphs. Razborov introduced flag algebras for general discrete structures with a set of very basic properties; for the details, we refer to the original paper [69].

2.5.1 Flags and flag densities

For an integer $k \geq 0$, a *type* σ of order k is a graph with $V(\sigma) = [k]$, where the only type of order zero is denoted by $\emptyset$. A *σ-flag* of order ℓ is a pair (G, θ), where G is a graph on ℓ vertices and θ is an injective homomorphism $\sigma \to G$ such that $\theta([k])$ induces a subgraph isomorphic to σ. Two σ-flags (G_1, θ_1) and (G_2, θ_2) are *isomorphic* if there exists an isomorphism f of G_1 and G_2 that preserves σ, i.e. $f(\theta_1(i)) = \theta_2(i)$ for every $i \in [k]$. If $F = (G, \theta)$ is a σ-flag, $X \subseteq V(G) \setminus \theta([k])$, and G_1 is a subgraph of G induced by $X \cup \theta([k])$, then we say that (G_1, θ) is a σ-flag *induced* by X.

For a fixed type σ, let $\mathcal{F}^\sigma$ be the set of all σ-flags, and $\mathcal{F}^\sigma_\ell$ be the set of all σ-flags of order ℓ. For $F_1, F_2 \in \mathcal{F}^\sigma$, where $F_1 = (G_1, \theta_1)$ and $F_2 = (G_2, \theta_2)$, we define a *density* $d(F_1, F_2)$ of F_1 in F_2 as the probability that a random subset of $V(G_2) \setminus \theta_2([k])$ of size $|V(G_1)| - k$ induces a σ-flag isomorphic to F_1. If $\sigma = \emptyset$, then $d(F_1, F_2)$ is just a probability that a random $|V(G_1)|$-element subset of $V(G_2)$ induces a copy of G_1 in G_2.

We shall also define a quantity $d(F_1, F_2; F)$ for σ-flags $F_1 = (G_1, \theta_1)$, $F_2 = (G_2, \theta_2)$, and $F = (G, \theta)$ as the probability that for a randomly chosen pair (X_1, X_2) of disjoint subsets of $V(G) \setminus \theta([k])$ with $|X_1| = |V(G_1)| - k$ and $|X_2| = |V(G_2)| - k$, X_1 induces a σ-flag isomorphic to F_1 and X_2 induces a σ-flag isomorphic to F_2.

2.5.2 Flag algebras

Fix a type σ of order k. Let $\mathbb{R}\mathcal{F}^\sigma$ denote the set of formal finite linear combinations of σ-flags with real coefficients. It is equipped with natural operations of addition and multiplication by a scalar, hence it can be treated as a vector space over $\mathbb{R}$.

Let $\mathcal{K}^\sigma$ be the linear subspace of $\mathbb{R}\mathcal{F}^\sigma$ generated by all vectors of the form

$$F - \sum_{F' \in \mathcal{F}^\sigma_\ell} d(F, F')F',$$

for all $\ell \geq k$ and all $F \in \mathcal{F}^\sigma$ of order at most ℓ.

For $F_1 \in \mathcal{F}^\sigma_{\ell_1}$ and $F_2 \in \mathcal{F}^\sigma_{\ell_2}$, let $\ell = \ell_1 + \ell_2 - k$ and define a product

$$F_1 \cdot F_2 = \sum_{F \in \mathcal{F}^\sigma_\ell} d(F_1, F_2; F)F.$$

This operation can be linearly extended to the whole $\mathbb{R}\mathcal{F}^\sigma \otimes \mathbb{R}\mathcal{F}^\sigma$.

Proposition 2.7 (Razborov [69]). *Let $\mathcal{A}^\sigma = \mathbb{R}\mathcal{F}^\sigma / \mathcal{K}^\sigma$ and $F_1, F_2 \in \mathbb{R}\mathcal{F}^\sigma$. The multiplication*

$$[F_1] \cdot [F_2] := [F_1 \cdot F_2]$$

is well defined with the multiplicative identity $1_\sigma := [(\sigma, \mathrm{id})]$ and turns $\mathcal{A}^\sigma$ into an algebra over $\mathbb{R}$.

We shall call $\mathcal{A}^\sigma$ a *flag algebra*. For convenience, we shall usually write F instead of $[F]$ if it does not lead to confusion. Also, when $c \in \mathbb{R}$, we shall write $c1_\sigma$ simply as c. Following this convention, observe that in a flag algebra for any $\ell \geq k$ we have $\sum_{F \in \mathcal{F}^\sigma_\ell} F = 1$.

2.5.3 Convergent sequences and limit homomorphisms

Consider an infinite sequence $(F_n)_{n \geq 1}$ of σ-flags, where $F_n \in \mathcal{F}^\sigma_{k_n}$. We shall say that this sequence is *increasing* if $k_n < k_{n+1}$ for every $n \geq 1$. An increasing sequence of σ-flags is *convergent* if the sequence of densities $(d(F, F_n))_{n \geq 1}$ is convergent for every $F \in \mathcal{F}^\sigma$.

For a convergent sequence $(F_n)_{n \geq 1}$, define a map $\Phi : \mathcal{F}^\sigma \to \mathbb{R}$ as $\Phi(F) = \lim_{n \to \infty} d(F, F_n)$. This map can be linearly extended to $\mathbb{R}\mathcal{F}^\sigma$. One may show that the kernel of this map contains $\mathcal{K}^\sigma$, therefore the linear map $\Phi : \mathcal{A}^\sigma \to \mathbb{R}$ given by $\Phi([F]) := \Phi(F)$ for all $F \in \mathbb{R}\mathcal{F}^\sigma$ is well defined.

Proposition 2.8 (Razborov [69]). *The linear map $\Phi : \mathcal{A}^\sigma \to \mathbb{R}$ preserves multiplication and identity, therefore is an algebra homomorphism.*

In particular, for any $F \in \mathcal{F}^\sigma$, we have $\Phi(F) \in [0, 1]$ and $\Phi(1_\sigma) = 1$. We shall call Φ a *limit homomorphism.*

Let $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ be the set of all algebra homomorphisms $\Phi : \mathcal{A}^\sigma \to \mathbb{R}$ with the property that $\Phi(F) \in [0, 1]$ for every $F \in \mathcal{F}^\sigma$. By definition, every limit homomorphism belongs to $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$. The following asserts that the converse is also true, i.e. every element of $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ is a limit homomorphism.

Theorem 2.9 (Lovász, Szegedy [61]). *For every $\Phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ there exists a convergent sequence $(F_n)_{n \geq 1}$ of σ-flags such that $\Phi(F) = \lim_{n \to \infty} d(F, F_n)$ for every $F \in \mathcal{F}^\sigma$.*

Therefore, $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ is exactly the set of all limit homomorphisms. Using Theorem 2.9, many problems from Extremal Graph Theory can be translated to the language of flag algebras and homomorphisms. For instance, the asymptotic version of Mantel's Theorem can be expressed as follows:

For any $\Phi \in \mathrm{Hom}^+(\mathcal{A}^\emptyset, \mathbb{R})$, if $\Phi(C_3) = 0$, then $\Phi(K_2) \leq \frac{1}{2}$.

To see that it indeed implies the asymptotic version of Mantel's Theorem, consider any increasing sequence $(F_n)_n$ of triangle-free graphs (or, equivalently, elements of $\mathcal{F}^{\emptyset}$) such that $|E(F_n)| \geq (1/4 + \varepsilon)\,|V(F_n)|^2$ for every n. One may easily show using Cantor's diagonal argument that every increasing sequence of flags contains a convergent subsequence, therefore without loss of generality we may assume that $(F_n)_n$ is already convergent. But then, it induces a limit homomorphism Φ that satisfies $\Phi(C_3) = 0$ and $\Phi(K_2) \geq 1/4 + \varepsilon$, a contradiction.

Sometimes, it is possible to derive the „finite" versions of theorems from their asymptotic counterparts, for instance by taking a sequence of blow-ups of the smallest counterexample. In particular, one may prove this way the exact version of Mantel's Theorem without the $o(n^2)$ error term.

Our next goal is to introduce tools that can be used to prove flag-algebraic statements as above.

2.5.4 Averaging operator and Cauchy-Schwarz inequality

For a fixed type σ and $F = (G, \theta) \in \mathcal{F}^{\sigma}$, define $q_{\sigma}(F) \in [0, 1]$ as the probability that a random injective function $\theta_0 : V(\sigma) \to V(G)$ yields a σ-flag (G, θ_0) isomorphic to F. Define an *averaging operator* $[\![\cdot]\!] : \mathcal{F}^{\sigma} \to \mathcal{F}^{\emptyset}$ as $[\![F]\!] = q_{\sigma}(F) \cdot G$. The linear extension of this operator to $\mathbb{R}\mathcal{F}^{\sigma}$ has the property that $[\![\mathcal{K}^{\sigma}]\!] \subseteq \mathcal{K}^{\emptyset}$, hence it induces a well-defined linear operator $[\![\cdot]\!] : \mathcal{A}^{\sigma} \to \mathcal{A}^{\emptyset}$.

For $f, g \in \mathcal{A}^{\sigma}$, write $f \geq g$ if $\Phi(f - g) \geq 0$ for every $\Phi \in \mathrm{Hom}^{+}(\mathcal{A}^{\sigma}, \mathbb{R})$.

Theorem 2.10 (Cauchy-Schwarz inequality [69]). *For any $f, g \in \mathcal{A}^{\sigma}$,*

$$[\![f^2]\!] \cdot [\![g^2]\!] \geq [\![f \cdot g]\!]^2.$$

In particular, $[\![f^2]\!] \geq 0$.

Since the operator $[\![\cdot]\!]$ is not an algebra homomorphism, Theorem 2.10 is a way of obtaining non-trivial flag algebra inequalities.

2.5.5 The semidefinite method

For the following, it is convenient for any type σ and $F \in \mathcal{F}^{\sigma}$ to linearly extend $d(\cdot, F)$ to the whole $\mathbb{R}\mathcal{F}^{\sigma}$.

Suppose that for some $f \in \mathcal{A}^{\emptyset}$ we want to find the minimum value of $c \in \mathbb{R}$ such that $f \leq c$. If we write f as a linear combination of flags on ℓ vertices:

$$f = \sum_{F \in \mathcal{F}^{\emptyset}_{\ell}} a_F \cdot F,$$

then we have

$$f \leq \max_{F \in \mathcal{F}^{\emptyset}_{\ell}} a_F \cdot \sum_{F \in \mathcal{F}^{\emptyset}_{\ell}} F = \max_{F \in \mathcal{F}^{\emptyset}_{\ell}} a_F.$$

One can improve this bound in the following way. Fix some type σ of order $1 \leq k \leq \ell - 2$ and enumerate all elements $F_1, \ldots, F_t$ of $\mathcal{F}^{\sigma}_m$, where $1 \leq m \leq (\ell + k)/2$ and $t = |\mathcal{F}^{\sigma}_m|$. Consider a positive semidefinite matrix $Q = (q_{ij})_{i,j=1}^{t}$. By Theorem 2.10, for a vector $v = (F_i)_{i=1}^{t}$ with $F_i \in \mathcal{F}^{\sigma}_m$, we have $[\![vQv^T]\!] \geq 0$. On the other hand,

$$[\![vQv^T]\!] = \sum_{F_i, F_j \in \mathcal{F}^{\sigma}_m} q_{ij} [\![F_i \cdot F_j]\!] = \sum_{F \in \mathcal{F}^{\emptyset}_{\ell}} c_F \cdot F,$$

where the coefficients $c_F = c_F(\sigma, m, Q)$ are given by the formula

$$c_F(\sigma, m, Q) = \sum_{F_i, F_j \in \mathcal{F}^{\sigma}_m} q_{ij} \cdot d([\![F_i \cdot F_j]\!], F).$$

Therefore, $f \leq \sum_{F \in \mathcal{F}_\ell^\emptyset} (a_F + c_F) \cdot F$ and it follows that

$$f \leq \max_{F \in \mathcal{F}_\ell^\emptyset} (a_F + c_F).$$

Since some of the coefficients c_F may be negative, we may obtain a strictly better bound for a suitable choice of Q. In fact, we can consider a sequence $(\sigma_i, m_i, Q_i)_{i=1}^s$ of triples for some $s \geq 1$, for which we get the following upper bound:

$$f \leq \max_{F \in \mathcal{F}_\ell^\emptyset} (a_F + \sum_{i=1}^{s} c_F(\sigma_i, m_i, Q_i)). \tag{2.1}$$

Once the types and integers are fixed, the problem of minimizing the right side of (2.1) over all positive semidefinite matrices Q_i can be formulated as a semidefinite programming problem (SDP) and as such can be solved numerically. However, numerical solutions need yet to be rounded to rational values in order to give a formal proof. In the process of rounding, the matrices Q_i may lose the property of being positive semidefinite, or the bound we get may become suboptimal. Still, there are a few methods that can be utilized to obtain the exact bound, which are described e.g. in Section 2.4.2 of [4] or in the appendix of [6]. In particular, one may provide the eigenvectors corresponding to the eigenvalue zero that should be preserved in the process of rounding the matrices. Some of them can be determined from the extremal constructions, but sometimes it is necessary to guess them from their numerical approximation for the matrices found by solving SDP.

Sometimes, we want to find the maximum value of some density with some additional assumptions about forbidden substructures. For instance, in Mantel's Theorem, we consider only triangle-free graphs. In this case, one just puts $c_F = 0 = a_F$ for all $F \in \mathcal{F}_\ell^\emptyset$ that contain a triangle. This can be made more rigorous by redefining $\mathcal{F}^\sigma$ as the family of those σ-flags which are triangle-free.

Vaughan wrote the program Flagmatic that reformulates a given flag-algebraic problem into an SDP, which can be solved using any publicly available SDP solver, and implements the rounding algorithm for the numerical solutions.

2.5.6 Explanation of the codes

We wrote some programs to prove several results in Chapters 5 and 7. All those programs require the SageMath [70] and Flagmatic software. Below, we include a short explanation of the codes. As an example, we consider the `graph14.sage` file, which corresponds to the problem described in Section 7.3.14.

```
from flagmatic.all import *
P=OrientedGraphProblem(5, density="4:12233414", type_orders=[3])
P.set_extremal_construction(field=QQ, target_bound=3/16)
P.add_sharp_graphs
    (0,9,10,11,12,13,156,157,158,159,167,168,169,170,171,180,181,182,187)
#3:
P.add_zero_eigenvectors(0,matrix(QQ,
    [(8,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,-1,1,1,0,0),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,0,-1,1,0,1,0),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,0,0,1,-1,0,0,1)]))
#3:1232
P.add_zero_eigenvectors(2,matrix(QQ,
    [(0,0,0,0,0,0,0,0,0,-1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0),
    (0,0,0,0,0,0,0,0,0,0,-1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0),
    (0,0,0,2,2,0,0,0,0,1,1,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0),
    (0,0,0,1,-1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)]))
#3:1223
P.add_zero_eigenvectors(3,matrix(QQ,
    [(0,0,0,2,2,0,0,0,0,1,1,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0),
```

```
      (0,0,0,1,-1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
      (0,0,0,0,0,0,0,0,0,-1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0),
      (0,0,0,0,0,0,0,0,0,0,-1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0)]))
11 #3:1213
12 P.add_zero_eigenvectors(4,matrix(QQ,
      [(0,2,2,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,0,0,0,0,0,0,0,0),
      (0,1,-1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
      (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,-1,0,0,0,0,0,0,0,0),
      (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,-1,0,0,0,0,0,0,0,0,0)]))
13 P.solve_sdp()
14 P.make_exact(16)
15 P.write_certificate("graph14.cert")
```

In line 1, we import Flagmatic libraries. In line 2, using the function `OrientedGraphProblem`, we define the problem P we want to solve. First two parameters of the function are mandatory. The first one defines the maximum order of flags used in the computations, while the second one defines the graph density we want to maximize. In the considered example, "`4:12233414`" means an oriented graph on 4 vertices (labeled with numbers from 1 to 4) and arcs 12, 23, 34, 14, i.e. graph . We can also specify other conditions for the problem. For instance, we limit the computations in the above code only to types of order 3.

At this moment, we have already defined the problem which Flagmatic can solve numerically using an SDP solver. However, in order to obtain an exact result, we need to provide more information for the rounding procedure. In line 3, we define the target bound (i.e. the value which we believe is the correct upper bound) and over which field we perform the computations — here, `QQ` stands for the field of rational numbers. In line 4, we specify which graphs densities are forced to satisfy the bound as equalities, they follow from the expected extremal constructions. Finally, in lines 5-12, we provide additional eigenvectors corresponding to the eigenvalue zero needed to properly round the semidefinite matrix.

It remains to perform the computations. In line 13, we numerically solve the SDP (by default it is performed by `csdp` [13], but in a few cases where double-double precision is needed, we used `sdpa-dd` [66]). In line 14, we round the output of the solver in the considered field. Finally, in the last line, we create a certificate, which contains a complete description of the problem, used flags, graphs, matrices, and the proven bound.

By default, Flagmatic is distributed together with an independent checker program written by Vaughan, `inspect_certificate.py`. A detailed explanation of the contents of certificates and how the checker program works can be found in [33].

Chapter 3

Generalized Turán-type problems for cycles

The results in this chapter are based on joint work with Andrzej Grzesik and are published in the article A. Grzesik and B. Kielak: *On the maximum number of odd cycles in graphs without smaller odd cycles*, J. Graph Theory **99(2)** (2022), 240–246. We came up independently with the idea of proving Claim 3.2 and the related definitions of sets A_i. Section 3.1 is an introduction written for the purpose of this thesis. Section 3.2 is Section 2 from [42] with no substantial changes, and Section 3.3 is Section 3 from [42] with the corrected statement of Conjecture 2 (Conjecture 3.5 in this thesis).

3.1 Introduction

For fixed connected graphs T and H, let $\mathrm{ex}(n,T,H)$ denote the maximum number of copies of T in an H-free graph on n vertices. For $T = K_2$, it is just the Turán number $\mathrm{ex}(n,H)$, which was already discussed in Chapter 1. Even though the systematic studies of this problem for $T \neq K_2$ were initiated by Alon and Shikhelman [3], some specific cases were considered earlier.

The first known result is due to Erdős [26] and Zykov [79], who independently determined $\mathrm{ex}(n,K_s,K_t)$ for $s < t$.

In 1984, Erdős [25] asked for the maximum number of copies of C_5 in a triangle-free n-vertex graph. He conjectured that the maximum is obtained by a balanced blow-up of C_5. Győri [44] proved an upper bound within a factor 1.03 of the optimal. Using flag algebras method, Grzesik [41] and, independently, Hatami et al. [49] proved that any triangle-free graph on n vertices has at most $(n/5)^5$ copies of C_5, which is a tight bound for n divisible by 5. Michael [65] presented a sporadic counterexample to the characterization of the extremal cases by presenting a graph on 8 vertices showing that not only a balanced blow-up of a C_5 can achieve the maximum. Recently, Lidický and Pfender [60], also using flag algebras, completely determined the extremal graphs for every n by showing that the graph pointed out by Michael is the only extremal graph which is not a balanced blow-up of a pentagon.

Since we will be interested mostly in the case when both T and H are cycles, let us mention other recent results regarding them. Bollobás and Győri [10] proved that $\mathrm{ex}(n,C_3,C_5) = \Theta(n^{3/2})$, Győri and Li [46] extended this result to obtain bounds for $\mathrm{ex}(n,C_3,C_{2k+1})$, which were later improved by Alon and Shikhelman [3] and by Füredi and Özkahya [36]. Gishboliner and Shapira [39] proved a correct order of magnitude of $\mathrm{ex}(n,C_k,C_\ell)$ for $k \geq 4$ and $\ell \geq 3$, and independently Gerbner et al. [38] for all even cycles, together with the tight asymptotic value of $\mathrm{ex}(n,C_4,C_{2k})$ and $\mathrm{ex}(n,C_6,C_8)$. Recently, Górski and Grzesik [40] determined $\mathrm{ex}(n,C_5,C_7)$ exactly, and Győri et al. [45] found an exact formula for $\mathrm{ex}(n,C_4,C_6)$ if $n > 3\binom{31}{4}$.

Here, we prove the following result from which we can conclude the exact asymptotics of $\mathrm{ex}(n,C_k,C_{k-2})$ for all odd $k \geq 7$, unknown before.

Theorem 3.1. *For each odd integer $k \geq 7$, any graph on n vertices without odd cycles of length smaller than k contains at most $(n/k)^k$ cycles of length k. Moreover, for $k \mid n$, a blow-up $C_k \odot I_{n/k}$ is the only graphs attaining this maximum. As a corollary,* $\mathrm{ex}(n, C_k, C_{k-2}) = (n/k)^k + o(n^k)$.

The corollary part of Theorem 3.1 follows from Lemma 2.3. Indeed, since C_{k-2} can be mapped homomorphically into C_ℓ for any odd ℓ less than k, if a large n-vertex graph G is C_{k-2}-free, then by Lemma 2.3 we can remove $o(n^2)$ edges from G to eliminate all copies of shorter cycles, thus the number of copies of C_k in G would change by at most $o(n^k)$.

The proof of Theorem 3.1 is based on the method developed by Král', Norin, and Volec [59].

3.2 Main result

Fix an odd integer $k \geq 7$, and let G be any n-vertex graph without C_ℓ for all odd ℓ between 3 and $k-2$. Since there are no odd cycles smaller than k, each k-cycle in G is induced.

We bound the number of k-cycles by bounding the probability that sampling vertices of G one by one at random results in a fixed induced k-cycle. However, instead of sampling the vertices in the cycle order, we do it with a small shift and sample the fourth vertex before the third. This is to avoid the situation that a particular 3-vertex induced path in G cannot be extended to a k-cycle, which happens, for example, when G is a blow-up of a k-cycle.

For any k-cycle $v_0 v_1 \ldots v_{k-1}$ contained in G, by a ***good sequence*** we mean a sequence $D = (z_i)_{i=0}^{k-1}$, where $z_i = v_i$ for $i \leq 1$ and $i \geq 4$, $z_2 = v_3$, and $z_3 = v_2$, i.e. v_2 and v_3 are in the reversed order. Note that there are $2k$ different good sequences corresponding to a single induced k-cycle. For any vertices v and w, by $d(v, w)$ we mean the minimum distance between the vertices v and w in G.

For a fixed good sequence D, we define the following sets:

$$\begin{aligned}
A_0(D) &= V(G),\\
A_1(D) &= N(z_0),\\
A_2(D) &= \{w \notin N(z_0) : d(z_1, w) = 2\},\\
A_3(D) &= N(z_1) \cap N(z_2),\\
A_4(D) &= \{w : z_0 z_1 z_3 z_2 w \text{ is an induced path}\},\\
A_i(D) &= \{w : z_0 z_1 z_3 z_2 z_4 \ldots z_{i-1} w \text{ is an induced path}\} \text{ for } 5 \leq i \leq k-2,\\
A_{k-1}(D) &= \{w : z_0 z_1 z_3 z_2 z_4 \ldots z_{k-2} w \text{ is an induced cycle}\}.
\end{aligned}$$

Define the ***weight*** $w(D)$ of a good sequence D as

$$w(D) = \prod_{i=0}^{k-1} |A_i(D)|^{-1} = \frac{1}{n} \prod_{i=1}^{k-1} |A_i(D)|^{-1}.$$

This quantity has the following probabilistic interpretation. Suppose we want to sample k vertices $w_0, \ldots, w_{k-1}$ so that $(w_i)_{i=0}^{k-1}$ is a good sequence. We start with choosing w_0 at random from all vertices of G. Next, we pick any neighbor of w_0 to be w_1. In general, w_i is a random vertex from the set $A_i((w_i)_{i=0}^{i-1})$. (Note that the definition of $A_i(D)$ depends only on first i elements of a sequence D.) Then, $w(D)$ is just the probability that the sequence $(w_i)_{i=0}^{k-1}$ obtained in this random process is equal to D.

In particular, the sum of the weights of all good sequences is at most one, since it is the sum of probabilities of pairwise disjoint events.

Fix a k-cycle $v_0 v_1 \ldots v_{k-1}$ in G, let $C = \{v_0, v_1, \ldots, v_{k-1}\}$ be the set of its vertices, and let $D_j = (v_j, v_{j+1}, v_{j+3}, v_{j+2}, v_{j+4}, \ldots, v_{j+k-1})$ for $0 \leq j \leq k-1$, where the indices are considered modulo k, be all the good sequences with the same orientation corresponding to this cycle (half of the total number of good sequences corresponding to this cycle).

If we prove that

$$\left(2\sum_{j=0}^{k-1} w(D_j)\right)^{-1} \leq M$$

for some number M, then $2\sum_{j=0}^{k-1} w(D_j) \geq M^{-1}$. Thus, by summing over all k-cycles (with both orientations) and using the fact that the sum of the weights of all good sequences is at most one, we conclude that the total number of k-cycles is bounded from above by M.

Let $n_{i,j} = |A_i(D_j)|$. Since

$$\left(2\sum_{j=0}^{k-1} w(D_j)\right)^{-1} = \left(2\sum_{j=0}^{k-1}\prod_{i=0}^{k-1} n_{i,j}^{-1}\right)^{-1} = n\left(\sum_{j=0}^{k-1}\left(\frac{n_{1,j}}{2}\right)^{-1}\prod_{i=2}^{k-1} n_{i,j}^{-1}\right)^{-1},$$

the maximum possible value of

$$n\left(\sum_{j=0}^{k-1}\left(\frac{n_{1,j}}{2}\right)^{-1}\prod_{i=2}^{k-1} n_{i,j}^{-1}\right)^{-1} \tag{3.1}$$

is an upper bound on the number of k-cycles in G.

Using the inequality between harmonic mean and geometric mean of k terms and the inequality between geometric mean and arithmetic mean of $k(k-1)$ terms, we obtain

$$\begin{aligned} n\left(\sum_{j=0}^{k-1}\left(\frac{n_{1,j}}{2}\right)^{-1}\prod_{i=2}^{k-1} n_{i,j}^{-1}\right)^{-1} &\leq \frac{n}{k}\left(\prod_{j=0}^{k-1}\frac{n_{1,j}}{2}\prod_{i=2}^{k-1} n_{i,j}\right)^{\frac{1}{k}} \\ &\leq \frac{n}{k}\left(\frac{1}{k(k-1)}\sum_{j=0}^{k-1}\left(\frac{n_{1,j}}{2}+\sum_{i=2}^{k-1} n_{i,j}\right)\right)^{k-1}. \end{aligned}$$

Claim 3.2. *The following inequality holds:*

$$\sum_{j=0}^{k-1}\left(\frac{n_{1,j}}{2}+\sum_{i=2}^{k-1} n_{i,j}\right) \leq n(k-1),$$

with equality if and only if each vertex of G is connected to two vertices of C at distance two.

Proof. It is enough to prove that the contribution of any vertex $w \in V(G)$ to the above sum is at most $k-1$, and that such a contribution can only occur if w is connected to two vertices of C at distance two.

Notice that any vertex $w \in V(G)$ has at most 2 neighbors in C, since otherwise it creates a shorter odd cycle. For the same reason, since $k \geq 7$, each vertex w satisfies the following property:

($\star$) There are at most three vertices in C at distance exactly 2 from w, and any two such vertices are not adjacent.

If w has no neighbors in C, then, for each j, it can contribute only to $n_{2,j}$. Moreover, if for some j we have $d(w, v_j) = 2$, then $d(w, v_{j-1}) > 2$ and $d(w, v_{j+1}) > 2$ by ($\star$), and so w does not contribute to $n_{2,j}$ and $n_{2,j-2}$. Therefore, such w contributes in total by at most $k-2$.

Assume, then, that w has exactly one neighbor in C — from symmetry, let it be v_0. Because of having only one neighbor, for each j, w does not contribute to $n_{3,j}$ and $n_{k-1,j}$. In order to contribute to $n_{i,j}$ for $i \notin \{2, 3, k-1\}$, w needs to be connected to v_{i+j-1}, and so it can contribute only to $n_{1,0}$ and $n_{i,k-i+1}$ for $4 \leq i \leq k-2$. Finally, w can contribute to $n_{2,j}$ only if $d(w, v_{j+1}) = 2$ and $w \notin N(v_j)$. By ($\star$), there are at most three vertices in C at distance 2 from w, but one of them is v_1 and $w \in N(v_0)$, so w contributes to $\sum_{j=0}^{k-1} n_{2,j}$ by at most 2. It follows that in this case w contributes to the considered sum in total by at most $k-3+\frac{1}{2}$.

Finally, assume that w has exactly two neighbors in C. These neighbors have to be at distance 2 in C, as otherwise it creates an odd cycle of length shorter than k. From symmetry, let v_{k-1} and v_1 be the neighbors of w. Then, $d(w, v_i) = 2$ for $i = k-2$, 0, 2, and there are no more i with this property by ($\star$). Therefore, w contributes only to $n_{1,k-1}$, $n_{1,2}$, $n_{2,k-3}$, $n_{3,k-2}$, and $n_{i,k-i}$ for $4 \leq i \leq k-1$, hence w contributes to the considered sum in total by $k-1$. $\square$

Using the above claim, we immediately get the wanted bound $(n/k)^k$ for (3.1). It follows that the total number of k-cycles in G is at most $(n/k)^k$, as desired.

If a graph G achieves this bound, then n needs to be divisible by k and we need to have equalities in all the inequalities we considered. In particular, for each k-cycle, all the other vertices of G need to be connected with exactly two vertices of the cycle, which are at distance 2. Since we used the AM-GM inequality, all the sets $A_{k-1}(D_j)$ are of the same size n/k. Moreover, they form a partition of $V(G)$ and it is easy to see that G must be a subgraph of $C_k \odot I_{n/k}$. Since k is odd, the only k-cycles in G are those which have exactly one vertex in each $A_{k-1}(D_j)$. Thus, if G maximizes the number of induces k-cycles, it must be indeed isomorphic to $C_k \odot I_{n/k}$.

3.3 Concluding remarks and open problems

In our proof, basically the only place where we are using that k is an odd number is to say that if a k-cycle is not induced (or, more generally, there is a short path in the graph between distant vertices of this cycle), then the graph contains a smaller odd cycle. This is not the case if k is an even number. Moreover, we do not have an analogue of Theorem 3.1 for even k, as forbidding any even cycle prevents big blow-ups of a single edge. Nevertheless, one can carefully analyze the proof to obtain the following result on *induced* even cycles.

Observation 3.3. *For each even integer $k \geq 8$, any graph on n vertices without induced cycles C_ℓ for $\ell = 3$ and $5 \leq \ell \leq k-1$ and without induced copies of graphs obtained from C_6 by adding one or two main diagonals contains at most $(n/k)^k$ induced cycles of length k.*

It seems possible that the same construction (balanced blow-up of a k-cycle) gives the best possible number of induced k-cycles also if we only forbid triangles.

Conjecture 3.4. *For each integer $k \geq 5$, any triangle-free graph on n vertices contains at most $(n/k)^k$ induced cycles of length k.*

Using Flagmatic, one can numerically check that Conjecture 3.4 should hold for $k \leq 8$.

If we forbid an ℓ-cycle for some odd ℓ, and try to maximize the number of k-cycles for some odd k larger than ℓ, then it seems that taking a sequence of blow-ups of $C_{\ell+2}$ is still optimal, although these blow-ups do not necessarily need to be balanced. We state it as a conjecture.

Conjecture 3.5. *For odd integers $k > \ell \geq 3$, we have that $\mathrm{ex}(n, C_k, C_\ell)$ is asymptotically realized by a sequence of blow-ups of $C_{\ell+2}$.*

Chapter 4

Oriented graphs and compressibility

The results in this chapter are based on joint work with Andrzej Grzesik, Justyna Jaworska, Aliaksandra Novik, and Tomasz Ślusarczyk, and are currently being prepared for publication. My main contribution is the proof of Theorem 4.14 and the proof of Theorem 4.27 for $\ell \geq 4$. I also contributed to the proofs of Theorem 4.12 and Theorem 4.19. The whole content of this chapter was written by me.

4.1 Introduction

Even though we will be mostly interested in oriented graphs, we shall discuss the Turán problem for directed graphs as well. To avoid ambiguity, we introduce the following notation. For a directed graph H, let $\mathrm{ex}_{\mathrm{d}}(n, H)$ denote the maximum number of arcs in an H-free directed graph on n vertices. Similarly, if H is an oriented graph, then $\mathrm{ex}_{\mathrm{o}}(n, H)$ denotes the maximum number of arcs in an H-free oriented graph on n vertices.

First known results on Turán number of directed graphs are due to Brown and Harary [17], who determined $\mathrm{ex}_{\mathrm{d}}(n, \overrightarrow{K_k})$ and $\mathrm{ex}_{\mathrm{d}}(n, \overrightarrow{T_k})$ for all $k \geq 2$. Later, Häggkvist and Thomassen [50] found the exact value of $\mathrm{ex}_{\mathrm{d}}(n, \overrightarrow{C_k})$ for every $k \geq 3$. Brown, Erdős, and Simonovits proved that for every family $\mathcal{H}$ of directed graphs there exists a sequence $(G_n)_{n \geq 1}$ of $\mathcal{H}$-free n-vertex digraphs, each of them being a certain blow-up of some fixed directed digraph D, such that $\mathrm{ex}_{\mathrm{d}}(n, \mathcal{H}) = |E(G_n)| + o(n^2)$ [15], and showed the existence of a finite algorithm which determines all such digraphs D [16]. Even though their result does not give much information about the digraph D itself, Valadkhan [76] observed that in the case of oriented graphs one may assume that D is the largest tournament whose blow-ups do not contain any $H \in \mathcal{H}$, which leads to the following crucial definition and theorem.

Definition 4.1. Let $\mathcal{H}$ be a family of oriented graphs. The *compressibility* of $\mathcal{H}$, denoted by $\tau(\mathcal{H})$, is the smallest $k \in \mathbb{N}$ such that for every tournament T on k vertices there exists a homomorphism $H \to T$ for some $H \in \mathcal{H}$. If no such k exists, then we put $\tau(\mathcal{H}) = \infty$. For brevity, we write $\tau(H) := \tau(\{H\})$ for an oriented graph H.

Theorem 4.2 (Valadkhan [76]). *For any family $\mathcal{H}$ of oriented graphs,*

$$\mathrm{ex}_{\mathrm{o}}(n, \mathcal{H}) = \left(1 - \frac{1}{\tau(\mathcal{H}) - 1}\right) \binom{n}{2} + o(n^2).$$

Therefore, compressibility plays the same role in the context of oriented graphs as chromatic number in the context of graphs and the Erdős-Stone Theorem. In particular, determining the compressibility of a graph or a family of graphs is asymptotically equivalent to solving the respective problem on the maximum number of arcs in oriented graphs of a given order.

Here, we focus on properties of $\tau(\mathcal{H})$ when $\mathcal{H}$ has a single member. (In contrast to the chromatic number, in general $\tau(\mathcal{H})$ may differ from $\min\{\tau(H) : H \in \mathcal{H}\}$, see Example 4.5.) If an oriented graph contains a directed cycle, and therefore it cannot be mapped homomorphically to any transitive tournament, then its compressibility is infinite. Therefore, we shall consider only acyclic oriented graphs. It is also easy to notice that $\overrightarrow{T_k}$ does not contain a homomorphic image of any acyclic oriented graph with a directed path of order greater k, hence $\tau(H)$ is always at least the maximum order $p(H)$ of a directed path in H. In fact, $\tau(H)$ can grow exponentially in terms of $p(H)$ as witnessed by transitive tournaments (Example 4.4) or particular orientations of complete bipartite graphs (Proposition 4.7). Therefore, it is natural to ask, as in [76], for which families of acyclic oriented graphs the growth is polynomial, or for which the trivial lower bound is optimal, i.e. $\tau(H) = p(H)$.

We show that the compressibility of acyclic oriented graphs with out-degree at most 2 is polynomial with respect to the maximum order of a directed path (Theorem 4.14), and that the same holds for a larger out-degree bound under the additional assumption that the Erdős-Hajnal conjecture holds (Theorem 4.12). Additionally, generalizing results for the square of a path, we determine the compressibility of acyclic oriented graphs with out-degree at most 2 having restricted structure (Theorem 4.19). Finally, generalizing the result by Valadkhan [76] for acyclic orientations of cycles, we prove that the equality $\tau(H) = p(H)$ holds for oriented graphs H with restricted distances of vertices to sinks and sources (Theorem 4.27).

4.2 Basic properties of compressibility

The compressibility of some particular graphs can be easily derived, for instance for directed paths.

Example 4.3. For any $k \geq 1$, $\tau(\overrightarrow{P_k}) = k$, as every tournament on k vertices contains a copy of $\overrightarrow{P_k}$, i.e. a Hamiltonian path, while there is no homomorphism $\overrightarrow{P_k} \to \overrightarrow{T_{k-1}}$.

If in the definition of compressibility we ask for the existence of an *injective* homomorphism from H to every tournament of a given order, then we obtain the definition of a *1-color oriented Ramsey number*. See [63] for more information on this concept. As some graphs have no homomorphism into smaller oriented graphs, bounds on their compressibility follow from known bounds on their 1-color oriented Ramsey number.

Example 4.4. Since the compressibility of a transitive tournament $\overrightarrow{T_k}$ is equal to the 1-color oriented Ramsey number of $\overrightarrow{T_k}$, standard probabilistic arguments [28, 72] imply that

$$c_1 2^{k/2} \leq \tau(\overrightarrow{T_k}) \leq c_2 2^k$$

for some constants $c_1, c_2 > 0$ and any $k \geq 1$. These are essentially the best known general bounds.

In general, the compressibility of a family of graphs $\mathcal{H}$ can differ from $\min\{\tau(H) : H \in \mathcal{H}\}$ significantly.

Example 4.5. If $\mathcal{H} = \{\overrightarrow{P_{2^k}}, \overrightarrow{T_k}\}$ for any $k \geq 1$, then $\tau(\overrightarrow{P_{2^k}}) = 2^k$ and $\tau(\overrightarrow{T_k}) \geq c2^{k/2}$ for some constant $c > 0$, but $\tau(\mathcal{H}) = k$, since each tournament T on k vertices either contains $\overrightarrow{C_3}$, and therefore there exists a homomorphism $\overrightarrow{P_{2^k}} \to T$, or is transitive.

Let $p(H)$ be the order of a longest directed path in H. By Example 4.3, $p(H)$ can be equivalently defined as the smallest k for which there exists a homomorphism $H \to \overrightarrow{T_k}$. In particular, Example 4.4 implies that the compressibility $\tau(H)$ is bounded exponentially in terms of $p(H)$. This motivates the following definition.

Definition 4.6. Let $\mathcal{G}$ be a family of acyclic oriented graphs. We say that $\mathcal{G}$ is *polynomially τ-bounded* if there exist constants c, d such that for every $H \in \mathcal{G}$, we have

$$\tau(H) \leq cp(H)^d.$$

Valadkhan [76] observed that containing a large transitive tournament is not a necessary condition to have $\tau(H)$ exponentially large in terms of $p(H)$. Even forbidding $\overrightarrow{T_3}$ is not enough to guarantee polynomial τ-boundedness.

Proposition 4.7 (Valadkhan [76]). *For $n \geq 1$, let H_n be the only acyclic orientation of $K_{n,n}$ such that $p(H_n) = 2n$. Then, $\tau(H_n) \geq 2^{n/2}$.*

Note that if $\tau(H) = 2$, i.e. H is a subgraph of $\overrightarrow{K_{s,t}}$ for some $s, t \in \mathbb{N}$, Theorem 4.2 implies only that $\mathrm{ex}_{\mathrm{o}}(n, H) = o(n^2)$ and one may ask about the order of magnitude of $\mathrm{ex}_{\mathrm{o}}(n, H)$. It turns out that many results for the unoriented bipartite graphs can be reproved in the oriented case, with possibly different constants. As an example, we shall reprove Bondy-Simonovits Theorem and Kővari-Sós-Turán Theorem for oriented graphs.

Observation 4.8. *For any oriented graph H there exists a subgraph H' of H such that $|E(H')| \geq |E(H)|/4$ and H' can be mapped homomorphically into $\overrightarrow{T_2}$.*

Proof. Let H_1 be the underlying graph of H. It is a well-known fact that there exists a bipartite subgraph H_2 of H_1 such that $|E(H_2)| \geq |E(H_1)|/2$. Let H_2' be the subgraph of H that corresponds to H_2 with a bipartition (A, B); we may choose the bipartition in such a way that at least half of the arcs of H_2' go from A to B. Choose H' to be a subgraph of H_2' consisting of all arcs going from A to B. It is straightforward to see that there exists a homomorphism $H' \to \overrightarrow{T_2}$. □

Corollary 4.9 (Bondy, Simonovits [12]). *For any $k \geq 2$, let D_{2k} be the only orientation of C_{2k} which can be mapped homomorphically to $\overrightarrow{T_2}$. Then, there exists a constant $c > 0$ such that*

$$\mathrm{ex}_{\mathrm{o}}(n, D_{2k}) \leq cn^{1+\frac{1}{k}}.$$

Proof. Let H be any oriented graph on $cn^{1+\frac{1}{k}}$ vertices. By Observation 4.8, we may assume that H is a subgraph of a blow-up of $\overrightarrow{T_2}$. Let H' be the underlying graph of H. By Bondy-Simonovits Theorem [12], if $c > 0$ is large enough, there exists a copy of C_{2k} in H', which corresponds to a copy of D_{2k} in H. □

Corollary 4.10 (Kővari, Sós, Turán [58]). *For any natural numbers $s, t \geq 1$ there exists a constant $c > 0$ such that*

$$\mathrm{ex}_{\mathrm{o}}(\overrightarrow{K_{s,t}}) \leq cn^{2-1/\min(s,t)}.$$

Proof. Let H be any oriented graph on $cn^{2-1/\min(s,t)}$ vertices. By Observation 4.8 we may assume that H is bipartite with a bipartition (A, B) and that all arcs of H go from A to B. Let H' be the underlying graph of H. For $c > 0$ large enough, we conclude from Kővari-Sós-Turán [58] that there exists a copy of $K_{s,t}$ in H' with s vertices in A and t vertices in B, which corresponds to a copy of $\overrightarrow{K_{s,t}}$ in H. □

4.3 Oriented graphs with bounded out-degree

For any integer $k \in \mathbb{N}$, let $\mathcal{D}_k$ be the family of all acyclic oriented graphs with out-degree bounded by k. In this section, we consider the question whether $\mathcal{D}_k$ is polynomially τ-bounded. First, we prove that this is implied by the following conjecture.

Conjecture 4.11. *For every tournament T there exists a constant $\varepsilon > 0$ such that every tournament on n vertices contains either T or a transitive tournament on n^{ε} vertices.*

Alon, Pach, and Solymosi [1] proved that Conjecture 4.11 is equivalent to the well-known Erdős-Hajnal Conjecture [27].

Theorem 4.12. *Conjecture 4.11 implies that $\mathcal{D}_k$ is polynomially τ-bounded for every $k \in \mathbb{N}$.*

Before we prove this theorem, let us introduce the following notion. For an oriented graph H, we say that a subset $X \subseteq V(H)$ is *dominated* in H if $X \subseteq N^+(v)$ for some $v \in V(H)$. We have the following easy observation.

Observation 4.13. *For any $k \geq 2$ and any tournament T, if all k-subsets of $V(T)$ are dominated in T, then for any $H \in \mathcal{D}_k$ there exists a homomorphism $H \to T$.*

Proof. Since H is acyclic, there is an order of the vertices of H in which all the arcs are directed backwards. We embed in T the vertices of H in this order using the fact that each vertex in H has out-degree at most k and each set of k vertices in T is dominated by some vertex of T. □

Proof of Theorem 4.12. Our goal is to prove that for every $k \in \mathbb{N}$ there exists a tournament T such that for each $H \in \mathcal{D}_k$ there exists a homomorphism $H \to T$. If such T exists, then Conjecture 4.11 implies that for every $H \in \mathcal{D}_k$, each tournament on $p(H)^{1/\varepsilon}$ vertices contains a copy of either T or $\overrightarrow{T}_{p(H)}$. In both cases, it contains a homomorphic image of H. Thus, $\tau(H) \leq p(H)^{1/\varepsilon}$.

Existence of such a tournament T follows from a probabilistic argument. Let $n \in \mathbb{N}$ be large enough and T be a random tournament on n vertices. For a k-vertex subset $X \subseteq V(T)$, let A_X be the event that X is not dominated in T, i.e. there is no vertex $v \in V(T)$ such that $X \subseteq N^+(v)$. Then, $A = \bigcup_{X \in \binom{V(T)}{k}} A_X$ is the event that some k-vertex subset of $V(T)$ is not dominated by any vertex. The probability of A can be bounded as follows:

$$\mathbb{P}(A) \leq \sum_{X \in \binom{V(T)}{k}} \mathbb{P}(A_X) = \binom{n}{k} \left(1 - \left(\frac{1}{2}\right)^k\right)^{n-k} \xrightarrow{n \to \infty} 0.$$

Therefore, for large enough n, the probability of the complement of A is positive, i.e. there exists a tournament T in which every set of k vertices is dominated by some other vertex. By Observation 4.13, there exists a homomorphism $H \to T$ for any $H \in \mathcal{D}_k$. □

In the case $k = 2$, one can notice that $\overrightarrow{C_3} \odot \overrightarrow{C_3}$ satisfies the assumption of Observation 4.13. As [1, Theorem 2.1] implies that the tournament $\overrightarrow{C_3} \odot \overrightarrow{C_3}$ satisfies Conjecture 4.11 with the constant $\varepsilon = 1/148$, we have

$$\tau(H) \leq cp(H)^{148}$$

for any $H \in \mathcal{D}_2$. We prove a much better bound.

Theorem 4.14. *There exists a constant c such that for every $H \in \mathcal{D}_2$ we have*

$$\tau(H) \leq cp(H)^4.$$

Before we prove this result, recall the notion of a domination graph. Define the *domination graph* of a tournament T as the spanning subgraph $\mathrm{dom}(T)$ of T consisting of those arcs from $E(T)$ that are not dominated in T. One of the most basic properties of the domination graph is the following easy observation, the proof of which is included for completeness.

Observation 4.15. *If T is a tournament and $vw, v'w'$ are two vertex disjoint arcs from $E(\mathrm{dom}(T))$, then any arc between the sets $\{v, w\}$ and $\{v', w'\}$ completely determines the orientation of all the remaining arcs between those four vertices — either vv', $v'w$, ww', $w'v \in E(T)$, or vw', $w'w$, wv', $v'v \in E(T)$.*

Proof. If the arcs of the tournament T between the sets $\{v, w\}$ and $\{v', w'\}$ are not forming a directed cycle, then there exists a vertex such that either v and w or v' and w' are its out-neighbors, which contradicts the fact that arcs vw and $v'w'$ are not dominated. Thus, the arcs between the sets $\{v, w\}$ and $\{v', w'\}$ are forming a directed cycle. Depending on its direction, we obtain one of the two possibilities listed in the statement of the observation. □

We are ready now to prove Theorem 4.14.

Proof of Theorem 4.14. We use induction on $p(H)$. If $p(H) = 2$, then $\tau(H) = 2$. If $p(H) = 3$, then $H \to \overrightarrow{T_3}$, and since any tournament on 4 vertices contains $\overrightarrow{T_3}$, we have $\tau(H) \leq 4$. Thus, for $p(H) \leq 3$ the inequality $\tau(H) \leq cp(H)^4$ holds for any $c \geq 2$.

Let T be any tournament on $cp(H)^4$ vertices for some constant $c > 0$ and $p(H) \geq 4$. We may assume that T does not contain a copy of $\overrightarrow{T_{p(H)}}$, as otherwise there would exists a homomorphism $H \to T$.

Assume first that $\mathrm{dom}(T)$ does not contain a matching on $cp(H)^3$ vertices. By removing from T the vertices of any maximum matching in $\mathrm{dom}(T)$, we obtain a tournament T' on at least $cp(H)^4 - 2cp(H)^3$ vertices, which is greater than $c(p(H)-1)^4$ for $p(H) \geq 4$. If we let H' be the subgraph of H obtained by removing all sources in H, then $p(H') = p(H) - 1$ and we can apply the induction hypothesis to find a homomorphism $H' \to T'$. Since every pair of vertices from $V(T')$ is dominated in T and the maximum out-degree of H is at most two, we can extend this homomorphism to $H \to T$. Therefore, we may assume that there exists a subgraph M of $\mathrm{dom}(T)$ which is a matching on at least $cp(H)^3$ vertices.

From Observation 4.15, it follows that for every arc $vw \in E(M)$ and every other vertex $u \in V(M)$, either $vu, uw \in E(T)$ or $wu, uv \in E(T)$. Therefore, if we pick one vertex from each arc in $E(M)$ and denote by T_M the subtournament of T induced by those vertices, then T_M can be considered as equipped with a special operation of *flipping* a vertex, i.e. reversing the orientations of all arcs incident to this vertex. Indeed, this operation corresponds to replacing this vertex by its neighbor in M.

We want to prove that there exists a subgraph of T_M isomorphic to $\overrightarrow{C_3} \odot \overrightarrow{C_3}$, because this implies that $H \to T$ by Observation 4.13. Note that $\overrightarrow{C_3} \odot \overrightarrow{C_3}$ consists of three clusters, each being a copy of $\overrightarrow{C_3}$. If we flip all vertices from one cluster, then this cluster will remain a copy of $\overrightarrow{C_3}$, but arcs between this cluster and remaining ones will reverse, resulting in a subgraph isomorphic to $\overrightarrow{T_3} \odot \overrightarrow{C_3}$. Therefore, it is enough to prove that every tournament on cn^3 vertices contains a copy of $\overrightarrow{T_3} \odot \overrightarrow{C_3}$ or $\overrightarrow{T_n}$. As $\overrightarrow{T_3} \odot \overrightarrow{C_3}$ is isomorphic to $(\overrightarrow{C_3} \Rightarrow \overrightarrow{C_3}) \Rightarrow \overrightarrow{C_3}$, we force its appearance in two steps using the following claim.

Claim 4.16. *For any oriented graph D and constants $c_0, \delta > 0$, if every $\overrightarrow{T_n}$-free tournament on $c_0 n^\delta$ vertices contains a copy of D, then there exists $c > 0$ such that every $\overrightarrow{T_n}$-free tournament on $cn^{\delta+1}$ vertices contains a copy of $D \Rightarrow \overrightarrow{C_3}$.*

Proof. Let T' be any $\overrightarrow{T_n}$-free tournament on $an^{\delta+1}$ vertices for $a \geq \max(3\sqrt[4]{8c_0}, 6)$. Assume additionally that T' contains at most $n^{3\delta+2}$ copies of $\overrightarrow{C_3}$. We want to find a lower bound for the number t' of copies of $\overrightarrow{T_1} \Rightarrow \overrightarrow{C_3}$ in T'. As every tournament on $an^{\delta+1}$ vertices contains at least $an^{\delta+1}/3$ vertices of out-degree at least $an^{\delta+1}/3$, we may choose the source of $\overrightarrow{T_1} \Rightarrow \overrightarrow{C_3}$ among those $an^{\delta+1}/3$ vertices. Now, since every $\overrightarrow{T_n}$-free tournament on $2n$ vertices contains at least n copies of $\overrightarrow{C_3}$, we can count the number of subsets of size $2n$ in the out-neighborhood restricted to $\lceil an^{\delta+1}/3 \rceil$ vertices and obtain

$$t' \geq \frac{an^{\delta+1}}{3} \cdot \frac{n\binom{\lceil an^{\delta+1}/3\rceil}{2n}}{\binom{\lceil an^{\delta+1}/3\rceil - 3}{2n-3}} = \frac{an^{\delta+2}}{3} \cdot \frac{\binom{\lceil an^{\delta+1}/3\rceil}{3}}{\binom{2n}{3}} \geq \frac{an^{\delta+2}}{3} \cdot \frac{(an^{\delta+1})^3}{(2n)^3 \cdot 3^3} = \frac{a^4 n^{4\delta+2}}{2^3 \cdot 3^4},$$

as every copy of $\overrightarrow{T_1} \Rightarrow \overrightarrow{C_3}$ will be counted this way at most $\binom{\lceil an^{\delta+1}/3\rceil - 3}{2n-3}$ times. Since there are at most $n^{3\delta+2}$ copies of $\overrightarrow{C_3}$ in T', there exists a copy of $\overrightarrow{C_3}$ which is dominated by at least

$$\frac{t'}{n^{3\delta+2}} \geq \frac{a^4}{2^3 \cdot 3^4} n^\delta \geq c_0 n^\delta$$

vertices of T'. Since any subtournament of T' of order at least $c_0 n^\delta$ contains a copy of D, we conclude that the tournament T' contains the desired copy of $D \Rightarrow \overrightarrow{C_3}$.

In order to prove the claim, consider any $\overrightarrow{T_n}$-free tournament T on $cn^{\delta+1}$ vertices for some $c \geq \max(3^4a^3c_0, 3a)$. From the previous paragraph, we may assume that every subtournament on $an^{\delta+1}$ vertices contains at least $n^{3\delta+2}$ copies of $\overrightarrow{C_3}$. By the same counting argument, we get that the number t of copies of $\overrightarrow{T_1} \Rightarrow \overrightarrow{C_3}$ in T satisfies

$$t \geq \frac{cn^{\delta+1}}{3} \cdot \frac{n^{3\delta+2}\binom{\lceil cn^{\delta+1}/3\rceil}{an^{\delta+1}}}{\binom{\lceil cn^{\delta+1}/3\rceil-3}{an^{\delta+1}-3}} = \frac{cn^{4\delta+3}}{3} \cdot \frac{\binom{\lceil cn^{\delta+1}/3\rceil}{3}}{\binom{an^{\delta+1}}{3}} \geq \frac{cn^{4\delta+3}}{3} \cdot \frac{(cn^{\delta+1})^3}{(an^{\delta+1})^3 \cdot 3^3} = \frac{c^4n^{4\delta+3}}{a^3 \cdot 3^4}.$$

Since there are at most $c^3n^{3\delta+3}$ copies of $\overrightarrow{C_3}$ in T, there exists a copy of $\overrightarrow{C_3}$ that is dominated by at least

$$\frac{t}{c^3n^{3\delta+3}} \geq \frac{c}{a^3 \cdot 3^4} \cdot n^\delta \geq c_0 n^\delta$$

vertices of T. Thus, T contains the desired copy of $D \Rightarrow \overrightarrow{C_3}$. □

Applying the above claim for $n = p(H)$, $D = \overrightarrow{C_3}$, $\delta = 1$, and $c_0 > 1$, and then for $D = (\overrightarrow{C_3} \Rightarrow \overrightarrow{C_3})$ and $\delta = 2$ we conclude that the tournament T_M on $cp(H)^3$ vertices contains a copy of $\overrightarrow{T_3} \odot \overrightarrow{C_3}$ or $\overrightarrow{T}_{p(H)}$, which ends the proof of Theorem 4.14. □

For certain subclasses of $\mathcal{D}_k$, it is possible to find homomorphisms into tournaments of even linear order. For instance, Draganić et al. proved the following result for powers of paths.

Theorem 4.17 (Draganić et al. [23])**.** *For every $n, k \geq 2$, every tournament on n vertices contains a k-th power of a directed path of order $n/2^{4k+6}k + 1$. Moreover, for $k = 2$, every tournament on n vertices contains a square of a directed path of order $\lceil 2n/3 \rceil$ and this value is optimal.*

A square of a directed path, considered in Theorem 4.17, is an oriented graph obtained from a directed path by adding arcs between vertices at distance 2. A generalization of this structure is an oriented graph obtained from a directed path by adding arcs between vertices at some different distance.

Definition 4.18. For any $2 \leq \ell < k$, let $\overrightarrow{P_k}(\ell)$ be the oriented graph on k vertices $v_1, \ldots, v_k$ with arcs v_iv_{i+1} for $1 \leq i \leq k-1$ and $v_iv_{i+\ell}$ for $1 \leq i \leq k-\ell$. In other words, $\overrightarrow{P_k}(\ell)$ is a directed path on k vertices with additional arcs between vertices at distance ℓ. Let also $\overrightarrow{C_k}(\ell)$ be the oriented graph on k vertices $w_0, w_1, \ldots, w_{k-1}$ with arcs $w_iw_{i+1 \pmod k}$ and $w_iw_{i+\ell \pmod k}$ for $0 \leq i < k$.

As $\overrightarrow{P_k}(\ell)$ is a subgraph of the ℓ-th power of $\overrightarrow{P_k}$, Theorem 4.17 implies that $\tau(\overrightarrow{P_k}(\ell))$ is linear in terms of $p(\overrightarrow{P_k}(\ell)) = k$. But the constant provided in Theorem 4.17 for large ℓ is very far from being optimal. The following theorem closes this gap and shows that for $\ell = 2$ and 3, the compressibility of $\overrightarrow{P_k}(\ell)$ differs from the compressibility of $\overrightarrow{P_k}$.

Theorem 4.19. *For every $2 \leq \ell < k$, the following holds*

- $\tau(\overrightarrow{P_k}(2)) = \lfloor \frac{3k-1}{2} \rfloor$,
- $\lfloor \frac{7k-1}{6} \rfloor \leq \tau(\overrightarrow{P_k}(3)) \leq 3k$,
- $\tau(\overrightarrow{P_k}(\ell)) = k$ *if* $\ell \geq 4$.

Proof. For $\ell = 2$, the graph $\overrightarrow{P_k}(2)$ is just a square of a path, and Theorem 4.17 implies that every tournament on $\lfloor \frac{3k-1}{2} \rfloor$ vertices contains a copy of $\overrightarrow{P_k}(2)$. On the other hand, there are tournaments on $\lfloor \frac{3k-1}{2} \rfloor - 1$ vertices that do not have a homomorphism from $\overrightarrow{P_k}(2)$. For odd k, we consider the tournament $\overrightarrow{P}_{(k-1)/2} \odot \overrightarrow{C_3}$, while for even k consider the tournament $\overrightarrow{T_1} \Rightarrow (\overrightarrow{P}_{k/2-1} \odot \overrightarrow{C_3})$. The considered tournaments have exactly $\lfloor \frac{3k-1}{2} \rfloor - 1$ vertices and any homomorphism of $\overrightarrow{P_k}$ into them maps some three consecutive vertices into a copy of $\overrightarrow{C_3}$, which cannot happen for the homomorphism of $\overrightarrow{P_k}(2)$.

If $\ell \geq 4$, then $\tau(\overrightarrow{P_k}(\ell)) \geq k$ as there exists no homomorphism $\overrightarrow{P_k}(\ell) \to \overrightarrow{T_{k-1}}$. To prove the upper bound, consider any tournament T on k vertices. Then, T admits a decomposition $T_1 \Rightarrow \ldots \Rightarrow T_m$ into strongly connected components. If any of those components is of size at least $\ell - 1$, then it contains a copy of $\overrightarrow{C_\ell}$, and since there is a homomorphism $\overrightarrow{P_k}(\ell) \to \overrightarrow{C_{\ell-1}}$, we have $\overrightarrow{P_k}(\ell) \to T$. Otherwise, all strongly connected components are of size strictly smaller than $\ell - 1$. This means that any function that maps the Hamiltonian path of $\overrightarrow{P_k}(\ell)$ into any Hamiltonian path of T induces a homomorphism $\overrightarrow{P_k}(\ell) \to T$.

We are left with the hardest case $\ell = 3$. To prove the lower bound, consider a tournament $\widetilde{T}$ on 7 vertices $v_1, \ldots, v_7$, with arcs $v_i v_j$ for $1 \leq i < j \leq 6$ and $N^+(v_7) = \{v_1, v_2, v_4\}$, see Figure 4.1. We want to prove that there exists no homomorphism $\overrightarrow{P_7}(3) \to \widetilde{T}$. This implies that there exists no homomorphism of $\overrightarrow{P_{6a+1}}(3) \to \overrightarrow{T_a} \odot \widetilde{T}$ for any integer $a \geq 1$ and the claimed lower bound follows.

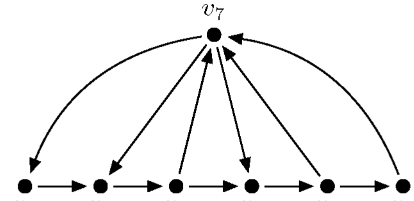


Figure 4.1: Tournament $\widetilde{T}$ from the proof of Theorem 4.19 for $\ell = 3$. The bottom vertices induce a transitive tournament.

Assume that $x_1, x_2, \ldots, x_7$ are the images of consecutive vertices of $\overrightarrow{P_7}(3)$ under some homomorphism $\overrightarrow{P_7}(3) \to \widetilde{T}$. As the vertices $v_1, \ldots, v_6$ induce a transitive tournament, there must exist the smallest i such that $x_i = v_7$. If $i = 1$, then since $x_1 x_4$ is an arc and $x_1 x_2 x_3 x_4$ is a path, we must have $x_4 = v_4$. But then it is not possible to find a path $x_4 x_5 x_6 x_7$ with an arc $x_4 x_7$. If $2 \leq i \leq 4$, then similarly $x_{i+3} = v_4$, hence $x_{i+2} \in \{v_1, v_2, v_3\}$. But since $x_{i-1} x_i$ is an arc, we have $x_{i-1} \in \{v_3, v_5, v_6\}$ and it is not possible for $x_{i-1} x_{i+2}$ to be an arc. If $5 \leq i \leq 6$, then by a symmetric argument we conclude that $x_{i-3} = v_3$, $x_{i-2} \in \{v_4, v_5, v_6\}$ and $x_{i+1} \in \{v_1, v_2, v_4\}$, hence $x_{i-2} x_{i+1}$ cannot be an arc. Finally, if $i = 7$, then we must have $x_j = v_j$ for every $1 \leq j \leq 7$, but in this case $x_4 x_7$ is not an arc. This finishes the proof of the lower bound.

In order to prove the upper bound, we apply the following theorem that characterizes the general structure of the domination graphs of tournaments. Here, by a *directed caterpillar* we mean a directed path with possible additional outgoing pendant arcs.

Theorem 4.20 (Fisher et al. [34]). *The domination graph of a tournament is either an odd directed cycle with possible outgoing pendant arcs and isolated vertices, or a forest of directed caterpillars.*

We prove by induction on k that for every tournament on $3k$ vertices there exists a homomorphism from $\overrightarrow{P_k}(3)$. For $k \leq 3$ an oriented graph $\overrightarrow{P_k}(3)$ is just a directed path $\overrightarrow{P_k}$, which can be mapped homomorphically into any tournament on k vertices (Example 4.3).

For $k > 3$, let T be any tournament on $3k$ vertices. Note that $\overrightarrow{P_k}(3) \to \overrightarrow{C_5}(3)$, so we may assume that T does not contain $\overrightarrow{C_5}(3)$. Denote vertices of $\overrightarrow{P_k}(3)$ by $w_1, \ldots, w_k$ with arcs of the form $w_i w_{i+1}$ and $w_i w_{i+3}$. Whenever we use the induction hypothesis to obtain a homomorphism $\overrightarrow{P_{k-1}}(3) \to T$, we think of this $\overrightarrow{P_{k-1}}(3)$ as of a subgraph of $\overrightarrow{P_k}(3)$ induced by vertices $w_2, w_3, \ldots, w_k$. In particular, in order to find a homomorphism $\overrightarrow{P_k}(3) \to T$, we only need to map w_1 to a vertex dominating the images of w_2 and w_4. This is possible exactly when the images of w_2 and w_4 induce an arc which does not belong to $E(\mathrm{dom}(T))$.

It turns out that if $\mathrm{dom}(T)$ contains a cycle of length at least five, two caterpillars, or a caterpillar with a directed path of length at least three, then T must contain $\overrightarrow{C_5}(3)$. It follows from the following observation.

Observation 4.21. *If* $\operatorname{dom}(T)$ *contains two vertex disjoint arcs, whose sources are not connected by an arc in* $\operatorname{dom}(T)$*, then* T *contains a copy of* $\overrightarrow{C_5}(3)$.

Proof. Let vw and $v'w'$ be the two arcs in $\operatorname{dom}(T)$, and without loss of generality let $v'v \in E(T) \setminus E(\operatorname{dom}(T))$. By Observation 4.15, all arcs between vw and $v'w'$ are then completely determined. Moreover, since $v'v \notin E(\operatorname{dom}(T))$, there exists a vertex u which dominates $v'v$, in particular it is neither w nor w'. Since vw and $v'w'$ are not dominated, we have that wu, $w'u \in E(T)$. Now, it is straightforward to check that vertices v, w, v', w' and u, in this order, induce a copy of $\overrightarrow{C_5}(3)$, as depicted in Figure 4.2. □

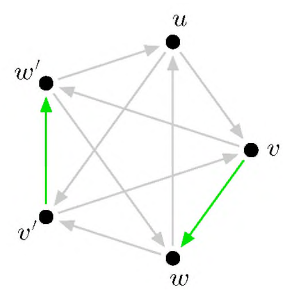


Figure 4.2: $\overrightarrow{C_5}(3)$ created in T using Observation 4.21. Green arcs belong to $E(\operatorname{dom}(T))$.

By Theorem 4.20 and Observation 4.21, $\operatorname{dom}(T)$ must be either a directed triangle with some outgoing arcs or a directed caterpillar with a longest directed path of length at most 2. In particular, there exist at most three vertices with a positive out-degree in $\operatorname{dom}(T)$, hence it is possible to find a subset $D \subseteq V(T)$ of size at most 3 such that each arc from $E(\operatorname{dom}(T))$ is incident to at least one vertex from D. Let T' be the subtournament of T induced by $V(T) \setminus D$. Since $|V(T')| \geq 3(k-1)$, by the induction hypothesis there exists a homomorphism $\overrightarrow{P_{k-1}}(3) \to T'$. Moreover, the arc induced by the images of w_2 and w_4 cannot belong to $E(\operatorname{dom}(T))$, hence we can extend this homomorphism to $\overrightarrow{P_k}(3) \to T$. □

4.4 Compressibility of ℓ-layered graphs

In this section, we study a class of acyclic oriented graphs H for which $\tau(H) = p(H)$. The considered class contains in particular graphs $\overrightarrow{P_k}(\ell)$ for $\ell \geq 4$, for which the equality holds by Theorem 4.19, as well as some graphs with out-degree not bounded by $p(H)$. It also generalizes the results of Valadkhan [76] for orientations of trees and cycles.

Definition 4.22. We say that an acyclic oriented graph H is *ℓ-layered* if for every vertex $v \in V(H)$ which is not a sink nor a source there exists a pair $(i, j) \in \mathbb{Z}_\ell^2$ such that the length of every directed path from any source of H to v is congruent to i modulo ℓ and the length of every directed path from v to any sink of H is congruent to j modulo ℓ. If a vertex v was assigned a pair (i, j), we will say that it is of type (i, j).

For $\ell \geq 2$, let $\mathcal{L}_\ell$ denote the family of all ℓ-layered acyclic oriented graphs.

Example 4.23. For any $3 \leq \ell < k$, the graph $\overrightarrow{P_k}(\ell)$ is $(\ell - 1)$-layered.

Example 4.24. Consider an acyclic oriented graph H and some integer $\ell \geq 2$, and replace each arc uv of H by a directed path of length ℓ from u to v. Then, the resulting graph, also called an $(\ell - 1)$-subdivision of H, is ℓ-layered.

Example 4.25. For any integers $k \geq 3$ and $\ell \geq 2$, each acyclic orientation of a cycle on k vertices is ℓ-layered.

Example 4.26. An acyclic oriented graph obtained from a directed path $v_1 v_2 \dots v_k$ by adding a new vertex v and an arc $v_{k-2}v$ is not ℓ-layered for any $\ell \geq 2$. It follows from the fact that the distance from v_1 to v is $k-2$, while from v_1 to v_k it is $k-1$. On the other hand, it is easy to observe that any acyclic orientation of a tree can be mapped homomorphically to some directed path, which is ℓ-layered for every $\ell \geq 2$.

Since the oriented graph in Proposition 4.7 is 2-layered, the class $\mathcal{L}_2$ is not polynomially τ-bounded. However, for $\ell \geq 3$ the situation is completely different.

Theorem 4.27. *Let $\ell \geq 3$ and $H \in \mathcal{L}_\ell$ with $p(H) \geq 6$. Then, $\tau(H) = p(H)$.*

Proof. Firstly, observe that H can be mapped homomorphically into $\overrightarrow{C_\ell} \Rightarrow \overrightarrow{T_1}$. Indeed, if we denote the consecutive vertices of $\overrightarrow{C_\ell}$ by $w_0, w_1, \dots, w_{\ell-1}$ and the only vertex of $\overrightarrow{T_1}$ by w, then we can define a map $H \to \overrightarrow{C_\ell} \Rightarrow \overrightarrow{T_1}$ in the following way: assign every source of H to w_0, every sink of H to w, and every vertex of type (i,j) to w_i. It is straightforward to check that this is indeed a homomorphism.

If T' is any tournament on 5 vertices containing a copy of $\overrightarrow{C_5}$, then some vertex of T' is contained in a copy of $\overrightarrow{C_3}$ and a copy of $\overrightarrow{C_4}$. Thus, there is a homomorphism $\overrightarrow{C_\ell} \to T'$ for any $\ell \geq 3$. In particular, there always exists a homomorphism $H \to T' \Rightarrow \overrightarrow{T_1}$. An analogous argument shows that there always also exists a homomorphism $H \to \overrightarrow{T_1} \Rightarrow T'$.

Fix now a tournament T on $p(H)$ vertices. Assume that T is not strongly connected. If at least one strongly connected component is of size at least $\min(5, \ell)$, then there exists a homomorphism $H \to T$ by the observation above. Therefore, we may assume that all strongly connected components of T are of size smaller than $\min(5, \ell)$. For each $v \in V(H)$, let $\ell(v)$ denote the length of any longest directed path in H starting at v. Choose any Hamiltonian path P in T with vertices in order $v_{p(H)-1}, \dots, v_0$. Since every strongly connected component of T is of size smaller than ℓ, we have $v_i v_j \in E(T)$ for any $i - j > \ell$. Define a map $H \to T$ by assigning each $v \in V(H)$ to $v_{\ell(v)}$. Since for each arc $vw \in E(H)$ we have either $\ell(v) - \ell(w) = 1$ or $\ell(v) - \ell(w) > \ell$, it follows that this map is indeed a homomorphism.

Since any strongly connected tournament on $p(H)$ vertices contains a strongly connected subtournament on 6 vertices, it is enough to show that there exists a homomorphism from H to any strongly connected tournament on 6 vertices.

Let us introduce the following tournaments on 5 vertices:

- T_a, obtained from $\overrightarrow{C_5}(3)$ by reversing the arc $w_1 w_4$;
- T_b, obtained from $\overrightarrow{C_5}(2)$ by reversing the arc $w_1 w_4$;
- T_c, obtained from $\overrightarrow{T_5}$ by reversing the arc between the sink and the source;
- T_d, obtained from T_c by reversing the arc $w_3 w_5$;
- T_e, obtained from T_c by reversing the arc $w_2 w_4$.

All of them are depicted in Figure 4.3. Let $\mathcal{T} = \{T_a, T_b, T_c, T_d, T_e\}$. By showing a series of claims we will prove that every strongly connected tournament on 6 vertices contains some tournament from $\mathcal{T}$, and that there exists a homomorphism from H to any tournament in $\mathcal{T}$.

Claim 4.28. *Every strongly connected tournament on 5 vertices is isomorphic to $\overrightarrow{C_5}(2)$ or some $T \in \mathcal{T}$.*

Proof. Let T be a strongly connected tournament on 5 vertices $w_1, \dots, w_5$ with arcs $w_5 w_1$ and $w_i w_{i+1}$ for $1 \leq i \leq 4$. If there are no vertices in T with out-degree equal to 3, then $d^+(w_i) = 2$ for every $1 \leq i \leq 5$ and T is isomorphic to $\overrightarrow{C_5}(2)$ (since $\overrightarrow{C_5}(2)$ and $\overrightarrow{C_5}(3)$ are isomorphic).

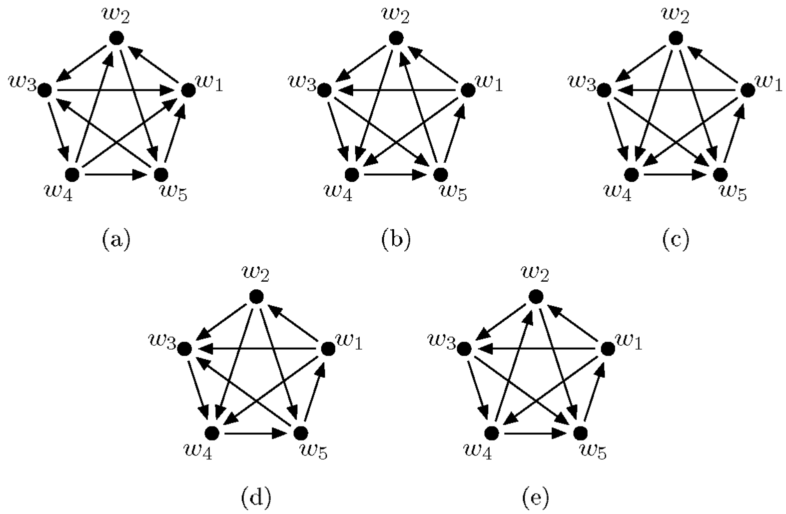


Figure 4.3: Tournaments used in the proof of Theorem 4.27.

Assume now that there is exactly one vertex v in T with out-degree 3. Then, there is also exactly one vertex w with out-degree 1. If $vw \in E(T)$, then by reversing an arc vw we obtain a tournament T' with all vertices having out-degree 2, hence T' is isomorphic to $\overrightarrow{C_5}(3)$ and T is isomorphic to either T_a or T_b. If $wv \in E(T)$, then the three remaining vertices of T are in out-neighborhood of v and in-neighborhood of w. They must induce a copy of $\overrightarrow{C_3}$, since v is the only vertex with out-degree 1. But then, T is isomorphic to T_e.

We are left with the case when there are two vertices with out-degree 3. It is easy to see that they must be neighbors in a copy of $\overrightarrow{C_5}$ contained in T, which determines all but one arc in T. Depending on the orientation of this remaining arc, we conclude that T is isomorphic either to T_c or to T_d. □

Claim 4.29. *Every strongly connected tournament on 6 vertices contains a copy of some $T \in \mathcal{T}$.*

Proof. By Claim 4.28, it is enough to find a strongly connected subtournament with a vertex of in-degree or out-degree equal to 3. Let T be any strongly connected tournament on 6 vertices. It must contain a copy of $\overrightarrow{C_5}$ and vertices of this copy induce a strongly connected subtournament T'. If T' is isomorphic to some element of $\mathcal{T}$, then we are done. Otherwise, by Claim 4.28, it must be isomorphic to $\overrightarrow{C_5}(2)$; let $w_1, \ldots, w_5$ be consecutive vertices of the outer directed cycle of T', and let w denote the remaining vertex of T. Since T is strongly connected, w has in-neighbors and out-neighbors in T'; without loss of generality, we may assume that $w_1 w, ww_2 \in E(T)$. If $ww_4 \in E(T)$, then the subtournament T_1 induced by vertices w, w_2, w_3, w_4, w_1 is strongly connected and in-degree of w_4 in T_1 is equal to 3. If $w_4 w \in E(T)$, then the subtournament T_2 induced by vertices w, w_2, w_4, w_5, w_1 is strongly connected and out-degree of w_4 is equal to 3. In both cases, T_1 or T_2 is isomorphic to some element of $\mathcal{T}$, which finishes the proof. □

To simplify the proof that H has a homomorphism to each $T \in \mathcal{T}$, we want to construct an oriented graph Q_ℓ such that H can be mapped homomorphically into Q_ℓ and then for each T provide a homomorphism from Q_ℓ. For every $0 \leq i < \ell$, let D_i be a directed cycle on a vertex set $\{(j, i-j) \in \mathbb{Z}_\ell^2 : 0 \leq j < \ell\}$ with arcs from $(j, i-j)$ to $(j+1, i-j-1)$ for every $0 \leq j < \ell$. Define Q_ℓ as a disjoint union of D_i, over all $0 \leq i < \ell$, and two additional vertices v_s, v_t, with arcs joining v_s to v_t, v_s to $(1, i)$, and v_t from $(i, 1)$ for all $0 \leq i < \ell$. Since the graph H is ℓ-layered, we have a natural homomorphism $H \to Q_\ell$ which maps all sources of H to v_s, all sinks of H to v_t, and all vertices of type (i, j) to the vertex (i, j) for every pair $(i, j) \in \mathbb{Z}_\ell^2$.

Claim 4.30. *Let T be a tournament on at least 5 vertices. Assume there exist vertices $u, v \in V(T)$ such that $uv \in E(T)$ and:*

- *$vw, wu, uz, zv \in E(T)$ for some $w, z \in V(T)$ and w is contained in a copy of $\overrightarrow{C_3}$,*
- *$ux, xy, yv \in E(T)$ for some $x, y \in V(T)$ and an arc xy is contained in a copy of $\overrightarrow{C_3}$ and a copy of $\overrightarrow{C_4}$.*

If $\ell = 3$ or $\ell = 4$, then there exists a homomorphism $Q_\ell \to T$.

Proof. Start defining the homomorphism $Q_\ell \to T$ by assigning v_s to u and v_t to v. It remains to define homomorphism $D_i \to T$ for every $0 \leq i < \ell$ such that the image of $(1, i)$ is in out-neighborhood of u and the image of $(i, 1)$ is in the in-neighborhood of v. Assign $(1, 1)$ to z, $(1, 2)$ to x, and $(2, 1)$ to y. If $\ell = 3$, then assign $(0, 1)$ to u and $(1, 0)$ to v. If $\ell = 4$, then assign $(0, 1)$ and $(1, 1)$ to z, $(3, 1)$ to u, and $(1, 3)$ to v. All of these assignments are depicted in Figure 4.4. Using the assumptions in the claim, it is straightforward to check that this can be extended to a homomorphism $Q_\ell \to T$. □

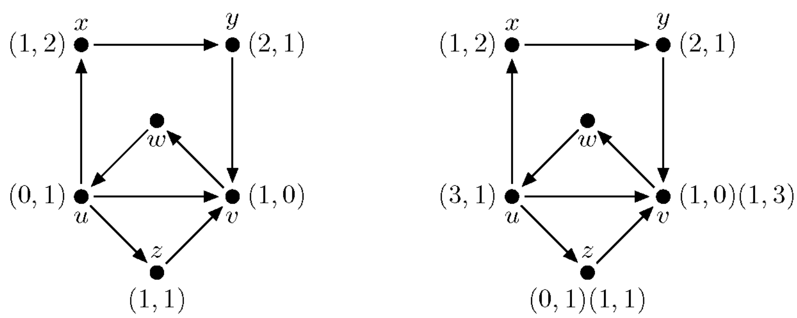


Figure 4.4: Partial homomorphisms $Q_3 \to T$ and $Q_4 \to T$ from the proof of Claim 4.30. The arc xy is assumed to be contained in a copy of $\overrightarrow{C_4}$ and a copy of $\overrightarrow{C_3}$, while the vertex z is assumed to be contained in a copy of $\overrightarrow{C_3}$.

Claim 4.31. *For every $3 \leq \ell \leq 5$ and every $T \in \mathcal{T}$, there exists a homomorphism $Q_\ell \to T$.*

Proof. If $\ell = 3$ or $\ell = 4$, it is enough for every $T \in \mathcal{T}$ to find vertices $u, v \in V(T)$ satisfying the assumptions of Claim 4.30. It is easy to verify that one can choose:

- w_4 as u and w_5 as v for T_a,
- w_1 as u and w_4 as v for T_b,
- w_1 as u and w_4 as v for T_c,
- w_5 as u and w_3 as v for T_d,
- w_1 as u and w_4 as v for T_e.

Consider now $\ell = 5$. Note that for every $T \in \mathcal{T}$ there is a copy of $\overrightarrow{C_5}$ with consecutive vertices $w_1, w_2, \ldots, w_5$, and denote it by C_T. Each D_i for $0 \leq i < 5$ can be mapped homomorphically into C_T in five different ways. We claim that for every $T \in \mathcal{T}$ there exists a homomorphism $Q_5 \to T$ that maps each D_i into C_T. Note that if the image of v_s is of out-degree k, then for every $0 \leq i < 5$ there are k choices for a homomorphism $D_i \to C_T$ that agrees with v_s, and if the image of v_t is of in-degree k', then there are k' choices for a homomorphism $D_i \to C_T$ that agrees with v_t. Moreover, for every $T \in \{T_a, T_b, T_c, T_d\}$ there exist vertices $u, v \in V(T)$ such that $d^+(u) = 3$, $d^-(v) = 3$, and $uv \in E(T)$. Therefore, if we choose u as the image of v_s and v as the image of v_t, then for each $0 \leq i < 5$ there exists a homomorphism $D_i \to C_T$ agreeing with v_s and v_t simply by the pigeonhole principle. Finally, for T_e it is straightforward to verify that one can choose w_1 as the image of v_s and w_4 as the image of v_t. $\square$

Claims 4.29 and 4.31 together imply for every $3 \leq \ell \leq 5$ that Q_ℓ can be mapped into any strongly connected tournament on 6 vertices. Hence, to finish the proof of the theorem, it is enough to show that for $\ell \geq 6$ the graph Q_ℓ also can be mapped homomorphically into every $T \in \mathcal{T}$. Note that for every $T \in \mathcal{T}$, each vertex of T is contained in a copy of $\overrightarrow{C_3}$. Therefore, if $v_1v_2v_3v_4$ is a directed path in D_i for some $0 \leq i < \ell$ and neither v_2 nor v_3 are neighbors of v_s or v_t in Q_ℓ, we can aim to find a homomorphism $D_i \to T$ which maps v_1 and v_4 to the same vertex of T, thus essentially reducing the length of D_i by 3. Since we can always perform this operation as long as the length of the cycle is at least 6, we can reduce the problem to the case $\ell \leq 5$, which was proved in Claim 4.31. $\square$

Note that the assumed bound $p(H) \geq 6$ in Theorem 4.27 cannot be improved. Indeed, $\overrightarrow{C_5}(2)$ does not contain two vertices u and v with paths of length 1, 2 and 3 from u to v, so the oriented graph H consisting of paths of lengths 1, 2, 3 and 4 with common endpoints is ℓ-layered with $p(H) = 5$ and $\tau(H) \geq 6$. Analogous constructions can be provided for $p(H) = 4$ and $p(H) = 3$. In the cases $p(H) \leq 5$ one can easily show that the best bounds are $\tau(H) \leq 2$ when $p(H) = 2$, $\tau(H) \leq 4$ when $p(H) = 3$, and $\tau(H) \leq 6$ when $p(H) \in \{4, 5\}$ and H is ℓ-layered.

4.5 Concluding remarks

It is straightforward to construct, for any $k > 0$, a sequence $(H_n)_{n\geq 1}$ of acyclic oriented graphs $H_n \in \mathcal{D}_k$ such that $p(H_n) = n$ and for every $H \in \mathcal{D}_k$ there exists a homomorphism $H \to H_{p(H)}$. Therefore, to understand the asymptotic behavior of the compressibility of acyclic oriented graphs with out-degree at most k, it suffices to examine the sequence $(H_n)_{n\geq 1}$. However, even for $k = 2$ we were able to compute $\tau(H_n)$ only for a few initial values of n, and we were unable to find a lower bound for $\tau(H_n)$ better than linear.

Let T be a tournament on 11 vertices $v_0, \dots, v_{10}$ with arcs v_iv_{i+j} for $j \in \{1, 3, 4, 5, 9\}$ and indices taken modulo 11. One can verify that every copy of $\overrightarrow{C_3}$ in T is dominated by some vertex, hence every $H \in \mathcal{D}_3$ can be mapped homomorphically into $T \odot \overrightarrow{C_3}$. Therefore, to prove that $\mathcal{D}_3$ is polynomially τ-bounded, it suffices to show that T satisfies Conjecture 4.11. It would be interesting to prove Conjecture 4.11 for this graph, especially with some low exponent.

Problem 4.32. *For which acyclic oriented graphs F is the family of F-free acyclic oriented graphs polynomially τ-bounded?*

Theorem 4.14 shows that it holds for $F = \overrightarrow{K_{1,3}}$. Also, by Proposition 4.7, if the family of F-free acyclic oriented graphs is polynomially τ-bounded, then F must be bipartite.

The following definitions and notation are taken from [71]. We say that an oriented graph H is an *o-clique* if every two vertices of H are joined by a directed path of length at most 2. Define the *absolute oriented clique number* of H, denoted by $\omega_{ao}(H)$, as the maximum size of an o-clique contained in H, and the *relative oriented clique number* of H, denoted by $\omega_{ro}(H)$, as the maximum size of a subset $S \subseteq V(H)$ such that every two vertices of S are joined in H by a directed path of length at most 2. It is clear that if H is an o-clique and T is a tournament, then any homomorphism $H \to T$ must be injective, and for a general oriented graph H we have $\omega_{ao}(H) \leq \omega_{ro}(H) \leq |V(T)|$. For $k \geq 3$, let $\mathcal{A}_k$ denote the family of all acyclic oriented graphs with absolute clique number at most k, and let $\mathcal{R}_k$ denote the family of all acyclic oriented graphs with relative clique number at most k. We have $\mathcal{R}_k \subseteq \mathcal{A}_k$ and one may observe that $\mathcal{D}_k \subseteq \mathcal{R}_{k^2+1}$.

Conjecture 4.33. *For $k \geq 3$, the families $\mathcal{A}_k$ and $\mathcal{R}_k$ are polynomially τ-bounded.*

Chapter 5

Generalized Turán-type problems for directed cycles

The results in this chapter are based on joint work with Andrzej Grzesik, Justyna Jaworska, Piotr Kuc, and Tomasz Ślusarczyk, and are currently being prepared for publication. My main contribution is the proof of Theorem 5.4. I also contributed to the proofs of Theorems 5.2, 5.3, 5.5, 5.6, and 5.7. The whole content of this chapter was written by me.

5.1 Introduction

In Chapter 3, we defined the generalized Turán number $\mathrm{ex}(n, T, H)$ as the maximum number of copies of T in an H-free graph on n vertices. Depending on the choice of T and H, the behavior of this quantity and the extremal constructions vary significantly, and even in the case when both T and H are cycles, we still did not reach full understanding of the problem.

In this chapter, we shall study this problem in the setting of oriented graphs. It is natural to define $\mathrm{ex}_{\mathrm{o}}(n, T, H)$ for oriented graphs T and H as the maximum number of copies of T in an H-free n-vertex oriented graph. For $T = \overrightarrow{T_2}$, we get just $\mathrm{ex}_{\mathrm{o}}(n, H)$, which was discussed in Chapter 4, and we are not aware of any published results for T other than $\overrightarrow{T_2}$. Still, some general results for unoriented graphs hold for oriented graphs as well. For instance, the proof of Proposition 2.1 in [3] translates immediately to the following proposition.

Proposition 5.1 (Alon, Shikhelman [3]). *Let T be a fixed oriented graph on k vertices. Then, $\mathrm{ex}_{\mathrm{o}}(n, T, H) = \Theta(n^k)$ if and only if there exists no homomorphism $H \to T$. Otherwise, we have $\mathrm{ex}_{\mathrm{o}}(n, T, H) \leq n^{k-\varepsilon(T,H)}$ for some constant $\varepsilon(T, H) > 0$.*

On the other hand, if we do not forbid any substructure, we end up with a problem of maximizing the number of copies of some fixed oriented graph T in n-vertex tournaments, which was considered in particular for directed cycles. Let $C(n, k)$ be the maximum number of copies of $\overrightarrow{C_k}$ in an n-vertex tournament, and $R(n, k)$ be the expected number of copies of $\overrightarrow{C_k}$ in a random n-vertex tournament. Define also $c(k) = \lim_{n\to\infty} C(n, k)/R(n, k)$. The value of $C(n, 3)$ was determined already by Kendall and Babington Smith [55], and independently by Szele [73]. Later, $C(n, 4)$ was determined by Beineke and Harary [8], and independently by Colombo [21]. Recently, Komarov and Mackey [56] proved that $c(5) = 1$. Bartley [7] and Day [22] conjectured that $c(k) = 1$ if and only if k is not divisible by 4, and they showed that $c(k) > 1$ for $4 \mid k$. Grzesik et al. [43] proved this conjecture and determined the value of $c(k)$ up to a small error term $o(1)$.

In the following, we shall investigate $\mathrm{ex}_{\mathrm{o}}(n, T, H)$ when both T and H are directed cycles. By Proposition 5.1, we may distinguish the sparse case when $T = \overrightarrow{C_k}$ and $H = \overrightarrow{C_{kt}}$ for $k \geq 3$ and $t \geq 2$, and the dense case when $T = \overrightarrow{C_k}$ and $H = \overrightarrow{C_\ell}$ for $k \nmid \ell$.

In the sparse case, which we discuss in Section 5.2, our first result states that $\varepsilon(\overrightarrow{C_k}, \overrightarrow{C_{kt}}) = 1$ in Proposition 5.1.

Theorem 5.2. *Let $k \geq 3$ and $t \geq 2$. Then,* $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_k}, \overrightarrow{C_{kt}}) = \Theta(n^{k-1})$.

We also find the exact asymptotics of $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_3}, \overrightarrow{C_6})$.

Theorem 5.3. *We have* $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_3}, \overrightarrow{C_6}) = n^2/4 + o(n^2)$.

In the dense case $k \nmid \ell$, which we discuss in Section 5.3, we determine asymptotically $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$ for $k \geq 6$ and sufficiently large ℓ such that k is odd or ℓ is even.

Theorem 5.4. *Let $k \geq 6$ and $\ell > 2k^2 - 4k + 1$. Assume that $2 \nmid k$ or $2 \mid \ell$. Define $d > 2$ as the smallest divisor of k which does not divide ℓ. Then,* $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell}) = \frac{n}{k} \cdot \left(\frac{n}{d}\right)^{k-1} + o(n^k)$.

We also determine $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$ asymptotically for $k \in \{3, 4, 5\}$ and $\ell > k$ not divisible by k.

Theorem 5.5.

a) $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_3}, \overrightarrow{C_\ell}) = \lceil \frac{n}{3} \rceil \lceil \frac{n-1}{3} \rceil \lceil \frac{n-2}{3} \rceil$ *for* $\ell = 4,\ 5$.

b) $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_3}, \overrightarrow{C_\ell}) = \frac{n^3}{27} + O(n^2)$ *for $\ell \geq 7$ not divisible by 3.*

Theorem 5.6. *For $\ell \geq 5$ not divisible by 4, we have* $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_4}, \overrightarrow{C_\ell}) = \left(\frac{n}{4}\right)^4 + o(n^4)$.

Theorem 5.7.

a) $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_5}, \overrightarrow{C_7}) = \frac{27}{16} \cdot \left(\frac{n}{5}\right)^5 + o(n^5)$.

b) $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_5}, \overrightarrow{C_\ell}) = \left(\frac{n}{5}\right)^5 + o(n^5)$ *for $\ell = 6$ and $\ell \geq 8$ not divisible by 5.*

5.2 Sparse case

In this section, we shall prove results regarding $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$ when $k \mid \ell$. Let us start by proving that $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_k}, \overrightarrow{C_{kt}}) = \Theta(n^{k-1})$ for $k \geq 3$ and $t \geq 2$.

Proof of Theorem 5.2. For the lower bound, just observe that if we substitute $I_{\lfloor \frac{n-1}{k-1} \rfloor}$ for every but one vertex of $\overrightarrow{C_k}$ and we add $n - (k-1)\lfloor \frac{n-1}{k-1} \rfloor$ isolated vertices, then we obtain a $\overrightarrow{C_{kt}}$-free oriented graph on n vertices with at least $\lfloor \frac{n}{k-1} \rfloor^{k-1}$ copies of $\overrightarrow{C_k}$.

For the upper bound, let G be any extremal graph on n vertices. We will say that an arc is *thick* if there exist at least k^2tn^{k-3} different copies of $\overrightarrow{C_k}$ containing this arc. Repeat the following procedure — as long as there exists an arc in G which is not thick, remove this arc from G. This way, we will remove $O(n^{k-1})$ copies of $\overrightarrow{C_k}$ from G and either every arc in G will be thick or G would have no arcs.

Assume that the theorem does not hold, and so there exist thick arcs in G. We will consider two cases.

Case 1: $t \leq k-1$.

Let $v, w \in V(G)$ be any two vertices of G. We claim that every two directed paths of length t which start in v and end in w share at least one internal vertex. Otherwise, let Q_1 and Q_2 be directed paths of length t that share only the endpoints. Then, since each arc of G is thick, we can find t arc-disjoint distinct copies $D_1, \ldots, D_t$ of $\overrightarrow{C_k}$, each containing exactly one arc of Q_1 and sharing no vertex outside of Q_1, which are vertex disjoint from $Q_2 - \{v, w\}$. But then, there exists a copy of $\overrightarrow{C_{kt}}$ in $D_1 \cup \ldots \cup D_t \cup Q_2$, which is a contradiction.

In particular, if we pick any two vertices v and w of G, then there exist at most $(t-1)^2 n^{t-2}$ paths of length t from v to w. To see this, just fix any path Q from v to w and observe that any other path must share at least one internal vertex with Q.

Now, we can count copies of $\overrightarrow{C_k}$ in G in the following way. First, we choose two vertices $v, w \in V(G)$. Then, we choose a directed path of length t from v to w. Finally, we choose remaining vertices to close the cycle. This way, we obtain at most

$$n^2(t-1)^2 n^{t-2} n^{k-t-1} = O(n^{k-1})$$

different copies of $\overrightarrow{C_k}$. Since each copy of $\overrightarrow{C_k}$ was counted at least once, we get the desired upper bound.

Case 2: $t \geq k$.

We prove by induction on t that $\mathrm{ex}_{\mathrm{o}}(\overrightarrow{C_k}, \overrightarrow{C_{kt}}) = O(n^{k-1})$. The basis of induction was handled in Case 1.

We claim that there are no copies of $\overrightarrow{C_{kt-k(k-2)}}$ in G. Otherwise, let C be such a copy and $vw \in E(C)$. Since vw is thick, we can find a directed path Q of length $k-1$ from w to v which is vertex disjoint from $C - \{v, w\}$. In the same way, we can find $k-1$ directed paths $Q_1, \ldots, Q_{k-1}$, such that the sum $Q_1 \cup \ldots \cup Q_{k-1}$ is a directed path from v to w of length $(k-1)^2$ which is vertex disjoint from $C - \{v, w\}$. But then, $C \cup Q_1 \cup \ldots \cup Q_{k-1}$ would contain a copy of $\overrightarrow{C_m}$, where $m = kt - k(k-2) - 1 + (k-1)^2 = kt$, a contradiction.

Since G is $\overrightarrow{C_{kt-k(k-2)}}$-free, we get by induction that it can have at most $O(n^{k-1})$ copies of $\overrightarrow{C_k}$, which finishes the proof.

$\square$

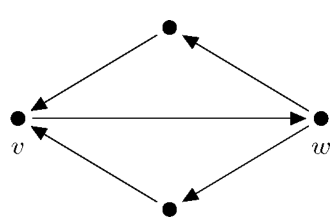
We shall also prove that $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_3}, \overrightarrow{C_6}) = n^2/4 + o(n^2)$.

Proof of Theorem 5.3. Let G be any extremal oriented graph on n vertices. We may assume that each arc of G is contained in at least one copy of $\overrightarrow{C_3}$. We will say that an arc of G is *thin* if it is contained in exactly one copy of $\overrightarrow{C_3}$; otherwise, we will say it is *thick*.

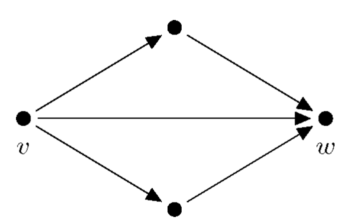
The crucial observation is that each copy of $\overrightarrow{C_3}$ in G contains at least one thin arc. Indeed, if there was a copy of $\overrightarrow{C_3}$ consisting only of thick arcs vw, wu, and uv, then we could find three more vertices v', w', u' in G such that $vw'u$, $uv'w$, and $wu'v$ are directed 3-cycles, and it would result in a directed 6-cycle $vw'uv'wu'$.

Since each copy of $\overrightarrow{C_3}$ contains at least one thin arc and any such arc cannot belong to any other copy of $\overrightarrow{C_3}$, we conclude that $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_3}, \overrightarrow{C_6})$ is bounded from above by the number of thin arcs in G.

Let vw be any thin arc. Observe that the following configurations are not allowed as subgraphs of G:



blow-up of a $\overrightarrow{C_3}$



blow-up of a $\overrightarrow{T_3}$

First configuration is not allowed just by the definition of a thin arc. For the second configuration, one can find two copies of $\overrightarrow{C_3}$, each sharing one of two upper arcs, to form a copy of $\overrightarrow{C_6}$.

Consider a subgraph G_0 of G consisting only of thin arcs of G. By Lemma 2.6, we can remove all copies of $\overrightarrow{C_3}$ and $\overrightarrow{T_3}$ from G_0 by removing $o(n^2)$ arc. But then, its underlying graph would be C_3-free, hence by Mantel's Theorem [64] it would contain at most $n^2/4$ edges. Therefore, G_0 has at most $n^2/4 + o(n^2)$ arcs, which finishes the proof. □

5.3 Dense case

In this section, we shall prove results regarding $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$ for $k \nmid \ell$. The situation is simple for $k = 3$, since there exists a construction which is asymptotically extremal for every value of ℓ.

Proof of Theorem 5.5. Balanced blow-up of $\overrightarrow{C_3}$ on n vertices contains $\lceil \frac{n}{3} \rceil \lceil \frac{n-1}{3} \rceil \lceil \frac{n-2}{3} \rceil$ copies of $\overrightarrow{C_3}$ and is $\overrightarrow{C_\ell}$-free for any $\ell \geq 4$ not divisible by 3, which gives the lower bound on $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_3}, \overrightarrow{C_\ell})$.

For the upper bound, we shall prove first the following claim.

Claim 5.8. *For any $n \geq 1$, we have $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_3}, \overrightarrow{T_3}) = \lceil \frac{n}{3} \rceil \lceil \frac{n-1}{3} \rceil \lceil \frac{n-2}{3} \rceil$.*

Proof. Since balanced blow-ups of $\overrightarrow{C_3}$ are $\overrightarrow{T_3}$-free, it is enough to prove the upper bound. Let G be any extremal oriented graph on n vertices and let G' be its underlying graph. Then, G' is K_4-free, since any orientation of K_4 contains a copy of $\overrightarrow{T_3}$. In particular, $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_3}, \overrightarrow{T_3}) \leq \mathrm{ex}(n, K_3, K_4)$. By the results of Erdös [26] and Zykov [79], $\mathrm{ex}(n, K_3, K_4) = \lceil \frac{n}{3} \rceil \lceil \frac{n-1}{3} \rceil \lceil \frac{n-2}{3} \rceil$, and the claim follows. □

Let $\ell \in \{4, 5\}$ and take any $\overrightarrow{C_\ell}$-free oriented graph G on n vertices. We may assume that each arc of G is contained in some copy of $\overrightarrow{C_3}$. By Claim 5.8, it is enough to show that G is $\overrightarrow{T_3}$-free. But if T was a copy of $\overrightarrow{T_3}$ in G, then we could find for each arc of T a copy of $\overrightarrow{C_3}$, and the union of those copies would contain a copy of $\overrightarrow{C_4}$ and a copy of $\overrightarrow{C_5}$. This finishes the proof for $\ell \in \{4, 5\}$.

Let now $\ell \geq 7$. We shall prove by induction that $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_3}, \overrightarrow{C_\ell}) = n^3/27 + O(n^2)$. For any $\overrightarrow{C_\ell}$-free oriented graph G on n vertices, let G' denote the subgraph of G obtained by removing successively all arcs of G which are contained in at most ℓ copies of $\overrightarrow{C_3}$. This way, we removed at most $\ell n^2/2$ copies of $\overrightarrow{C_3}$. As in Case 2 of the proof of Theorem 5.2, one can show that G' is $\overrightarrow{C_{\ell-3}}$-free, hence the result follows by induction. □

Let us now discuss the case $k = 4$. Since every copy of $\overrightarrow{C_4}$ must be induced in $\overrightarrow{C_3}$-free oriented graphs, $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_4}, \overrightarrow{C_3})$ is not greater than the maximum number of induced copies of $\overrightarrow{C_4}$ in an n-vertex oriented graph. The latter is shown by Hu et al. [52] to be asymptotically maximized by iterated blow-ups of $\overrightarrow{C_4}$, which are also $\overrightarrow{C_3}$-free. Hence, $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_4}, \overrightarrow{C_3}) = n^4/252 + o(n^4)$.

For $\ell \geq 5$ not divisible by 4, we shall prove that $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_4}, \overrightarrow{C_\ell}) = (n/4)^4 + o(n^4)$, which is attained e.g. by balanced blow-ups of $\overrightarrow{C_4}$.

Proof of Theorem 5.6. We need only to show the upper bound. Let G be any $\overrightarrow{C_\ell}$-free oriented graph on n vertices. By Lemma 2.6, we may remove from G all homomorphic images of $\overrightarrow{C_\ell}$ by removing at most $o(n^2)$ arcs, hence by removing at most $o(n^4)$ copies of $\overrightarrow{C_4}$. In particular, we may assume that $5 \leq \ell \leq 7$.

Case $\ell = 5$ or $\ell = 7$

We may assume that each arc in G is contained in some copy of $\overrightarrow{C_4}$. But then, G must be $\overrightarrow{T_3}$-free, as otherwise we would find a copy of a homomorphic image of $\overrightarrow{C_\ell}$. Therefore, it is enough to prove the following.

Claim 5.9. *For every $n \geq 1$, we have $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_4}, \overrightarrow{T_3}) \leq (n/4)^4$.*

Proof. Proof is based on a method developed by Kral', Norin, and Volec [59]. For any directed 4-cycle $v_0v_1v_2v_3$ contained in G, by a ***good sequence*** we mean a sequence $D = (z_i)_{i=0}^3$, where $z_0 = v_1$, $z_1 = v_0$, $z_2 = v_2$, and $z_3 = v_3$, i.e. v_0 and v_1 are in the reversed order. Note that there are 4 different good sequences corresponding to a single 4-cycle.

For a fixed good sequence D, we define the following sets:

$$\begin{aligned} A_0(D) &= V(G), \\ A_1(D) &= N^-(z_0), \\ A_2(D) &= N^+(z_0), \\ A_3(D) &= N^+(z_2) \cap N^-(z_1). \end{aligned}$$

Define the ***weight*** $w(D)$ of a good sequence D as

$$w(D) = \prod_{i=0}^{3} |A_i(D)|^{-1} = \frac{1}{n} \prod_{i=1}^{3} |A_i(D)|^{-1}.$$

Recall (from Chapter 3) that this quantity has the following probabilistic interpretation. Suppose we want to sample four vertices $w_0, \ldots, w_3$ so that $(w_i)_{i=0}^3$ is a good sequence. We start with choosing w_0 at random from all vertices of G. Next, we pick some in-neighbor of w_0 to be w_1. In general, w_j is a random vertex from the set $A_j((w_i)_{i=0}^{j-1})$. (Note that the definition of $A_j(D)$ depends only on first j elements of a sequence D.) Then, $w(D)$ is just the probability that the sequence $(w_i)_{i=0}^3$ obtained in this random process is equal to D.

In particular, the sum of the weights of all good sequences is at most one, since it is the sum of probabilities of pairwise disjoint events.

Fix a 4-cycle $v_0v_1v_2v_3$ in G, let $C = \{v_0, v_1, v_2, v_3\}$ be the set of its vertices, and let $D_j = (v_{j+1}, v_j, v_{j+2}, v_{j+3})$, for $0 \leq j \leq 3$, where the indices are considered modulo 4, be all good sequences corresponding to this cycle.

If we prove that

$$\left(\sum_{j=0}^{3} w(D_j) \right)^{-1} \leq M$$

for some number M, then $\sum_{j=0}^{3} w(D_j) \geq M^{-1}$. Thus, by summing over all directed k-cycles and using the fact that the sum of weights of all good sequences is at most one, we get that the total number of copies of $\overrightarrow{C_4}$ in G is bounded from above by M.

Put $n_{i,j} = |A_i(D_j)|$. Since

$$\left(\sum_{j=0}^{3} w(D_j) \right)^{-1} = \left(\sum_{j=0}^{3} \prod_{i=0}^{3} n_{i,j}^{-1} \right)^{-1} = n \left(\sum_{j=0}^{3} \prod_{i=1}^{3} n_{i,j}^{-1} \right)^{-1},$$

the maximum possible value of

$$n \left(\sum_{j=0}^{3} \prod_{i=1}^{3} n_{i,j}^{-1} \right)^{-1} \tag{5.1}$$

is an upper bound on the number of copies of $\overrightarrow{C_4}$ in G.

Using the inequality between harmonic mean and geometric mean of 4 terms and the inequality between geometric mean and arithmetic mean of $4 \cdot 3 = 12$ terms, we obtain

$$n \left(\sum_{j=0}^{3} \prod_{i=1}^{3} n_{i,j}^{-1} \right)^{-1} \leq \frac{n}{4} \left(\prod_{j=0}^{3} \prod_{i=1}^{3} n_{i,j} \right)^{\frac{1}{4}} \leq \frac{n}{4} \left(\frac{1}{12} \sum_{j=0}^{3} \sum_{i=1}^{3} n_{i,j} \right)^{3}.$$

Claim 5.10. *The following inequality holds:*

$$\sum_{j=0}^{3}\sum_{i=1}^{3} n_{i,j} \leq 3n,$$

with equality if and only if each vertex of G has exactly one in-neighbor and one out-neighbor in C, and those neighbors are not adjacent.

Proof. For any $w \in V(G)$, let $n_{i,j}(w) = 1$ if $w \in A_i(D_j)$ and $n_{i,j}(w) = 0$ otherwise. Then,

$$\sum_{j=0}^{3}\sum_{i=1}^{3} n_{i,j} = \sum_{w \in V(G)} \sum_{j=0}^{3}\sum_{i=1}^{3} n_{i,j}(w),$$

hence it is enough to show that for every $w \in V(G)$ we have

$$\sum_{j=0}^{3}\sum_{i=1}^{3} n_{i,j}(w) \leq 3. \tag{5.2}$$

The crucial observation is that w can have at most two neighbors in C, as otherwise we would find a copy of $\overrightarrow{T_3}$. Therefore,

$$\sum_{j=0}^{3} n_{3,j}(w) \leq 1$$

and

$$\sum_{j=0}^{3} (n_{2,j}(w) + n_{3,j}(w)) = d_C^+(w) + d_C^-(w) \leq 2.$$

In particular, if (5.2) is an equality, then $w \in A_3(D_j)$ for some $j \in \{0, 1, 2, 3\}$, and the claim follows. □

From the claim above we immediately obtain that the maximum value of (5.1), i.e. the maximum number of 4-cycles in G, is at most

$$\frac{n}{4}\left(\frac{1}{12}3n\right)^3 = \left(\frac{n}{4}\right)^4,$$

which finishes the proof of Claim 5.10. □

Case $\ell = 6$

Since G does not contain a copy of any homomorphic image of $\overrightarrow{C_6}$, it is in particular $\overrightarrow{C_3}$-free. Therefore, it is enough to prove the following.

Claim 5.11. *For every $n \geq 1$, we have* $\mathrm{ex}_o(n, \overrightarrow{C_4}, \{\overrightarrow{C_3}, \overrightarrow{C_6}\}) \leq (n/4)^4$.

The outline of the proof is the same as of the proof of Claim 5.9, with the following changes. First, we change the definition of a good sequence — if $v_0v_1v_2v_3$ is a directed 4-cycle, then by a good sequence we mean a sequence $D = (z_i)_{i=0}^3$, where $z_0 = v_0, z_1 = v_2, z_2 = v_1$, and $z_3 = v_3$, i.e. v_1 and v_2 are in the reversed order. Therefore, the sets $A_i(D)$ shall be defined as follows:

$$\begin{aligned}
A_0(D) &= V(G),\\
A_1(D) &= \{v \in V(G) : d_G(v, z_0) = 2 \text{ and } d_G(z_0, v) = 2\},\\
A_2(D) &= N^+(z_0) \cap N^-(z_1),
\end{aligned}$$

$$A_3(D) = N^-(z_0) \cap N^+(z_1).$$

Note that $z_1 \in A_1(D)$, since each copy of $\overrightarrow{C_4}$ in G must be induced.

We also need to reprove Claim 5.10. Recall that $n_{i,j} = |A_i(D_j)| = \sum_{v \in V(G)} n_{i,j}(w)$. It is enough to show that for each $w \in V(G)$, we have

$$\sum_{j=0}^{3} \sum_{i=1}^{3} n_{i,j}(w) \leq 3. \tag{5.3}$$

Assume that $w \in A_1(D_i)$ and $w \in A_1(D_{i'})$ for two different $i, i' \in \{0,1,2,3\}$. If $2 \mid i - i'$, then we would find a copy of a homomorphic image of $\overrightarrow{C_6}$, which must contain a copy of $\overrightarrow{C_6}$ or $\overrightarrow{C_3}$, a contradiction. Hence, we may assume that $i' \equiv i + 1 \pmod 4$. But now, if $w \in A_2(D_m)$ or $w \in A_3(D_m)$ for any $m \in \{0,1,2,3\}$, then we would find a copy of $\overrightarrow{C_3}$ in G. Therefore, the left-hand side of (5.3) for such w is not greater than 2.

Moreover, if $w \in A_2(D_i) = A_3(D_{i+2})$ for any $i \in \{0,1,2,3\}$, then $w \in A_1(D_i)$ as well; in particular, this is possible only for at most one value of i. This implies that the left-hand side of (5.3) for such w is not greater than 3, which proves Claim 5.11, and as a consequence finishes the proof of Theorem 5.6. $\square$

For $k = 5$, we have a few possible scenarios. It turns out that $\mathrm{ex}_{\circ}(n, \overrightarrow{C_5}, \overrightarrow{C_3}) = \frac{1}{512}n^5 + o(n^5)$ and it is asymptotically maximized by the blow-ups $\overrightarrow{C_4} \odot \overrightarrow{T}_{n/4}$. The proof of the upper bound can be obtained via flag algebra method by running the following code using Flagmatic.

```
from flagmatic.all import *
problem = OrientedGraphProblem(5, forbid="3:122331", density=["5:1223344551","
    5:122334455113","5:12233445511324"], types=["3:12","3:1223","3:122313"])
problem.set_extremal_construction(field=QQ, target_bound=15/64)
problem.add_sharp_graphs(149,164,168,188,193,265,267,285,312,316)
# type 3:12
problem.add_zero_eigenvectors(0,matrix(QQ,
    [(0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,1,0),
    (0,0,0,0,0,0,0,-1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0),
    (0,0,0,0,0,0,0,-1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
    (0,0,0,0,0,-1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)]))
# type 3:1223
problem.add_zero_eigenvectors(1,matrix(QQ,
    [(0,0,0,0,0,0,0,0,1,0,0,0,1,0,1,0,0,0,1,0,0),
    (0,0,0,0,0,0,0,0,-1,0,0,1,0,0,0,0,0,0,0,0,0),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,1,0),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,1,0,0,0,0,0)]))
# type 3:121323
problem.add_zero_eigenvectors(2,matrix(QQ,
    [(1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,0,0,1),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,0,0,1),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,0,1,0),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,1,0,0),
    (0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,1,1,0,0),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,0,1),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,1,0),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,1,0,0),
    (0,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,1,0,1,0),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,1),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,0,1,0),
    (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,1,0,0)]))
problem.solve_sdp()
problem.make_exact()
P.write_certificate("exC5C3.cert")
```

Let us determine $\mathrm{ex}_{\circ}(n, \overrightarrow{C_5}, \overrightarrow{C_\ell})$ for $\ell \geq 6$. The following observation from Number Theory would be useful in our considerations.

Observation 5.12 (Frobenius Coin Problem [67]). *For two relatvely prime integers $k, m \geq 1$, the largest integer which cannot be expressed as $ak + bm$, where $a, b \geq 0$ are non-negative integers, is equal to $km - k - m$.*

Proof of Theorem 5.7. Let G be any $\overrightarrow{C_\ell}$-free oriented graph on n vertices for some $\ell \geq 6$. By Lemma 2.6, we may remove from G all homomorphic images of $\overrightarrow{C_\ell}$ by removing at most $o(n^2)$ arcs, hence by removing at most $o(n^5)$ copies of $\overrightarrow{C_5}$. We may also assume that each arc in G is contained in some copy of $\overrightarrow{C_5}$.

Suppose that G contains a copy of $\overrightarrow{T_3}$. Then, since each arc is contained in a copy of $\overrightarrow{C_5}$, there exists a vertex which is contained in a copy of $\overrightarrow{C_5}$, $\overrightarrow{C_6}$, and $\overrightarrow{C_9}$. In particular, if $\ell = 5\alpha_1 + 6\alpha_2 + 9\alpha_3$ for some non-negative integers α_1, α_2, and α_3, then G does not contain a copy of $\overrightarrow{T_3}$. With the help of Observation 5.12, it is easy to verify that only $\ell \in \{7, 8, 13\}$ are not of this form. Consider the following cases.

Case $\ell \geq 6$, $\ell \notin \{7, 8, 13\}$

We already know that G is $\overrightarrow{T_3}$-free. Since $\ell = 5\alpha_1 + 6\alpha_2 + 9\alpha_3 = 5\alpha_1 + 3(2\alpha_2 + 3\alpha_3)$, we conclude that G is $\overrightarrow{C_3}$-free as well. In particular, the underlying graph of G is C_3-free, hence the number of copies of $\overrightarrow{C_5}$ in G is at most $\mathrm{ex}(n, C_5, C_3)$. The latter is known to be maximized in a balanced blow-up of C_5 [60]. Since $\overrightarrow{C_5} \odot I_{n/5}$ is $\overrightarrow{C_\ell}$-free, we have $\mathrm{ex}_o(n, \overrightarrow{C_5}, \overrightarrow{C_\ell}) = (n/5)^5 + o(n^5)$.

Case $\ell \in \{7, 8, 13\}$

Introduce the following equivalence relation in $V(G)$ — vertices v and w are in the same equivalence class if and only if $N^+(v) = N^+(w)$ and $N^-(v) = N^-(w)$. It is easy to see that there is no copy of $\overrightarrow{C_5}$ that would contain two vertices from the same equivalence class. Therefore, if $[v]$ and $[w]$ are two equivalence classes such that there exists no copy of $\overrightarrow{C_5}$ containing both v and w, then it is possible to remove all vertices from one equivalence class and add the same number of vertices to the other equivalence class without decreasing the number of copies of $\overrightarrow{C_5}$ in G. Since this operation decreases the number of equivalence classes, we may assume that G is connected and for every two equivalence classes $[v]$ and $[w]$ there exists a copy of $\overrightarrow{C_5}$ in G that contains both v and w.

Let R be an oriented graph which vertices are the equivalence classes; two vertices $[v], [w] \in V(R)$ are joined by an arc if and only if $vw \in E(G)$.

For $m \geq 4$, let Q_m denote an oriented graph obtained from $\overrightarrow{C_m}$ by reversing a single arc. Since every arc in G is contained in a copy of $\overrightarrow{C_5}$, we conclude that G and R are Q_4-free if $\ell = 7$, and are Q_5-free if $\ell \in \{8, 13\}$.

Claim 5.13. *For every vertex in $V(R)$, its out-neighborhood and in-neighborhood in R are tournaments.*

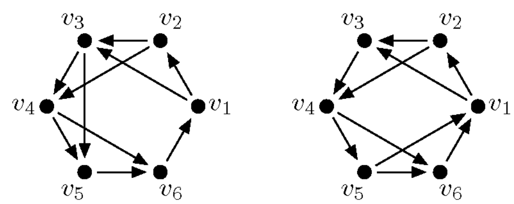
Proof. Let us present the argument for the out-neighborhood, as the same reasoning works for the in-neighborhood as well. By contradiction, assume there exist $[u], [v], [w] \in V(R)$ such that v and w are non-adjacent and $uv, uw \in E(G)$. Since there exists a copy of $\overrightarrow{C_5}$ in G containing both v and w, there exist directed paths of length 2 and 3 joining v and w. Therefore, we can find copies of homomorphic images of Q_4 and Q_5 in G, and the claim follows. $\square$

Let $C = v_1 \dots v_t$ be a directed t-cycle in G for some $t \geq 5$, which corresponds to a directed t-cycle in R. Assume that w is adjacent to C and does not belong to any equivalence class $[v_i]$. We claim that in this case, C can be extended to a directed $(t+1)$-cycle. Indeed, assume that $v_1 w \in E(G)$. (If $wv_1 \in E(G)$, we can argue by a symmetric argument.) Then, by Claim 5.13, w and v_2 are adjacent. If $wv_2 \in E(G)$, then we are done. Otherwise, $v_2 w \in E(G)$ and we can successively repeat the same argument. Therefore, either we find some $i \in [t]$ such that $v_i w, wv_{i+1} \in E(G)$, or $v_i w \in E(G)$ for every $i \in [t]$. But the second possibility cannot happen, since G would contain a copy of Q_4 and Q_5.

Start with a directed 5-cycle and extend it to the largest possible directed cycle C using the argument above. If $\ell = 7$, then the length t of C is equal to either 5 or 6.

- If $t = 5$, then it is easy to see that C must be induced, since every arc of G must be contained in a copy of $\overrightarrow{C_5}$.

- If $t = 6$ and $C = v_1 \dots v_6$, then the only possible diagonals in C are of the form $v_i v_{i+2}$. Indeed, any arc of the form $v_i v_{i+3}$ creates a copy of Q_4, and any arc of the form $v_i v_{i+4}$ forces v_{i+1} and v_{i+4} to be joined by an arc, which we already excluded. One may verify that there are two possible edge-maximal arrangements, G_1 and G_2, which are illustrated in Figure 5.1.



(a) Graph G_1 (b) Graph G_2

Figure 5.1: Oriented graphs used in the proof of Theorem 5.7.

As a consequence, G is a blow-up of $\overrightarrow{C_5}$, G_1, or G_2, but is straightforward to check that it is optimal to consider a blow-up of G_1 or G_2. More specifically, if we substitute I_{n_i} for v_i in G_1, then one should put $n_i = n/5$ for $i \in \{1, 6\}$ and $n_i = 3n/20$ for $2 \le i \le 5$, and if we substitute I_{n_i} by v_i in G_2, then one should put $n_i = n/5$ for $i \in \{1, 4\}$ and $n_i = 3n/20$ for $i \in \{2, 3, 5, 6\}$. One may compute that such blow-ups contain $\frac{27}{16}\left(\frac{n}{5}\right)^5 + o(n^5)$ copies of $\overrightarrow{C_5}$.

We are left with the case $\ell \in \{8, 13\}$. Then, the length t of C is equal to either 5, 6, or 7. Recall that G is $\overrightarrow{C_3}$-free and $\overrightarrow{C_4}$-free, hence each copy of $\overrightarrow{C_5}$ in G is induced.

- If $t = 5$, then again C must be induced.

- If $t = 6$ and $C = v_1 \dots v_6$, then the only possible diagonals in C are of the form $v_i v_{i+2}$. Moreover, C can have at most two such diagonals. Suppose that $v_2 v_4 \in E(G)$ and $v_1 v_3 \notin E(G)$. Then, since the arc $v_3 v_4$ is contained in some copy of $\overrightarrow{C_5}$ and this copy must also contain the arc $v_2 v_3$, there must exist a directed path of length 3 from v_4 to v_2. But this creates a copy of $\overrightarrow{C_4}$, a contradiction.

- If $t = 7$ and $C = v_1 \dots v_7$, then the only possible diagonals in C are of the form $v_i v_{i+2}$ or $v_i v_{i+3}$. If $v_1 v_4 \in E(G)$, then also $v_1 v_3, v_2 v_4 \in E(G)$. Moreover, since $v_2 v_4$ is contained in a copy of $\overrightarrow{C_5}$ and this copy must also contain arcs $v_4 v_5$ and $v_5 v_6$, we must have $v_7 v_2 \in E(G)$. Analogously, by considering the arc $v_1 v_3$, we conclude that $v_3 v_5 \in E(G)$. At this moment, introducing any more diagonals would result in creating a copy of $\overrightarrow{C_3}$ or $\overrightarrow{C_4}$. On the other hand, the arc $v_1 v_2$ is not contained in any copy of $\overrightarrow{C_5}$.

  Therefore, C can have only diagonals of the form $v_i v_{i+2}$, and it cannot have all diagonals of this form. Suppose that $v_2 v_4 \in E(G)$ and $v_1 v_3 \notin E(G)$. Then, since the arc $v_3 v_4$ is contained in some copy of $\overrightarrow{C_5}$ and this copy must also contain the arc $v_2 v_3$, there must exist a directed path of length 3 from v_4 to v_2. But this creates a copy of $\overrightarrow{C_4}$, a contradiction.

It follows that C is a blow-up of $\overrightarrow{C_5}$, hence G contains at most $(n/5)^5 + o(n^5)$ copies of $\overrightarrow{C_5}$, which finishes the proof. □

Finally, we prove that if $k \ge 7$ is odd or ℓ is even, then for ℓ sufficiently large, $\mathrm{ex}_{\circ}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$ is asymptotically maximized in a blow-up of $\overrightarrow{C_d}$ for a proper choice of d.

Proof of Theorem 5.4. Recall that $k \geq 6$ and $\ell > 2k^2 - 4k + 1$ is not divisible by k, and $d > 1$ is defined as the smallest divisor of k which does not divide ℓ. We also assume that k is odd or ℓ is even, which in particular implies that $d > 2$.

Let G be any $\overrightarrow{C_\ell}$-free oriented graph on n vertices. By Lemma 2.6, we may remove from G all homomorphic images of $\overrightarrow{C_\ell}$ by removing at most $o(n^2)$ arcs, hence by removing at most $o(n^k)$ copies of $\overrightarrow{C_k}$. We may also assume that each vertex and each arc in G is contained in some copy of $\overrightarrow{C_k}$.

For any integer $m > 2$, let Q_m denote the graph obtained from $\overrightarrow{C_m}$ by reversing a single arc.

Claim 5.14. *For any $3 \leq m \leq d+1$, G is Q_m-free.*

Proof. If G contains a copy of Q_m, then it also contain a copy of a homomorphic image of $\overrightarrow{C_{k+m-2}}$, which implies the existence of homomorphic images of directed cycles of length $ak + b(k+m-2)$ for any integers $a, b \geq 0$. Therefore, it is enough to show that $\ell = ak + b(k+m-2)$ for some integers $a, b \geq 0$.

Let s denote the greatest common divisor of $m-2$ and k. If s does not divide ℓ, then $s \in \{1, 2\}$, as otherwise we would get a contradiction with the definition of d. But $s \neq 2$, because we assumed that $2 \nmid k$ or $2 \mid \ell$. Therefore, we have $s = 1$. Since $\ell > 2k^2 - 4k + 1$, by Observation 5.12, ℓ can be expressed as $ak + b(k+m-2)$ for some integers $a, b \geq 0$.

On the other hand, if s divides ℓ, then ℓ/s can be expressed as $ak/s + b(k+m-2)/s$ for some integers $a, b \geq 0$, hence again $\ell = ak + b(k+m-2)$. $\square$

The rest of the proof is based on a method developed by Kral', Norin, and Volec [59], which we used already in proofs of Theorem 5.6 and 5.7.

For any directed k-cycle $v_0 \ldots v_{k-1}$ contained in G, by a ***good sequence*** we mean a sequence $D = (v_i)_{i=0}^{k-1}$.

For a fixed good sequence D, we define the following sets:

$$\begin{aligned} A_0(D) &= V(G), \\ A_i(D) &= N^+(v_{i-1}) \text{ for } 1 \leq i \leq k-2, \\ A_{k-1}(D) &= N^+(v_{k-2}) \cap N^-(v_0). \end{aligned}$$

We then define the ***weight*** $w(D)$ of a good sequence D as

$$w(D) = \prod_{i=0}^{k-1} |A_i(D)|^{-1} = \frac{1}{n} \prod_{i=1}^{k-1} |A_i(D)|^{-1}.$$

By symmetry, define

$$\begin{aligned} B_0(D) &= V(G), \\ B_i(D) &= N^-(v_{k-i}) \text{ for } 1 \leq i \leq k-2, \\ B_{k-1}(D) &= N^+(v_{k-1}) \cap N^-(v_1) \end{aligned}$$

and

$$\overline{w}(D) = \prod_{i=0}^{k-1} |B_i(D)|^{-1} = \frac{1}{n} \prod_{i=1}^{k-1} |B_i(D)|^{-1}.$$

For a sequence $D = (v_i)_{i=0}^{k-1}$, let $D_j = (v_{j+1}, v_{j+2}, \ldots, v_{j+k-1})$, where the indices are considered modulo k.

If we prove that

$$\left(\frac{1}{2} \sum_{j=0}^{k-1} (w(D_j) + \overline{w}(D)) \right)^{-1} \leq M$$

for some number M, then $\frac{1}{2}\sum_{j=0}^{k-1}(w(D_j)+\overline{w}(D)) \geq M^{-1}$. Thus, by summing over all directed k-cycles and using the fact that the sum of the weights of all good sequences is at most one, we conclude that the total number of copies of $\overrightarrow{C_k}$ in G is bounded from above by M.

Put $n_{i,j} = |A_i(D_j)|$ and $\overline{n}_{i,j} = |B_i(D_j)|$. Since

$$\left(\frac{1}{2}\sum_{j=0}^{k-1}(w(D_j)+\overline{w}(D_i)\right)^{-1} = 2n\left(\sum_{j=0}^{k-1}\left(\prod_{i=1}^{k-1} n_{i,j}^{-1} + \prod_{i=1}^{k-1}\overline{n}_{i,j}^{-1}\right)\right)^{-1},$$

the maximum possible value of

$$2n\left(\sum_{j=0}^{k-1}\left(\prod_{i=1}^{k-1} n_{i,j}^{-1} + \prod_{i=1}^{k-1}\overline{n}_{i,j}^{-1}\right)\right)^{-1} \tag{5.4}$$

is an upper bound on the number of copies of $\overrightarrow{C_k}$ in G.

Using the inequality between harmonic mean and geometric mean of 4 terms and the inequality between geometric mean and arithmetic mean of $2k(k-1)$ terms, we obtain

$$\begin{aligned} 2n\left(\sum_{j=0}^{k-1}\left(\prod_{i=1}^{k-1} n_{i,j}^{-1} + \prod_{i=1}^{k-1}\overline{n}_{i,j}^{-1}\right)\right)^{-1} &\leq \frac{n}{k}\left(\prod_{j=0}^{k-1}\prod_{i=1}^{k-1} n_{i,j}\overline{n}_{i,j}\right)^{\frac{1}{2k}} \\ &\leq \frac{n}{k}\left(\frac{1}{2k(k-1)}\sum_{j=0}^{k-1}\sum_{i=1}^{k-1}(n_{i,j}+\overline{n}_{i,j})\right)^{k-1}. \end{aligned}$$

Claim 5.15. *We have*

$$\sum_{j=0}^{k-1}\sum_{i=1}^{k-1}(n_{i,j}+\overline{n}_{i,j}) \leq 2k(k-1)n/d.$$

Proof. For any $w \in V(G)$, put $n_{i,j}(w) = 1$ if $w \in A_i(D_j)$, and $n_{i,j}(w) = 0$ otherwise; similarly, $\overline{n}_{i,j}(w) = 1$ if $w \in B_i(D_j)$, and $\overline{n}_{i,j}(w) = 0$ otherwise. Then,

$$\sum_{j=0}^{k-1}\sum_{i=1}^{k-1}(n_{i,j}+\overline{n}_{i,j}) = \sum_{w\in V(G)}\sum_{i=1}^{k-1}\sum_{j=0}^{k-1}(n_{i,j}(w)+\overline{n}_{i,j}(w)).$$

Observe that w can have at most 2 neighbors among d consecutive vertices in D, as otherwise G would contain a copy of Q_m for some $3 \leq m \leq d+1$. Therefore, w has at most $2k/d$ neighbors in D.

For any fixed $i \in [k-1]$, the sum $\sum_{j=0}^{k-1}(n_{i,j}(w)+\overline{n}_{i,j}(w))$ is equal to the number of neighbors of w in D, hence is at most $2k/d$. Also, $\sum_{j=0}^{k-1} n_{k-1,j}(w)$ is not greater than the number of in-neighbors of w in D, and $\sum_{j=0}^{k-1}\overline{n}_{k-1,j}(w)$ is not greater than the number of out-neighbors of w in D, hence $\sum_{j=0}^{k-1}(n_{k-1,j}(w)+\overline{n}_{k-1,j}(w)) \leq 2k/d$ as well. By summing over all $w \in V(G)$, we get the desired inequality. □

From the claim above, we conclude that

$$M \leq \frac{n}{k}\cdot\left(\frac{n}{d}\right)^{k-1}$$

and $\mathrm{ex}_\circ(n, \overrightarrow{C_k}, \overrightarrow{C_\ell}) \leq \frac{n}{k}\cdot\left(\frac{n}{d}\right)^{k-1} + o(n^k)$. □

5.4 Concluding remarks

The assumption in Theorem 5.4 that k is odd or ℓ is even is necessary. For instance, if we want to maximize the number of copies of $\overrightarrow{C_6}$ in oriented graphs without directed $3(2t+1)$-cycles for large enough t, then it seems optimal to consider a random orientation of a complete balanced bipartite graph. On the other hand, if $4 \mid k$ and $2 \nmid \ell$ for sufficiently large ℓ, then one should instead consider a balanced blow-up of $\overrightarrow{C_4}$. We state these observations as a conjecture.

Conjecture 5.16. *Let $k \geq 4$ be divisible by 2 and ℓ be not divisible by 2. For sufficiently large ℓ, if $4 \mid k$, then $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell}) = \frac{n}{k} \cdot \left(\frac{n}{4}\right)^{k-1} + o(n^k)$. Otherwise, $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell}) = \frac{n}{2k} \cdot \left(\frac{n}{4}\right)^{k-1} + o(n^k)$.*

It seems that if $\ell < k$, then the extremal constructions are more varied. For instance, very interesting is the case of $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_5}, \overrightarrow{C_4})$. One may check that the blow-up of $\overrightarrow{C_7}(3)$ (see Definition 4.18) contains more copies of $\overrightarrow{C_5}$ than the iterated blow-up of $\overrightarrow{C_5}$ on the same number of vertices. We conjecture that one cannot do essentially better.

Conjecture 5.17. *We have $\mathrm{ex}_{\mathrm{o}}(n, \overrightarrow{C_5}, \overrightarrow{C_4}) = \frac{n^5}{2401} + o(n^5)$.*

Chapter 6

Inducibility of graphs

The results in this chapter are based on joint work with Łukasz Bożyk and Radosław Żak, and are currently being prepared for publication. My main contribution is developing the computer search algorithm for extremal constructions, implementing it, and using it to find the constructions described in Section 6.3. The whole content of this chapter was written by me.

6.1 Introduction

Fix a graph H on k vertices. For a graph G, let $N(H,G)$ denote the number of induced copies of H in G. Define

$$i(H) = \lim_{n\to\infty} \frac{\max\{N(H,G) : |V(G)| = n\}}{\binom{n}{k}}.$$

We call $i(H)$ the *inducibility* of a graph H. We have $i(H) = 1$ for H being a complete graph or an empty graph, but in most of the remaining cases the problem of determining $i(H)$ is open. The concept of inducibility was introduced in 1975 by Pippenger and Golumbic [68], who observed that

$$i(H) \geq \lim_{n\to\infty} \frac{N(H, H^{\odot n})}{\binom{k^n}{k}} \geq \frac{k!}{k^k - k} \tag{6.1}$$

for any graph H on k vertices, and conjectured the following:

Conjecture 6.1 (Pippenger, Golumbic [68]). *If H is a cycle of length $k \geq 5$, then $i(H) = \frac{k!}{k^k-k}$.*

Brown and Sidorenko proved that the inducibility of any complete bipartite graph is realized by complete bipartite graphs.

Theorem 6.2 (Brown, Sidorenko [18]). *For any $1 \leq s \leq t$,*

$$i(K_{s,t}) = \binom{s+t}{t} \max_{x\in[0,1]} \left(x^s(1-x)^t + x^t(1-x)^s\right).$$

They also showed that the inducibility of complete multipartite graphs is realized by complete multipartite graphs, possibly with a different number of parts. This result was improved by Bollobás et al. Below, by $T_r(n)$, we mean a *Turán graph*, i.e. a graph on n vertices which is a balanced blow-up of K_r.

Theorem 6.3 (Bollobás et al. [9]).

- *For any $t \geq 2$ and $n \geq 1$, $T_3(n)$ is the only graph on n vertices that maximizes the number of induced copies of $K_3 \odot I_t$.*

- *If $r \geq 4$ and $t > 1 + \log r$, then $T_r(n)$ is the unique extremal graph for $i(K_r \odot I_t)$. If $t < \log(r+1)$ and n is large enough, then $T_r(n)$ is not extremal for $i(K_r \odot I_t)$.*

Regarding Conjecture 6.1, Balogh et al. [5] confirmed it for $k = 5$. Pippenger and Golumbic proved in their original paper a general bound on $i(C_k)$ for $k \geq 5$ within a multiplicative factor of $2e$. This was recently improved to $128e/81$ by Hefetz and Tyomkin [50] and to 2 by Král', Norin, and Volec [59]. Nevertheless, Conjecture 6.1 remains open for $k > 5$.

There are also results for more general classes of graphs. Hatami, Hirst, and Norine [48] proved that for any graph H there exists k_0 such that $i(H \odot I_k)$ is realized by a sequence of (not necessarily balanced) blow-ups of H if $k > k_0$. Later, Yuster showed that (6.1) is almost an inequality for almost every graph H in the following sense.

Theorem 6.4 (Yuster [78])**.** *For any $k \geq 1$, let $G(k, 1/2)$ be a random graph on k constructed by joining every pair of vertices with an edge independently at random with probability $1/2$. Then,*

$$\lim_{k\to\infty} \mathbb{P}\left(i(G(k,1/2)) \leq \frac{k!}{k^k - k}\left(1 + \frac{4}{k^{k^{1/3}}}\right)\right) = 1.$$

Fox, Huang, and Lee [35] claim to have proven that $\mathbb{P}\left(i(G(k,1/2)) = \frac{k!}{k^k-k}\right) \xrightarrow{k\to\infty} 1$, although the proof is still not published.

Let us turn attention to small graphs. It is a simple task to determine the inducibility of all graphs on three vertices. Hirst [51] determined the inducibility of all graphs on four vertices with a single exception of P_4.

6.2 Bounds on inducibility of P_4

How to create a large number of induced copies of P_4 in a graph? The simplest idea is to consider iterated blow-ups of P_4, which give a lower bound $i(P_4) \geq 2/21 \geq 0.0952$. However, taking iterated blow-ups of C_5 results in a much better bound $i(P_4) \geq 6/31 \geq .0.1935$. Exoo [32] realized that one can do even better by taking iterated blow-ups of G_{17}, a Paley graph on 17 vertices, which gives $i(P_4) \geq 60/307 \geq 0.1954$. The best known construction is due to Evan-Zohar and Linial [31], who showed that a sequence $K_4 \otimes (K_3 \otimes K_3)^{\odot n}$ implies $i(P_4) \geq 1173/5824 \geq 0.201407$.

For the upper bound, Exoo [32] proved that $i(P_4) \leq 1/3$, which was later improved by using the flag algebra method to $i(P_4) \leq 0.2064$ by Hirst [51] and to $i(P_4) \leq 0.204513$ by Vaughan [77].

In conclusion, we only know that $0.201407 \leq i(P_4) \leq 0.204513$ and it is not clear which of these bounds is closer to the truth. We shall propose a new construction which improves the lower bound.

6.3 The construction

We shall construct a graph G_{32} on 32 vertices with 8800 induced copies of P_4. A straightforward calculation shows that iterated blow-ups of G would imply the bound $i(P_4) \geq 6600/32767 \geq 0.201422$.

Let G be a hypercube on $2^4 = 16$ vertices with vertex set $V(G) = \mathbb{Z}_2^4$, which we shall also treat as a vector space, and with edges between the vertices which differ on exactly one coordinate. Take two copies G_1 and G_2 of a square of G. Define an automorphism $f : \mathbb{Z}_2^4 \to \mathbb{Z}_2^4$ of vector spaces by putting

$$f(0,0,0,1) = (0,0,0,1), \qquad f(0,0,1,0) = (1,0,1,0),$$
$$f(0,1,0,0) = (0,1,1,1), \qquad f(1,0,0,0) = (1,1,1,0).$$

Define G_{32} as a disjoint union of G_1 and G_2 with the following additional edges — if $v \in G_1$, then its neighborhood in G_2 is equal to $\{f(v) + w : w \in A\}$, where

$$A = \{(0,0,0,0), (0,0,0,1), (0,1,0,1), (0,1,1,1), (1,0,0,0), (1,0,1,1)\}.$$

Graph G_{32} is shown in Figure 6.1. One can verify by writing a computer program that G_{32} has indeed 8800 induced copies of P_4.

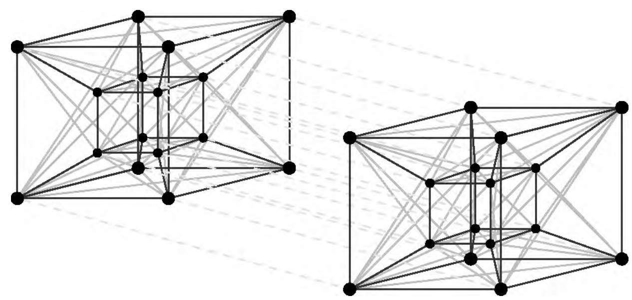


Figure 6.1: Graph G_{32}, which iterated blow-ups imply the new lower bound on $i(P_4)$. The actual edges between hypercubes are given by the function f.

6.4 Comments and remarks

We are fairly convinced that taking iterated blow-ups is not the most effective way of constructing large graphs with many induced copies of P_4. This can be observed by searching for extremal graphs on a relatively small number n of vertices. It seems that iterated blow-ups of C_5 are not extremal for any $n \geq 17$. For instance, for $n = 17$, the best subgraph of an iterated blow-up of C_5 has only 506 induced copies of P_4, while the Paley graph on 17 vertices has 680 of them. There exists a graph on 18 vertices, which is a particular Cayley graph of a direct product $S_3 \times \mathbb{Z}_3$, whose iterated blow-ups imply the bound $i(P_4) \geq 1152/5831 \geq 0.1975$. The smallest graph we found whose iterated blow-ups imply $i(P_4) \geq 0.2$ has 21 vertices and is not a Cayley graph, but is self-complementary. Finally, graph G_{32} is neither self-complementary nor a Cayley graph, but still possesses a quite large number of symmetries, since its automorphism group is of order 320.

Based on this evidence, one may reach the conclusion that the more vertices we consider, the better constructions may be involved. This can also explain the large gap between the lower and the upper bound on $i(P_4)$, since flag algebras consider only densities of very small graphs (the number of vertices almost never exceeds 10 in „real-life" applications) because of the computational complexity.

Since we found our constructions by the computer program, we shall also outline the algorithm which we used in the computer search. The idea is to start from a random graph on a fixed number of vertices. Then, apply the following greedy *saturation procedure* — as long as removing or adding a single edge from the graph results in increasing the number of induced copies of P_4, remove or add this edge. If the resulting graph, denoted by G, is extremal, terminate. Otherwise, repeatedly apply the following procedure. First, randomly perturbate a small fraction of the edges of G. Next, apply the saturation procedure and let G' denote the obtained graph. If G' has more induced copies of P_4 than G, put $G := G'$.

The random perturbing of the edges is necessary in order to escape from local maxima, which the algorithm encounters rather frequently. Also, we usually choose a random regular graph as the initial graph, since the graphs with a large number of induced copies of P_4 seem to be regular or almost regular as well.

Chapter 7

Inducibility of oriented graphs

The results in this chapter are based on joint work with Łukasz Bożyk and Andrzej Grzesik, and are published in the article Ł. Bożyk, A. Grzesik, and B. Kielak: *On the inducibility of oriented graphs on four vertices*, Discrete Math. **345(7)** (2022), 112874. My main contribution is the proof of the upper bound for Graph 24. I also contributed to finding other constructions presented in Section 7.3. I rewrote the introduction and adjusted the content of Section 7.2 in comparison to the respective sections in [14]. Section 7.3 is the respective section from [14] with no substantial changes.

7.1 Introduction

The concept of inducibility, discussed in Chapter 6 for undirected graphs, can be also introduced in the setting of oriented graphs.

Fix an oriented graph H on k vertices. For an oriented graph G, let $N(H,G)$ denote the number of induced copies of H in G. Define

$$i(H) = \lim_{n\to\infty} \frac{\max\{N(H,G) : |V(G)| = n\}}{\binom{n}{k}}.$$

We call $i(H)$ the *inducibility* of an oriented graph H. Even though we use the same notation as in the undirected case, it should not lead to any confusion.

In general, little is known about the inducibility of oriented graphs. Huang determined the inducibility of directed stars and of complete bipartite digraphs.

Theorem 7.1 (Huang [54]).

- *For every $k \geq 3$,*

$$i(\overrightarrow{S_k}) = \max_{0\leq x\leq 1} \frac{kx(1-x)^{k-1}}{1-x^k}.$$

- *For integers $2 \leq s \leq t$,*

$$i(\overrightarrow{K_{s,t}}) = \binom{s+t}{t}\left(\frac{s}{s+t}\right)^s\left(\frac{t}{s+t}\right)^t,$$

which is achieved by the balanced blow-up $\overrightarrow{K}_{\frac{s}{s+t}n,\frac{t}{s+t}n}$ of $\overrightarrow{K_{s,t}}$.

Recently, Hu et al. [53] determined the inducibility of all other orientations of stars on at least 7 vertices. There is also an analogue of Conjecture 6.1 for oriented graphs:

Conjecture 7.2. *If H is a directed cycle of length $k \geq 4$, then $i(H) = \frac{k!}{k^k-k}$.*

The conjectured value is achieved in the sequence of iterated blow-ups of a directed k-cycle. The case $k = 4$ was proved by Hu et al. [52], who also determined the inducibility of all other orientations of C_4. Note that the case $k = 5$ follows from the result on undirected cycles [5]. Choi, Lidický, and Pfender [20] made a similar conjecture that the inducibility of a directed path on $k - 1$ vertices is also achieved in the sequence of iterated blow-ups of a directed k-cycle.

The inducibility of all oriented graphs on 3 vertices is known. For a directed triangle $\overrightarrow{C_3}$ it is $1/4$ and is achieved by a sequence of any regular tournaments. The inducibility of a graph with one arc is $3/4$, since it is the same as the inducibility of the undirected complement of $K_{1,2}$ solved in [68], and is attained by a sequence of disjoint unions of two arbitrary tournaments of equal size. The inducibility of a directed star $\overrightarrow{S_3}$ was determined by Huang to be $2\sqrt{3} - 3$ (Theorem 7.1) and is achieved by an iterative construction. The case of a directed path $\overrightarrow{P_3}$ was announced to be solved by Hladký, Král' and Norin [20] — the inducibility is equal to $2/5$ and is achieved in the sequence of iterated blow-ups of a directed 4-cycle.

We shall consider the inducibility of all oriented graphs on four vertices. Up to isomorphism, there are 42 such graphs. Since $i(H) = i(\overleftarrow{H})$, where $\overleftarrow{H}$ is the graph obtained from H by reversing all arcs, the number of non-isomorphic cases to consider can be reduced to 30.

In Section 7.3 we present upper bounds and constructions providing lower bounds for all oriented graph on four vertices. The results for the directed star [54], all 4-vertex tournaments [19], and all orientations of C_4 [52] were known before. The results for some other cases follow from the results on the inducibility of undirected graphs [31]. In the remaining cases, the upper bounds were obtained mostly using Flagmatic. In several cases, the presented constructions give lower bounds that are matching the upper bounds, while in the remaining ones, when the construction is complex, the lower bounds differ by at most 0.004 in one case, and by at most 0.001 in all the other cases. This indicates that the constructions might be optimal and the applied flag algebras computations were not sufficient to obtain the matching upper bounds. All of the results are summarized in Table 7.1. Whenever the constant is irrational, it is defined in the appropriate subsection of Section 7.3. Description of the used notation and explanations of the pictograms applied to illustrate the constructions are contained in Section 7.2.

It is worth mentioning that the obtained results indicate the structure of constructions giving the inducibility to be far richer than the intuition suggests. Yuster [78] and independently Fox, Huang, and Lee [35] proved that for almost all graphs H, the inducibility is attained by the iterated blow-ups of H. For small graphs the situation is different. Out of 28 non-trivial non-isomorphic graphs considered here, only 2 of them have this property, whereas such constructions as those in Subsections 7.3.4, 7.3.20, or 7.3.23 show that the inducibility can be attained by very sophisticated and complex structures.

7.2 Preliminaries

We shall introduce the following notation in addition to the one defined in Chapter 2.

For any $\alpha \in [0, \frac{1}{2}]$, define the *circular* graph $S^1(\alpha)$ as the oriented graph with vertex set $[0, 1)$ and arcs from x to $x + a \pmod 1$ for each $x \in [0, 1)$ and each $a \in [0, \alpha)$. Let T_n^{reg} denote any regular tournament on n vertices. For an undirected graph G, a random oriented graph G^{rand} is obtained from G by orienting every edge of G independently at random. If G and H are oriented graphs, we write $G \overset{p}{\Rightarrow} H$ to denote a random oriented graph constructed from $G \sqcup H$ by joining each vertex of G with each vertex of H by an arc independently at random with probability p. For $p = 1$, we just write $G \Rightarrow H$, which coincides with the notation introduced in Chapter 2.

While depicting constructions corresponding to obtained lower bounds, we use the following conventions. First of all, the illustrations show the limit structure of each construction instead of the finite graphs forming a sequence giving the lower bound. This allows us to see the structure of this sequence without caring about getting integer sizes of particular clusters. Each cluster is assigned a real number from the interval $(0, 1)$ which corresponds to the fraction of vertices present in that cluster in the limit (in a corresponding n-vertex graph, a cluster of size α is assumed to

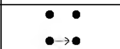
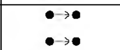
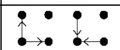
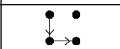
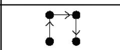
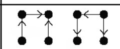
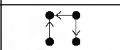
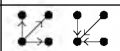
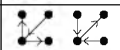
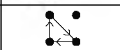
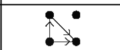
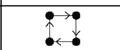
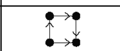
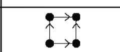
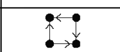
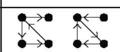
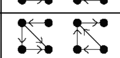
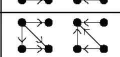
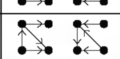
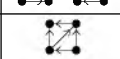
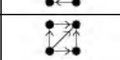
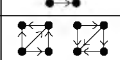
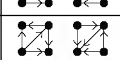
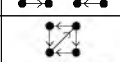
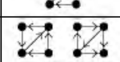
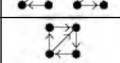
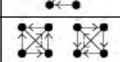
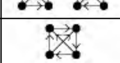
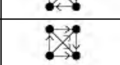
Id	$H, \overleftarrow{H}$	Upper bound		Lower bound		
				Value	Approximation	Construction
1		1	=	1	1	I_n
2		72/125	=	72/125	0.576	$I_5 \odot \overrightarrow{T_n}$
3		3/8	=	3/8	0.375	$\overrightarrow{T_n} \sqcup \overrightarrow{T_n}$
4		≈ 0.235046	>	c_4	0.234309	G^4
5		≈ 0.204123	>	64/315	0.203175	$\overrightarrow{C_4}^{\odot n} \sqcup \overrightarrow{C_4}^{\odot n}$
6		≈ 0.193552	>	6/31	0.193548	$\overrightarrow{C_5}^{\odot n}$
7		≈ 0.105867	>	c_7	0.102124	G^7
8		81/512	=	81/512	0.158203	$I_n \overset{3/4}{\Rightarrow} I_n$
9		c_9	=	c_9	0.423570	G^9
10		81/400	=	81/400	0.2025	G^{10}
11		1/8	=	1/8	0.125	$T_n^{\mathrm{reg}} \sqcup T_n^{\mathrm{reg}}$
12		1/2	=	1/2	0.5	$\overrightarrow{T_n} \sqcup \overrightarrow{T_n}$
13		2/21	=	2/21	0.095238	$\overrightarrow{C_4}^{\odot n}$
14		3/16	=	3/16	0.1875	$K_{n,n}^{\mathrm{rand}}$
15		c_{15}	=	c_{15}	0.189000	G^{15}
16		3/8	=	3/8	0.375	$\overrightarrow{K_{n,n}}$
17		≈ 0.095640	>	2/21	0.095238	$H^{\odot n}$
18		≈ 0.189030	>	c_{18}	0.189000	G^{18}
19		≈ 0.317681	>	c_{19}	0.317678	G^{19}
20		≈ 0.119760	>	c_{20}	0.119537	G^{20}
21		c_{21}	=	c_{21}	0.227173	$S^1(\frac{9+\sqrt{3}}{26})$
22		≈ 0.244055	>	c_{22}	0.244053	G^{22}
23		≈ 0.177784	>	c_{23}	0.177630	G^{23}
24		3/8	=	3/8	0.375	$\overrightarrow{T_n} \to I_n$
25		4/9	=	4/9	0.444444	$\overrightarrow{C_3} \odot I_n$
26		≈ 0.113205	>	c_{26}	0.112567	G^{26}
27		≈ 0.148148	>	4/27	0.148148	$S^1(4/9)$
28		c_{28}	=	c_{28}	0.157501	G^{28}
29		1/2	=	1/2	0.5	$S^1(1/2)$
30		1	=	1	1	$\overrightarrow{T_n}$

Table 7.1: Summary of bounds on the inducibility of graphs on four vertices.

have roughly αn vertices). If no cluster sizes occur, all clusters in the picture are meant to have equal sizes. Every cluster forming an independent set is depicted as a white (empty) circle, while a cluster forming an arbitrary tournament is depicted with a gray (shaded) circle. When some other structure appears inside the cluster, then there is a letter indicating the structure. In particular, **T** for the transitive tournament and **R** for the random tournament. Iterated constructions are marked by a dot inside the circle — this means that the cluster consists of a copy of the entire construction. We also provide a special pictogram for a circular graph $S^1(\alpha)$. All the above pictograms are summarized in Figure 7.1.

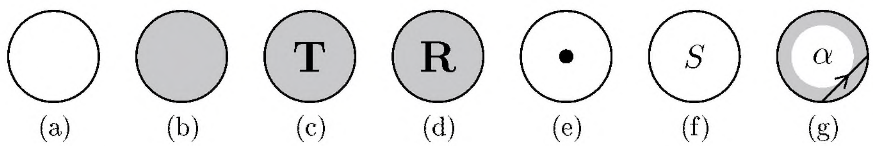


Figure 7.1: (a) Empty graph. (b) Arbitrary tournament. (c) Transitive tournament. (d) Random tournament. (e) Iterative structure. (f) Cluster with structure S. (g) Circular graph with parameter α.

7.3 Graphs

In the following subsections, for each oriented graph on four vertices we present the proven upper and lower bound on its inducibility and provide schematic picture and description of a construction giving the lower bound. If a pair $H, \overleftarrow{H}$ is considered, only a construction for the graph H is presented; by reversing arcs, one may obtain an analogous construction for the graph $\overleftarrow{H}$.

Each value of the upper bound which is not preceded with the approximation symbol ($\approx$) is proven exactly and meets the lower bound. Whenever we use Flagmatic, we indicate on how many vertices it is performed. Unless it is possible to make some reduction by forbidding certain structures in the extremal construction, the computations are performed using graphs on at most six vertices. It is possible to increase this number, which will result in slight improvement of the upper bounds, but will cause much longer running time of the program.

In almost all cases with sharp bounds, the standard application of the flag algebras semidefinite method is insufficient due to rounding errors made in the rationalization process. In order to overcome this, one needs to provide additional eigenvectors for the eigenvalue zero of the numerically obtained semidefinite matrix, which are not implied by the extremal construction. Such eigenvectors were found using the method described in the appendix of [6] and in Section 2.4.2 of [4]. For each calculation, we published the applied Flagmatic code with all commands and added eigenvectors, as well as a certificate useful for verification of the obtained bound. Explanation how to understand the codes is written in Section 2.5.6 of Chapter 2. All the codes and certificates are available at `https://arxiv.org/abs/2010.11664`.

We use for two graphs the following version of Lemma 2.1 from [54], which can be proved using essentially the same method.

Lemma 7.3. *Let H be an oriented graph with the following property — for every $v, w \in V(H)$ not joined by an arc, $N^+(v) = N^+(w)$ and $N^-(v) = N^-(w)$. Then, for any $n \in \mathbb{N}$, there exists an oriented graph G on n vertices satisfying the same property and such that $N(H, G) = \max\{N(H, G') : |V(G')| = n\}$.*

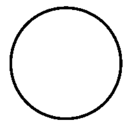
7.3.1 Graph 1

Graph:

Upper bound: 1

Lower bound: 1

Construction:



Empty graph.

7.3.2 Graph 2

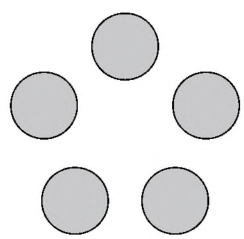
Graph: 

Upper bound: $72/125$

As every connected component of this graph is a transitive tournament, the inducibility of this graph is equal to the inducibility of the undirected graph $K_2 \sqcup I_2$ (an edge on four vertices), which was explicitly determined in [31].

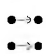
Lower bound: $72/125$

Construction:



Balanced union of arbitrary five tournaments.

7.3.3 Graph 3

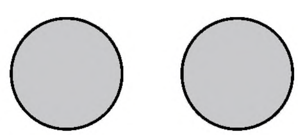
Graph: 

Upper bound: $3/8$

As every connected component of this graph is a transitive tournament, the inducibility of this graph is equal to the inducibility of the undirected graph $K_2 \sqcup K_2$, which is equal to the inducibility of its complement $K_{2,2}$, whose value is known [11, 68].

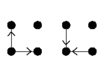
Lower bound: $3/8$

Construction:



Balanced union of arbitrary two tournaments.

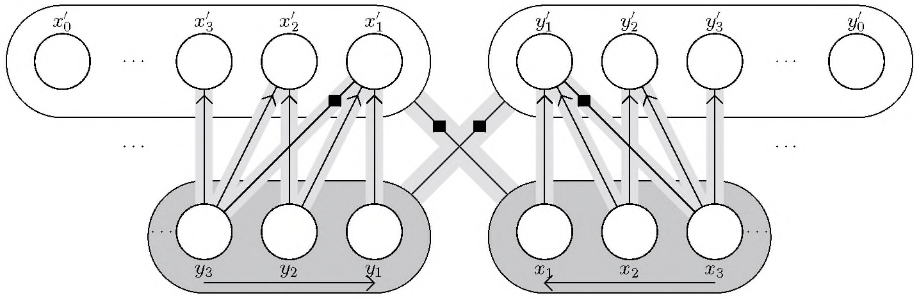
7.3.4 Graph 4

Graph: 

Upper bound (by Flagmatic on 6 vertices): ≈ 0.235046

Lower bound: ≈ 0.234309 obtained by cluster size optimization, assuming that $x_i = y_i$ for all $i \geq 1$, $x'_i = y'_i$ for all $i \geq 0$, $x'_1 = x'_i = 0$ for $i \geq 6$, and $(x_i)_{i \geq 5}$ is a geometric series.

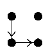
Construction:



$$x'_0 + y'_0 + \sum_{i \geq 1} x_i + y_i + x'_i + y'_i = 1$$

Construction G^4 is the following. Split vertices into parts X_i and Y_i, $i \geq 0$. Consider also partitions $X_0 = \bigcup_{i=0}^{\infty} X'_i$ and $Y_0 = \bigcup_{i=0}^{\infty} Y'_i$. Add all arcs from X_j to X_i and from Y_j to Y_i for $0 \leq i < j$. Finally, add all arcs from X_i to $\bigcup_{j=1}^{i} Y'_i$ and from Y_i to $\bigcup_{j=1}^{i} X'_i$ for $i \geq 1$.

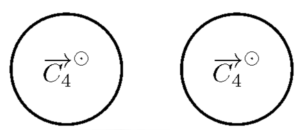
7.3.5 Graph 5

Graph: 

Upper bound (by Flagmatic on 6 vertices): ≈ 0.204123

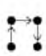
Lower bound: $64/315 \approx 0.203175$

Construction:



Balanced union of two iterated blow-ups of $\overrightarrow{C_4}$.

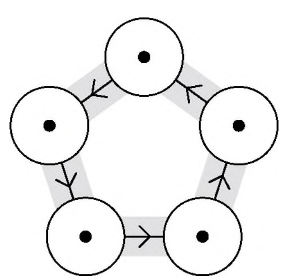
7.3.6 Graph 6

Graph: 

Upper bound (by Flagmatic on 6 vertices): ≈ 0.193552

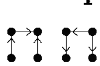
Lower bound: $6/31 \approx 0.193548$

Construction:



Iterated blow-up of $\overrightarrow{C_5}$.

7.3.7 Graph 7

Graph: 

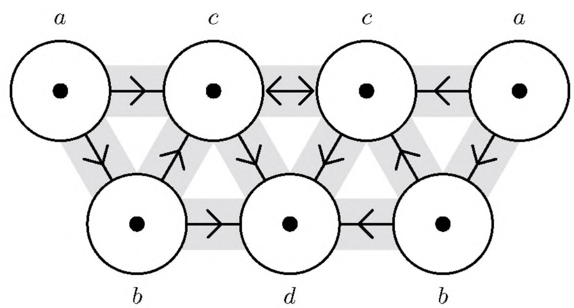
Upper bound (by Flagmatic on 6 vertices): ≈ 0.105867

Lower bound: $c_7 \approx 0.102124$, where

$$c_7 = \max_{\substack{a,b,c,d \in [0,1] \\ 2a+2b+2c+d=1}} \frac{24((a+b)^2c^2 + 2abd(b+2c))}{1 - 2a^4 - 2b^4 - 2c^4 - d^4}$$

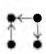
and the maximum is attained at $(a, b, c, d) \approx (0.117446, 0.159343, 0.146896, 0.152629)$.

Construction:



Construction G^7 is a weighted iterated blow-up of a graph on 7 vertices. More specifically, let $2a + 2b + 2c + d = 1$ and split vertices into 7 parts — two of size a, two of size b, two of size c, and one of size d. Add arcs between appropriate parts to make them complete bipartite and orient them as shown in the picture. Finally, iterate this process inside each of the seven parts. Note that the direction of arcs between parts of size c is not important, hence many non-isomorphic examples of graphs with this bound met can be found.

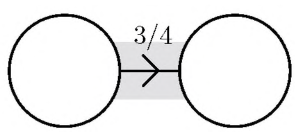
7.3.8 Graph 8

Graph: 

Upper bound (by Flagmatic on 6 vertices): $81/512$

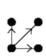
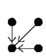
Lower bound: $81/512$

Construction:



Union of two balanced empty graphs with random arcs from the first to the second with probability $\frac{3}{4}$.

7.3.9 Graph 9

Graph:  

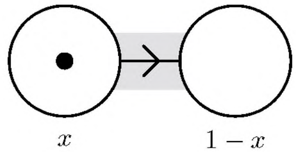
Upper bound: c_9, where

$$c_9 = \max_{x \in [0, \frac{1}{2}]} \frac{32(1-2x)x^3}{1-(1-2x)^4} = 4 - \frac{6}{\sqrt[3]{\sqrt{2}-1}} + 6\sqrt[3]{\sqrt{2}-1}$$

and the maximum is attained at $x \approx 0.37346$. Proved by Huang [54], also included as an example application of the Flagmatic software [77].

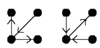
Lower bound: $c_9 \approx 0.423570$

Construction:



Split vertices into two parts, A of size x and B of size $1-x$. Then, put all possible arcs from A to B and iterate this process inside A.

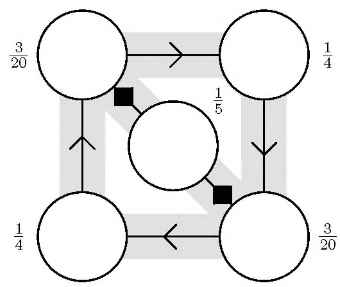
7.3.10 Graph 10

Graph: 

Upper bound (by Flagmatic on 5 vertices): $81/400$

Lower bound: $81/400$

Construction:



Construction G^{10} is the following. Split vertices into 5 parts — A of size $\frac{3}{20}$, B of size $\frac{1}{4}$, C of size $\frac{3}{20}$, D of size $\frac{1}{4}$, and E of size $\frac{1}{5}$. Put all arcs from A to B, from B to C, from C to D, from D to A, from C to E, and from A to E. In [53], there is also a general probabilistic construction for any orientation of a star.

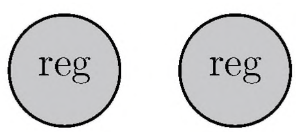
7.3.11 Graph 11

Graph: 

Upper bound (by Flagmatic on 5 vertices): $1/8$

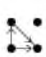
Lower bound: $1/8$

Construction:



Balanced union of two arbitrary regular tournaments.

7.3.12 Graph 12

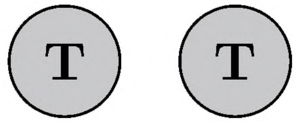
Graph: 

Upper bound: $1/2$

As every connected component of this oriented graph is a transitive tournament, the inducibility of this graph is equal to the inducibility of the undirected graph $K_3 \sqcup I_1$, which is equal to the inducibility of $K_{1,3}$ determined in [18].

Lower bound: $1/2$

Construction:



Balanced union of two transitive tournaments.

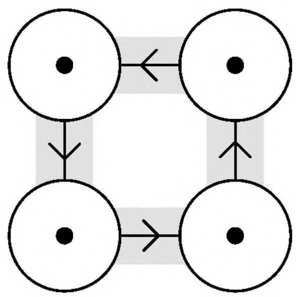
7.3.13 Graph 13

Graph: 

Upper bound: 2/21 proved in [52] using the flag algebra method with additional stability arguments.

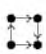
Lower bound: 2/21

Construction:



Iterated blow-up of $\overrightarrow{C_4}$.

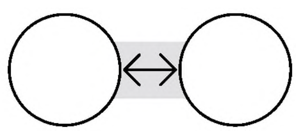
7.3.14 Graph 14

Graph: 

Upper bound (by Flagmatic on 5 vertices): 3/16

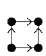
Lower bound: 3/16

Construction:



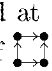
Balanced complete bipartite graph with randomly oriented edges. In [52], it is additionally proven that every extremal graph is within edit distance $o(n^2)$ from a pseudorandom orientation of a balanced complete bipartite graph.

7.3.15 Graph 15

Graph: 

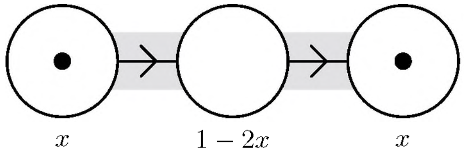
Upper bound: c_{15} proved in [52], where

$$c_{15} = \max_{x \in [0, \frac{1}{2}]} \frac{12x^2(1-2x)^2}{1-2x^4} = 9\left(\sqrt{2}-2\right) + 6\sqrt{2\left(\sqrt{2}-1\right)}$$

and the maximum is attained at $x \approx 0.25202$. It is shown in [52] that the unique limit object maximizing the inducibility of  is given by the construction presented below.

Lower bound: $c_{15} \approx 0.189000$

Construction:



Construction G^{15} is the following. Split vertices into three parts, A of size x, B of size $1-2x$, and C of size x. Put all possible arcs from A to B and from B to C. Iterate this process inside A and inside C.

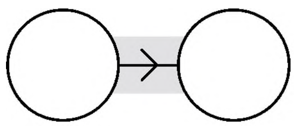
7.3.16 Graph 16

Graph: 

Upper bound: 3/8
It is equal to the inducibility of undirected $K_{2,2}$ solved in [11, 68].

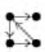
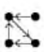
Lower bound: 3/8

Construction:



Complete balanced bipartite graph with edges oriented from the first part to the second.

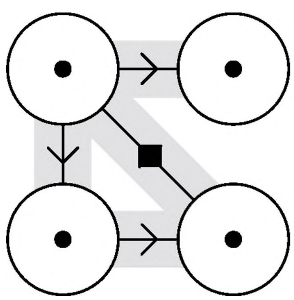
7.3.17 Graph 17

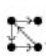
Graph:  

Upper bound (by Flagmatic on 6 vertices): ≈ 0.095640

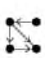
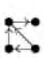
Lower bound: $2/21 \approx 0.095238$

Construction:



Iterated blow-up of graph .

7.3.18 Graph 18

Graph:  

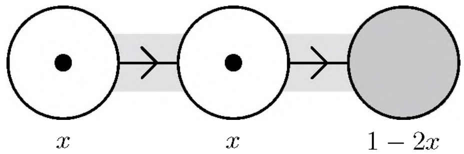
Upper bound (by Flagmatic on 6 vertices): ≈ 0.189030

Lower bound: $c_{18} \approx 0.189000$, where

$$c_{18} = \max_{x\in[0,\frac{1}{2}]} \frac{12x^2(1-2x)^2}{1-2x^4} = 9\left(\sqrt{2}-2\right) + 6\sqrt{2\left(\sqrt{2}-1\right)}$$

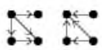
and is attained at $x \approx 0.25202$.

Construction:



Construction G^{18} is the following. Split vertices into three parts, A of size x, B of size x, and C of size $1-2x$. Put all possible arcs from A to B and from B to C, put also all possible arcs inside C to make it an arbitrary tournament. Iterate this process inside A and inside B.

7.3.19 Graph 19

Graph: 

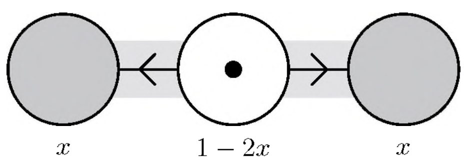
Upper bound (by Flagmatic on 6 vertices): ≈ 0.317681

Lower bound: $c_{19} \approx 0.317678$, where

$$c_{19} = \max_{x \in [0, \frac{1}{2}]} \frac{24(1-2x)x^3}{1-(1-2x)^4} = \frac{3}{2}\left(2 - \frac{3}{\sqrt[3]{\sqrt{2}-1}} + 3\sqrt[3]{\sqrt{2}-1}\right)$$

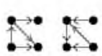
and the maximum is attained at $x \approx 0.37346$.

Construction:



Construction G^{19} is the following. Split vertices into three parts — A of size x, B of size x, and C of size $1-2x$. Put all possible arcs from C to A, from C to B and all possible arcs in parts A and B to make them arbitrary tournaments. Iterate this process inside part C.

7.3.20 Graph 20

Graph: 

Upper bound (by Flagmatic on 6 vertices): ≈ 0.119760

Lower bound: $c_{20} \gtrsim 0.119537$, where

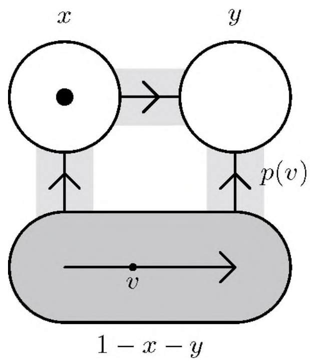
$$c_{20} = \max \frac{24y(1-x-y)}{1-x^4}(xyI_1 + y(1-x-y)I_2 + (1-x-y)^2 I_3),$$

$$I_1 = \int_0^1 p(a)(1-p(a))\,\mathrm{d}a, \quad I_2 = \int_0^1 \int_0^a p(a)^2 p(b)(1-p(b))\,\mathrm{d}b\,\mathrm{d}a,$$

$$I_3 = \int_0^1 \int_0^a \int_0^b (1-p(c))p(b)(1-p(a))\,\mathrm{d}c\,\mathrm{d}b\,\mathrm{d}a$$

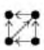
and the maximum is taken over all $x, y \in [0,1]$ and functions $p : [0,1] \longrightarrow [0,1]$. The above lower bound is obtained by maximizing over values of $x, y \in [0,1]$ and p being a polynomial of degree at most 7.

Construction:



Construction G^{20} is the following. Split vertices into three parts — A of size x, B of size y, and C of size $1 - x - y$, fix also a function $p : [0,1] \to [0,1]$. Put all possible arcs from A to B, from C to A, and inside part C to make it a transitive tournament. Treat C as a finite subset of $[0,1]$, with order $<$ on $[0,1]$ preserving the transitive order of vertices of C. For each vertex $v \in C$ and $w \in B$, put an arc from v to w independently at random with probability $p(v)$. Finally, iterate this process inside part A.

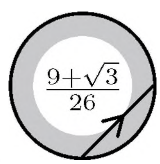
7.3.21 Graph 21

Graph: 

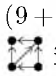
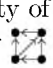
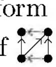
Upper bound (by Flagmatic on 5 vertices): $(28 + 6\sqrt{3})/169$

Lower bound: $(28 + 6\sqrt{3})/169 \approx 0.227173$

Construction:

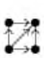


Circular graph with parameter $(9 + \sqrt{3})/26$.

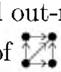
To determine the density of  in the circular graph $S^1(\alpha)$ with parameter $\alpha = (9 + \sqrt{3})/26$, note that if the vertices of  are denoted by v, x, y, z in such a way that vx, xy, and yz are arcs, then every its embedding has the vertices in order v, x, y, z along the circle. Fix v on the circle and parameterize the remaining vertices by their oriented arc-distance from v. Since z is non-adjacent to v, its position must be in the interval $(\alpha, 1-\alpha)$, while x and y need to form an ordered pair in the interval of vertices adjacent to both v and z. Therefore, the density of  in $S^1(\alpha)$ is equal to

$$4! \int_{\alpha}^{1-\alpha} \frac{1}{2}(\alpha - (z - \alpha))^2 \, \mathrm{d}z = \frac{28 + 6\sqrt{3}}{169}.$$

7.3.22 Graph 22

Graph: 

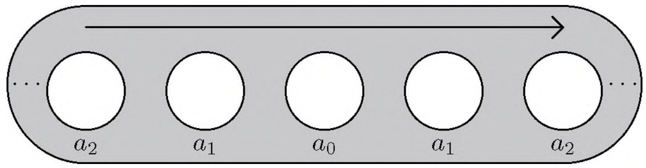
Upper bound (by Flagmatic on 8 vertices): ≈ 0.244055

By Lemma 7.3, we may assume that the oriented graphs maximizing the number of induced copies of  have the property that the in- and out-neighborhood of non-neighbors are the same. In particular, we can consider the inducibility of  in a family of graphs which have no induced copy of $T_2 \sqcup I_1$ (an arc plus an isolated vertex) or $\overrightarrow{P_3}$, simplifying this way the computations in Flagmatic.

Lower bound: $c_{22} \approx 0.244053$, where

$$c_{22} = \max_{0\le 2y\le 1-q\le 1} \left(1 - \frac{2y}{1-q}\right)^2 \frac{12y^2}{(1-q)^2} + \frac{24qy^3}{(1-q)^2}\left(\frac{1}{1+q+q^2} - \frac{y}{1-q^4}\right).$$

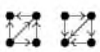
Construction:



$$a_0 + 2a_1 + 2a_2 + \cdots = 1$$

Construction consists of a sequence of clusters A^i for $i \in \mathbb{Z}$, where each A^i is an empty graph, and each pair of vertices from different clusters is joined by an arc pointing to the cluster with the larger index. Let the cluster A^i be of the size $a_{|i|}$, where $a_0 + 2a_1 + 2a_2 + \ldots = 1$. Letting the cluster sizes decrease geometrically, i.e., $a_i = yq^{i-1}$ for $i = 1, 2, \ldots$ and $a_0 = 1 - \frac{2y}{1-q}$, we obtain the desired density.

7.3.23 Graph 23

Graph: 

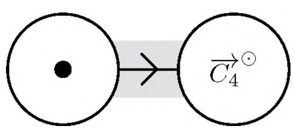
Upper bound (by Flagmatic on 6 vertices): ≈ 0.177784

Lower bound: $c_{23} \approx 0.177630$, where

$$\begin{aligned} c_{23} &= \max_{x\in[0,1]} \frac{\frac{8}{5}x(1-x)^3 + \frac{8}{315}(1-x)^4}{1-x^4} \\ &= \frac{4}{105}\left(61^{2/3}\sqrt[3]{\sqrt{7690}-63} - 61^{2/3}\sqrt[3]{63+\sqrt{7690}} + 42\right) \end{aligned}$$

and the maximum is attained at $x \approx 0.24063$.

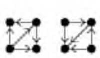
Construction:



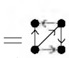
x $\qquad$ $1 - x$

Construction G^{23} is the following. Split vertices into two parts — A of size x and B of size $1 - x$. Put arcs in B to make it an iterated blow-up of $\overrightarrow{C_4}$ and all possible arcs from A to B. Iterate the process inside A.

7.3.24 Graph 24

Graph: 

Upper bound: 3/8

Proof. Let G be a graph on n vertices maximizing the number of induced copies of $H =$  and introduce the following equivalence relation in $V(G)$: $v \sim w$ if and only if $N^+(v) = N^+(w)$ and $N^-(v) = N^-(w)$. By Lemma 7.3, since H has the property that $v \sim w$ if and only if v and w are not joined by an arc, we may assume that G satisfies the same property.

Define T as the set of all vertices which belong to the equivalence classes of size one, and let $t = |T|$. Let m be the number of the remaining equivalence classes; we shall write them as

$B_1, \ldots, B_m$, where $b_i = |B_i|$ and $b_1 \le \ldots \le b_m$. Note that each induced copy of H which is not disjoint from T contains exactly two vertices in T (joined by an arc, which can be oriented in any way) and two vertices in some B_i, with arcs oriented from T to B_i. Hence, the number of induced copies of H in G does not depend on the orientation of arcs between vertices in T. Also, by reorienting the arcs in such a way that $T \Rightarrow B_i$ for every $i \in [m]$, we will not decrease the number of induced copies of H. The following claim gives the orientation of arcs between the remaining pairs of equivalence classes.

Claim 7.4. *We can reorient the arcs between B_i and B_j for every $i < j$ so that $B_i \Rightarrow B_j$ without decreasing the number of induced copies of H.*

Proof. Consider a partition $I \cup J = [m-1]$ of indices such that $B_i \Rightarrow B_m$ for $i \in I$ and $B_m \Rightarrow B_j$ for $j \in J$. Let us modify the graph G by reorienting all arcs incident to B_j and B_m towards B_m for each $j \in J$. Each induced copy of H removed in this way has exactly one vertex in B_m, exactly two vertices in B_j for some $j \in J$, and one more vertex which belongs neither to B_j nor B_m and has outgoing arcs to B_j. The number of lost copies is then equal to

$$b_m \sum_{j \in J} \binom{b_j}{2} \left(t + \sum_{i \in I, B_i \Rightarrow B_j} b_i + \sum_{j' \in J, B_{j'} \Rightarrow B_j} b_{j'} \right). \tag{7.1}$$

The created copies of H contain exactly two vertices in B_m and at least one vertex in B_j for some $j \in J$, therefore the number of created copies is equal to

$$\binom{b_m}{2} \left(t \sum_{j \in J} b_j + \sum_{i \in I, j \in J} b_i b_j + \sum_{j < j' \in J} b_j b_{j'} \right). \tag{7.2}$$

Now, we can compare both values by looking at the appropriate terms in both formulas. For each pair $j \neq j' \in J$, we have exactly one of the following two terms

$$b_m \binom{b_j}{2} b_{j'}, \quad b_m \binom{b_{j'}}{2} b_j$$

in (7.1) containing $b_j b_{j'}$, and exactly one term

$$\binom{b_m}{2} b_j b_{j'}$$

in (7.2). Since $b_m \ge b_i$ for all $i \in [m-1]$, we have

$$\binom{b_m}{2} b_j b_{j'} \ge b_m \binom{b_j}{2} b_{j'} \quad \text{and} \quad \binom{b_m}{2} b_j b_{j'} \ge b_m \binom{b_{j'}}{2} b_j.$$

Also, for each $i \in I$ and $j \in J$, we have at most one term containing $b_i b_j$ in (7.1) — it is equal to $b_m \binom{b_j}{2} b_i$, which is not greater than $\binom{b_m}{2} b_i b_j$ in (7.2). Finally, for t and any index $j \in J$, we have one term $b_m \binom{b_j}{2} t$ in (7.1), which is again not greater than $\binom{b_m}{2} b_j t$ in (7.2). Therefore, (7.1) is not greater than (7.2), i.e. the number of induced copies of H in G did not decrease after reorienting all arcs towards B_m.

We can repeat this argument for the subgraph of G induced by the union of T and $B_1, \ldots, B_{m-1}$ — just note that by reorienting arcs in this subgraph, we will not decrease the number of induced copies of H which contain at least one vertex in B_m. By repeating this iterative procedure, we will eventually obtain the property that $B_i \Rightarrow B_j$ whenever $i < j$. $\square$

So far, we proved that the graph G has the following structure:

$$\overrightarrow{T \Rightarrow B_1 \Rightarrow B_2 \Rightarrow \ldots \Rightarrow B_m}.$$

We would like to show that $m = 1$, so that G is just of the form $T \Rightarrow B_1$. Assume that this is not the case, i.e. $m \geq 2$, and G is an extremal graph with b_m being maximal and — for this value of b_m — also t being maximal.

Claim 7.5. *If $m \geq 2$, then $t > b_m$.*

Proof. Since G is an extremal graph, introducing a single arc in B_1 shall not increase the number of induced copies of H. Such an arc removes $\binom{t}{2}$ induced copies of H and creates $\sum_{i=2}^{m} \binom{b_i}{2}$ induced copies of H. Therefore, we have

$$\binom{t}{2} \geq \sum_{i \geq 2} \binom{b_i}{2}.$$

In particular, $t \geq b_m$. Moreover, if $m > 2$, then $t > b_m$, since $\binom{b_2}{2} \geq 1$. Also, if $m = 2$ and $t = b_m$, then we can move those two vertices which we just joined by an arc to T; this way, we will not decrease the number of induced copies of H in G, but we will increase the size of T, contradicting our assumptions about the maximality of the graph G. It follows that we must have $t > b_m$. □

We are ready to give an argument contradicting the maximality of G. Remove one vertex from T and add one vertex to B_m. This way, we destroy the following number of induced copies of H:

$$(t-1)\sum_{i=1}^{m} \binom{b_i}{2} + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} b_i \binom{b_j}{2},$$

which can be written in the following way:

$$(t-1)\binom{b_m}{2} + (t-1)\sum_{i=1}^{m-1} \binom{b_i}{2} + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m-1} b_i \binom{b_j}{2} + \sum_{i=1}^{m-1} b_i \binom{b_m}{2}. \tag{7.3}$$

On the other hand, we create the following number of induced copies of H:

$$b_m \left(\binom{t-1}{2} + (t-1)\sum_{i=1}^{m-1} b_i + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m-1} b_i b_j \right),$$

which can be written in the following way:

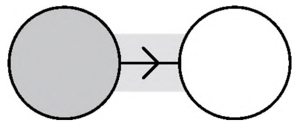
$$b_m \binom{t-1}{2} + (t-1)\sum_{i=1}^{m-1} \frac{b_i b_m}{2} + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m-1} b_i b_j b_m + \sum_{i=1}^{m-1} (t-1)\frac{b_i b_m}{2}. \tag{7.4}$$

Since $t - 1 \geq b_m$ by Claim 7.5, it is straightforward to see that each summand of (7.3) is not greater than the appropriate summand of (7.4), and that there is at least one strictly smaller summand (e.g. the last one).

It follows that the graph G must be of the form $T \Rightarrow B_1$. The number of induced copies of H in G is then equal to $\binom{t}{2}\binom{n-t}{2}$, which is maximized for $t \in \{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil\}$. This gives the desired upper bound for the inducibility of H. □

Lower bound: 3/8

Construction:

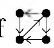


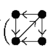
Full join from an arbitrary tournament to an empty graph of the same size.

7.3.25 Graph 25

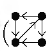
Graph: 

Upper bound: 4/9

Proof. Note that in each induced copy of  there are exactly two vertices for which out-degree, in-degree, and non-degree are equal to 1. Basing on this observation and using AM-GM inequality, we get the following bound on the number of induced copies of  in any n-vertex graph G:

$N($$, G) \leq \frac{1}{2} \sum_{v \in V(G)} d^+(v) d^-(v) d'(v) \leq \frac{n(n-1)^3}{54}.$

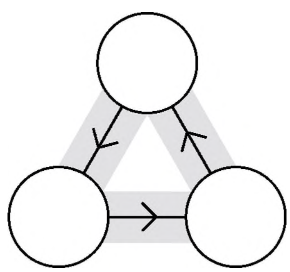
In particular,

$i($$) \leq \lim_{n \to \infty} \frac{n(n-1)^3}{54} \Big/ \binom{n}{4} = \frac{4}{9}.$

$\square$

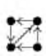
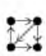
Lower bound: 4/9

Construction:



Balanced blow-up of $\overrightarrow{C_3}$.

7.3.26 Graph 26

Graph:  

Upper bound (by Flagmatic on 6 vertices): ≈ 0.113205

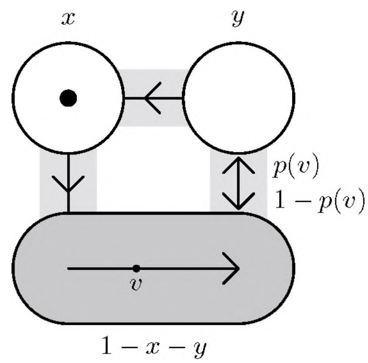
Lower bound: $c_{26} \gtrsim 0.112567$, where

$$c_{26} = \max \frac{24y^2(1-x-y)}{1-x^4}(xI_1 + (1-x-y)I_2),$$

$$I_1 = \int_0^1 p(a)(1-p(a))\,\mathrm{d}a, \quad I_2 = \int_0^1 \int_0^a (1-p(b))^2 p(a)(1-p(a))\,\mathrm{d}b\,\mathrm{d}a$$

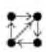
and the maximum is taken over all $x, y \in [0,1]$ and functions $p : [0,1] \longrightarrow [0,1]$. The above lower bound is obtained by maximizing over $x, y \in [0,1]$ and p being a polynomial of degree at most 7.

Construction:



Construction G^{26} is the following. Split vertices into three parts — A of size x, B of size y, and C of size $1-x-y$, fix also a function $p:[0,1]\to[0,1]$. Put all possible arcs from B to A, from A to C, and inside part C to make it a transitive tournament. Treat C as a finite subset of $[0,1]$, with order $<$ on $[0,1]$ preserving the transitive order of vertices of C. For each $v\in C$ and $w\in B$, put an arc between v and w oriented independently at random, with probability $p(v)$ from v to w and with probability $1-p(v)$ from w to v. Finally, iterate this process inside part A.

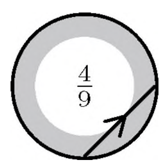
7.3.27 Graph 27

Graph: 

Upper bound (by Flagmatic on 6 vertices): ≈ 0.148148

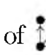
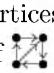
Lower bound: $4/27 \approx 0.148148$

Construction:

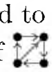


Circular graph with parameter 4/9.

The exact value of the upper bound found by Flagmatic is equal to $\frac{4919131754658232715 51}{332041393326771929 0880}$, the decimal expansion of which agrees with $4/27$ up to the tenth decimal place. We believe it is possible to obtain the proof of the optimal upper bound by providing eigenvectors for the eigenvalue zero and solving the semidefinite problem with sufficiently good precision. Unfortunately, computations on 6 vertices generate too large numerical errors on `csdp`, while computations on a solver with higher precision require too much computational power. Thus, we were unable to prove the optimal bound.

To determine the density of  in the circular graph $S^1(4/9)$, note that each embedding of a directed triangle can be extended to an embedding of  by adding any vertex non-adjacent to one of the vertices of the triangle, and the set of such vertices has measure $1/3$. Moreover, each embedding of  is counted this way exactly once.

In order to find the density of a directed triangle with arcs vx, xy, and yv, we fix v on the circle and parameterize the remaining vertices by their oriented arc-distance from v. To create a directed triangle, the position of x must be in the interval $(1/9, 4/9)$, while y needs to be in the interval of vertices appropriately connected to v and x. Since this way we count each directed triangle three times, we need to divide the obtained value by 3.

Therefore, the density of  in $S^1(4/9)$ is equal to

$$\frac{4!}{9}\int_{1/9}^{4/9}\left(x+\frac{4}{9}-\frac{5}{9}\right)\mathrm{d}x=\frac{4}{27}.$$

7.3.28 Graph 28

Graph:  

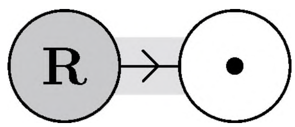
Upper bound: c_{28} proved in [19] (Theorem 4).

Lower bound: $c_{28} \approx 0.157501$, where

$$c_{28} = \max_{x \in [0,1]} \frac{x^4/8 + x^3(1-x)}{1-(1-x)^4} = \frac{8 - 3^{7/3} + 3^{5/3}}{8}$$

and the maximum is attained at $x \approx 0.85642$. It is proved in [19] that the extremal construction for sufficiently large number of vertices is precisely the one presented below.

Construction:



x $\quad$ $1-x$

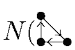
Construction G^{28} is the following. Split vertices into two parts — A of size x and B of size $1-x$. Let A be a random tournament and put all possible arcs from A to B. Finally, iterate the process inside B.

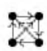
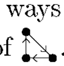
7.3.29 Graph 29

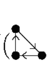
Graph: 

Upper bound: $1/2$

Proof. Let G be an n-vertex oriented graph with the maximum number of induced copies of . As adding an arc between two non-neighbors may only increase the number of induced copies of , we may assume that G is a tournament, i.e., $d^+(v) + d^-(v) = n-1$ for every $v \in V(G)$. Note that

$N($$, G) \leq \frac{1}{2} \sum_{v \in V(G)} N($$, G - v),$

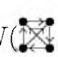
as in each induced copy of  there are two ways to select vertex v in such a way that the remaining three vertices form an induced copy of . Furthermore, for every tournament H on k vertices, we have

$2N($$, H) + \binom{k}{3} = 3N($$, H) + N($$, H) = \sum_{v \in V(H)} d^+(v)d^-(v) \leq \frac{k(k-1)^2}{4}$

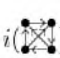
by the AM-GM inequality, and in consequence

$N($$, H) \leq \frac{(k-1)k(k+1)}{24}.$

Applying this inequality for $k = n-1$ and $H = G - v$ for every $v \in V(G)$, and plugging to the previous estimation, we finally get

$N($$, G) \leq \frac{n^2(n-1)(n-2)}{48},$

hence

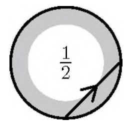
$i($$) \leq \lim_{n \to \infty} \frac{n^2(n-1)(n-2)}{48} \Big/ \binom{n}{4} = \frac{1}{2}.$

□

An independent proof can be found in [19].

Lower bound: 1/2

Construction:



Circular graph with parameter 1/2.

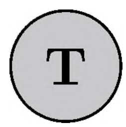
7.3.30 Graph 30

Graph: 

Upper bound: 1

Lower bound: 1

Construction:



Transitive tournament.

Bibliography

[1] N. Alon, J. Pach, and J. Solymosi: *Ramsey-type theorems with forbidden subgraphs*, Combinatorica **21** (2001), 155–170.

[2] N. Alon and A. Shapira: *Testing subgraphs in directed graphs*, J. Comput. System Sci. **69** (2004), 354–382.

[3] N. Alon and C. Shikhelman: *Many T copies in H-free graphs*, J. Combin. Theory Ser. B **121** (2016), 146–172.

[4] R. Baber: *Some results in extremal combinatorics*, Doctoral thesis, University College London (2011), 68–75; available at `https://www.ucl.ac.uk/~ucahjmt/Thesis.pdf`.

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

[6] J. Balogh, P. Hu, B. Lidický, O. Pikhurko, B. Udvari, and J. Volec: *Minimum number of monotone subsequences of length 4 in permutations*, Combin. Probab. Comput. **24** (2015), 658–679.

[7] J.P. Bartley: *Topics in Extremal and Probabilistic Combinatorics*, PhD Thesis, Queen Mary University of London, 2018.

[8] L. Beineke and F. Harary: *The maximum number of strongly connected subtournaments*, Can. Math. Bull. **8** (1965), 491–498.

[9] B. Bollobás, Y. Egawa, A. Harris, and G. P. Jin: *The maximal number of induced r-partite subgraphs*, Graphs Combin. **11** (1995), 1–19.

[10] B. Bollobás and E. Győri: *Pentagons vs. triangles*, Discrete Math. **308** (2008) 4332–4336.

[11] B. Bollobás, C. Nara, and S.-i. Tachibana: *The maximal number of induced complete bipartite graphs*, Discrete Math. **62** (1986), 271–275.

[12] J. A. Bondy and M. Simonovits: *Cycles of even length in graphs*, J. Combin. Theory Ser. B **16** (1974), 97–105.

[13] B. Borchers: *CSDP, A C library for semidefinite programming*, Optim. Methods Softw. **11** (1999), 613–623.

[14] Ł. Bożyk, A. Grzesik, and B. Kielak: *On the inducibility of oriented graphs on four vertices*, Discrete Math. **345(7)** (2022), 112874.

[15] W.G. Brown, P. Erdős, and M. Simonovits: *Extremal problems for directed graphs*, J. Combin. Theory Ser. B **15** (1973), 77–93.

[16] W.G. Brown, P. Erdős, and M. Simonovits: *Algorithmic solution of extremal digraph problem*, Trans. Amer. Math. Soc. **292** (1985), 421–449.

[17] W.G. Brown and F. Harary: *Extremal digraphs*, Combinatorial theory and its applications (Vol. I), North-Holland Publishing Company, 1970, 135–198.

[18] J. Brown and A. Sidorenko: *The inducibility of complete bipartite graphs*, J. Graph Theory **18(6)** (1994), 629–645.

[19] D. Burke, B. Lidický, F. Pfender, and M. Phillips: *Inducibility of 4-Vertex Tournaments*, `https://arxiv.org/abs/2103.07047`.

[20] I. Choi, B. Lidický, and F. Pfender: *Inducibility of directed paths*, `https://arxiv.org/abs/1811.03747`

[21] U. Colombo: *Sui circuiti ne grafi completi*, Boll. Unione Math. Ital. **19** (1964), 153–170.

[22] A.N. Day: *A collection of problems in extremal combinatorics*, PhD Thesis, Queen Mary University of London, 2017.

[23] N. Draganić, F. Dross, J. Fox, A. Girão, F. Havet, D. Korándi, W. Lochet, D. M. Correia, A. Scott, and B. Sudakov: *Powers of paths in tournaments*, Combin. Probab. Comput. **30**(6) (2021), 1–5.

[24] P. Erdős: *On sequences of integers no one of which is divisible by any other and on some related problems*, Tomsk. Gos. Univ. Ucen. Zap. **2** (1938), 74–82.

[25] P. Erdős: *On some problems in graph theory, combinatorial analysis and combinatorial number theory*, Graph Theory and Combinatorics, Proc. Conf. Hon. P. Erdős, Cambridge 1983, 1–17 (1984).

[26] P. Erdős: *On the number of complete subgraphs contained in certain graphs*, Publ. Math. Inst. Hung. Acad. Sci. **7** (1962), 459–464.

[27] P. Erdős and A. Hajnal: *Ramsey-type theorems*, Discrete Math. **25** (1989), 37–52.

[28] P. Erdős and L. Moser: *On the representation of directed graphs as unions of orderings*, Publ. Math. Inst. Hung. Acad. Sci. **9** (1964), 125–132.

[29] P. Erdős and A. Stone: *On the structure of linear graphs*, Bull. Amer. Math. Soc. **52** (1946), 1087–1091.

[30] L. Euler: *Elementa doctrinae solidorum (in Latin)*, Novi Comment. Acad. Sci. Petropol. **4** (1758), 109–140.

[31] C. Even-Zohar and N. Linial: *A note on the inducibility of 4-vertex graphs*, Graphs Combin. **31(5)** (2015), 1367–1380.

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

[33] V. Falgas-Ravry, E.R. Vaughan: *Applications of the Semi-Definite Method to the Turán Density Problem for 3-Graphs*, Combin. Probab. Comput. **22(1)** (2013), 21–54.

[34] D. C. Fisher, J. R. Lundgren, S. K. Merz, and K. B. Reid: *Domination graphs of tournaments and digraphs*, Congr. Numer. **108** (1995), 97–107.

[35] J. Fox, H. Huang, and C. Lee: *A solution to the inducibility problem for almost all graphs*, manuscript.

[36] Z. Füredi and L. Özkahya: *On 3-uniform hypergraphs without a cycle of a given length*, Discrete Appl. Math. **216** (2017), 582–588.

[37] Z. Füredi and M. Simonovits: *The history of degenerate (bipartite) extremal graph problems*, Erdős Centennial **25** (2013), 169–264.

[38] D. Gerbner, E. Győri, A. Methuku, and M. Vizer: *Generalized Turán problems for even cycles*, J. Combin. Theory Ser. B **145** (2020), 169–213.

[39] L. Gishboliner and A. Shapira: *A generalized Turán problem and its applications*, Proc. of STOC 2018, 760–772.

[40] M. Górski and A. Grzesik: *On the maximum number of 5-cycles in graphs without 7-cycles*, manuscript.

[41] A. Grzesik: *On the maximum number of five-cycles in a triangle-free graph*, J. Combin. Theory Ser. B **102** (2012), 1061–1066.

[42] A. Grzesik and B. Kielak: *On the maximum number of odd cycles in graphs without smaller odd cycles*, J. Graph Theory **99(2)** (2022), 240–246.

[43] A. Grzesik, D. Kral', L. M. Lovász, and J. Volec: *Cycles of a given length in tournaments*, J. Combin. Theory Ser. B **158** (2023), 117–145.

[44] E. Győri: *On the number of C_5's in a triangle-free graph*, Combinatorica **9** (1989) 101–102.

[45] E. Győri, Z. He, Z. Lv, N. Salia, C. Tompkins, K. Varga, and X. Zhu: *Exact results for generalized extremal problems forbidding an even cycle*, `https://arxiv.org/abs/2208.02538`

[46] E. Győri and H. Li: *The maximum number of triangles in C_{2k+1}-free graphs*, Combin. Probab. Comput. **21** (2012), 187–191.

[47] R. Häggkvist and C. Thomassen: *On pancyclic digraphs*, J. Combin. Theory Ser. B **20** (1976), 20–40.

[48] H. Hatami, J. Hirst, and S. Norine: *The inducibility of blow-up graphs*, J. Combin. Theory Ser. B **109** (2014), 196–212.

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

[50] D. Hefetz and M. Tyomkin: *On the inducibility of cycles*, J. Combin. Theory Ser. B **133** (2018), 243–258.

[51] J. Hirst: *The inducibility of graphs on four vertices*, J. Graph Theory **75(3)** (2014), 231–243.

[52] P. Hu, B. Lidický, F. Pfender, and J. Volec: *Inducibility of orientations of C_4*, manuscript.

[53] P. Hu, J. Ma, S. Norin, and H. Wu: *Inducibility of oriented stars*, `https://arxiv.org/abs/2008.05430`

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

[55] M.G. Kendall, B. Babington Smith: *On the method of paired comparisons*, Biometrika **31** (1940), 324–345.

[56] N. Komarov and J. Mackey: *On the number of 5-cycles in a tournament*, J. Graph Theory **86** (2017), 341–356.

[57] J. Komlós and M. Simonovits: *Szemerédi's regularity lemma and its applications in graph theory*, Combinatorics, Paul Erdős is Eighty (Vol. 2), János Bolyai Mathematical Society, Keszthely (Hungary) (1993), Budapest (1996), 295–352.

[58] T. Kővari, V. T. Sós, and P. Turán: *On a problem of K. Zarankiewicz*, Colloq. Math. **3** (1954), 50–57.

[59] D. Král', S. Norin, and J. Volec: *A bound on the inducibility of cycles*, J. Combin. Theory Ser. A **161** (2019), 359–363.

[60] B. Lidický and F. Pfender: *Pentagons in triangle-free graphs*, European J. Combin. **74** (2018), 85–89.

[61] L. Lovász and B. Szegedy: *Limits of dense graph sequences*, J. Combin. Theory Ser. B **96** (2006), 933–957.

[62] J. Ma and Y. Qiu: *Some sharp results on the generalized Turán numbers*, European J. Combin. **84** (2020), 103026.

[63] Y. Manoussakis and Z. Tuza: *Ramsey numbers for tournaments*, Theoret. Comput. Sci. **263** (2001), 75–85.

[64] W. Mantel: *Vraagstuk XXVIII*, Wiskundige Opgaven **10** (1907), 60–61.

[65] T.S. Michael: *Cycles of length 5 in triangle-free graphs: a sporadic counterexample to a characterization of equality*, Bull. ICA **67** (2013), 6–8.

[66] M. Nakata: *A numerical evaluation of highly accurate multiple-precision arithmetic version of semidefinite programming solver: SDPA-GMP, -QD and -DD*, Proc. IEEE Multi-Conference on Systems and Control (2010), 29–34, available at `http://sdpa.sourceforge.net/`.

[67] A. Nijenhuis and H.S. Wilf: *Representations of Integers by Linear Forms in Nonnegative Integers*, J. Number Theory **4** (1972), 98–106.

[68] N. Pippenger and M.C. Golumbic: *The inducibility of graphs*, J. Combin. Theory Ser. B **19(3)** (1975), 189–203.

[69] A. Razborov: *Flag Algebras*, J. Symb. Log. **72** (2007), 1239–1282.

[70] The Sage Developers: *SageMath, the Sage Mathematics Software System*, available at `https://www.sagemath.org/`.

[71] É. Sopena: *Homomorphisms and colourings of oriented graphs: An updated survey*, Discrete Math. **339** (2016), 1993–2005.

[72] R. Stearns: *The voting problem*, Amer. Math. Monthly **66** (1959), 761–763.

[73] T. Szele: *Kombinatorische Untersuchungen über den gerichtelen vollstandigen Graphen*, Mat. Fiz. Lapok **50** (1943), 223–256.

[74] E. Szemerédi: *Regular partitions of graphs*, Colloques Internationaux C.N.R.S. Nº 260 — Problèmes Combinatoires et Théorie des Graphes, Orsay (1976), 399–401.

[75] P. Turán: *On an extremal problem in graph theory (in Hungarian)*, Mat. Fiz. Lapok **48** (1941), 436–452.

[76] P. Valadkhan: *Extremal oriented graphs and Erdős-Hajnal conjecture*, Master's thesis, Simon Fraser University, 2009.

[77] E.R. Vaughan: *Flagmatic: A tool for researchers in extremal graph theory*, further developed by J. Sliacan, available at `https://github.com/jsliacan/flagmatic-2.0`.

[78] R. Yuster: *On the exact maximum induced density of almost all graphs and their inducibility*, J. Combin. Theory Ser. B **136** (2019), 81–109.

[79] A.A. Zykov: *On some properties of linear complexes (in Russian)*, Mat. Sb. (N.S.), **24(66)**:2 (1949), 163–188.