Tournaments with forbidden substructures and the Erdös-Hajnal Conjecture

Krzysztof Choromanski

Submitted in partial fulfillment of the

Requirements for the degree

of Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2013

ABSTRACT

Tournaments with forbidden substructures and the Erdös-Hajnal Conjecture

Krzysztof Choromanski

A celebrated Conjecture of Erdös and Hajnal states that for every undirected graph H there exists $\epsilon(H) > 0$ such that every undirected graph on n vertices that does not contain H as an induced subgraph contains a clique or a stable set of size at least $n^{\epsilon(H)}$. In 2001 Alon, Pach and Solymosi proved ([2]) that the conjecture has an equivalent directed version, where undirected graphs are replaced by tournaments and cliques and stable sets by transitive subtournaments. This dissertation addresses the directed version of the conjecture and some problems in the directed setting that are closely related to it. For a long time the conjecture was known to be true only for very specific small graphs and graphs obtained from them by the so-called *substitution procedure* proposed by Alon, Pach and Solymosi in [2]. All the graphs that are an outcome of this procedure have *nontrivial homogeneous sets*. Tournaments without nontrivial homogeneous sets are called *prime*. They play a

central role here since if the conjecture is not true then the smallest counterexample is prime. We remark that for a long time the conjecture was known to be true only for some prime graphs of order at most 5. There exist 5-vertex graphs for which the conjecture is still open, however one of the corollaries of the results presented in the thesis states that all tournaments on at most 5 vertices satisfy the conjecture. In the first part of the thesis we will establish the conjecture for new infinite classes of tournaments containing infinitely many prime tournaments. We will first prove the conjecture for so-called *constellations*. It turns out that almost all tournaments on at most 5 vertices are either constellations or are obtained from constellations by substitutions. The only 5-vertex tournament for which this is not the case is a tournament in which every vertex has outdegree 2. We call this the tournament C_5. Another result of this thesis is the proof of the conjecture for this tournament. We also present here the structural characterization of the tournaments satisfying the conjecture in almost linear sense. In the second part of the thesis we focus on the upper bounds on coefficients $\epsilon(H)$ for several classes of tournaments. In particular we analyze how they depend on the structure of the tournament. We prove that for almost all h-vertex tournaments $\epsilon(H) \leq \frac{4}{h}(1 + o(1))$. As a byproduct of the methods we use here, we get upper bounds for $\epsilon(H)$ of undirected graphs. We also present upper bounds on $\epsilon(H)$ of tournaments with small nontrivial homogeneous sets, in particular prime tournaments. Finally we analyze tournaments with big $\epsilon(H)$ and explore some of their structural properties.

Table of Contents

Acknowledgments

I would like to thank very much Professor Maria Chudnovsky for telling me about the Erdös-Hajnal Conjecture and significant involvement in my research. Not only did she familiarize me with the problem but she was also constantly encouraging me to work on it. And besides we managed to prove lots of interesting results related to the conjecture together. I would like also to thank Professor Joel Spencer. His willingness to listen to me presenting my new results whenever I had some was something that I will remember for a long time. Professor Spencer has a unique talent to ask the right questions which is probably one of the most difficult things a mathematician should try to learn. The results about pseudo-celebrities are completely inspired by a conversation I had with him some time ago at Courant Institute. He predicted the existence of the family of tournaments satisfying the conjecture in almost linear sense and made me start working on its structural characterization. My research would certainly look completely different without Professor Paul Seymour. I was very lucky to work with him and publish several papers together. His ability to extend methods that work in some specific settings for very general scenarios and find elegant proofs that capture the heart of the matter had a huge impact on my own research. I would like to thank Professor Dave Bayer for convincing me to apply to IEOR Department for doctoral studies and work with Professor Chudnovsky. Finally I would like to thank my Parents for making me be interested in mathematics when I was a child and for their great support throughout all these years.

1

Introduction

1.1 Notation and basic definitions

We use $||$ to denote the size of the set. Let G be a graph. We denote by $V(G)$ the set of its vertices and by $E(G)$ the set of its edges. By $|G|$ we denote the number of vertices of G and call it the *size of* G. For a graph G and a subset $X \subseteq V(G)$ we denote by $G|X$ the subgraph of G induced by X and by $G \backslash X$ the graph obtained from G be deleting all vertices from X and all edges with at least one endpoint in X.

If an undirected graph G does not contain another undirected graph H as an induced subgraph then we say that G is H-free. A *clique* in an undirected graph is a set of pairwise adjacent vertices and a *stable set* in an undirected graph is a set of pairwise nonadjacent vertices. We denote by $\omega(G)$ the size of the largest clique and by $\alpha(G)$ the size of the largest stable set of the undirected graph G. For an undirected graph G we denote by G^c *the complement of* G, i.e. the graph with the same set of vertices as G and such that there is an edge between two vertices in G^c iff there is no edge between these two vertices in G.

An undirected graph G is *bipartite* if its vertex set $V(G)$ can be partitioned into two subsets V_1, V_2 such that no edge of G has both endpoints in V_1 or in V_2. We call subsets V_1, V_2 the *color classes* of G. Bipartite undirected graph G with color classes V_1, V_2 is *complete* if for any two given $v_1 \in V_1$, $v_2 \in V_2$ there is an edge $\{v_1, v_2\} \in E(G)$. A complete bipartite graph G in which the sizes of two its color classes differ by at most 1 is called a *bi-clique*. A *matching* in a bipartite graph is the set of edges such that no two of them are incident to the same vertex. A matching is called *perfect matching* if every vertex of a bipartite graph is incident to some edge of it.

A *tournament* is a directed graph such that for every two vertices v and w exactly one of the directed edges (v, w) or (w, v) exists. If (v, w) is an edge of the tournament then we say that v is *adjacent to* w and w is *adjacent from* v. In such a case vertex w is an *outneighbour* of a vertex v and vertex v is an *inneighbour* of a vertex w. The *outdegree* of a vertex v of a tournament T is the number of vertices of T adjacent from it. For two disjoint sets of vertices V_1, $V_2 \subseteq V(T)$ we say that V_1 is *complete to* V_2 (or equivalently V_2 is *complete from* V_1) if every vertex of V_1 is adjacent to every vertex of V_2. A tournament is *transitive* if it contains no directed cycle. For the set of vertices $V = \{v_1, v_2, ..., v_k\}$ we say that an ordering $(v_1, v_2, ..., v_k)$ is *transitive* if v_1 is adjacent to all other vertices of V, v_2 is adjacent to all other vertices of V but v_1, etc. If a tournament T does not contain another tournament H as a subtournament then we say that T is *H-free.* In this definition we consider tournaments to be unlabelled.

A coloring of a tournament H is an assignment of colors to its vertices such that no directed triangle is monochromatic. The minimal number of colors needed to color H is called *the chromatic number of* H and will be denoted as $\chi(H)$.

We denote by $t(H)$ the number of directed triangles of the n-vertex tournament H. A tournament H is called *δ-dense* if it contains at least $\delta|H|^3$ directed triangles (not necessarily edge-disjoint).

Let us say that $\epsilon \geq 0$ is an *EH-coefficient* for a tournament H if there exists $c > 0$ such that every H-free tournament G satisfies $\alpha(G) \geq c|G|^{\epsilon}$. (We introduce c in the definition of the Erdős-Hajnal coefficient to eliminate the effect of tournaments G of bounded order; now, whether ϵ is an EH-coefficient for H depends only on arbitrarily large tournaments not containing H.) For a fixed tournament H, define $\xi(H)$ to be the supremum of all ϵ for which the following holds: for some n_0 and every $n > n_0$ every H-free tournament with $n \geq n_0$ has a transitive subtournament of size at least n^ϵ. We call $\xi(H)$ the *EH-supremum of H.* For a fixed undirected graph G, define $\xi(G)$ to be the supremum of all ϵ for which the following holds: for some n_0 and every $n > n_0$ every G-free undirected graph with $n \geq n_0$ has a clique or a stable set of size at least n^ϵ. We call $\xi(G)$ the *EH-supremum of G.* The Erdös-Hajnal Conjecture is true if and only if $\xi(H) > 0$ for every H.

There exist tournaments H (so-called *celebrities*) that satisfy the conjecture in the strongest - linear sense. For a celebrity H there exists $c(H) > 0$ such that if a n-vertex tournament does not contain H as a subtournament then it contains a transitive subtournament of size at least $c(H)n$. All tournaments H with this property were described in [7]. However the question whether there exist tournaments satisfying the conjecture in almost linear sense remained open. A tournament H satisfies the conjecture in almost linear sense if it is not a celebrity but for every $0 < \epsilon < 1$ there exists n_ϵ such that for every $n > n_\epsilon$ every H-free n-vertex tournament contains a transitive subtournament of size at least n^ϵ. We prove in the thesis that tournaments with this property exist and describe all of them (this chapter is based on a joint work with Maria Chudnovsky and Paul Seymour, see [12]). As a corollary we describe all tournaments H with $\xi(H) > \frac{5}{6}$.

The EH-supremum for a tournament H is *not* necessarily itself an EH-coefficient for H; indeed, most of the Chapter 2 concerns finding tournaments H with $\xi(H) = 1$

for which 1 is not an EH-coefficient.

A subset of vertices $S \subseteq V(G)$ of a tournament G is called *homogeneous* if for every $v \in V(G)\backslash S$ the following holds: either $\forall_{w \in S}(w, v) \in E(G)$ or $\forall_{w \in S}(v, w) \in E(G)$. Similarly, a subset of vertices $S \subseteq V(G)$ of an undirected graph G is called *homogeneous* if for every $v \in V(G)\backslash S$ the following holds: either $\forall_{w \in S}\{w, v\} \in E(G)$ or $\forall_{w \in S}\{w, v\} \in E(G^c)$. A homogeneous set S is called *nontrivial* if $|S| > 1$ and $S \neq V(G)$. A graph is called *prime* if it does not have nontrivial homogeneous sets.

All logarithms used in the thesis except those in Chapter 2 are natural logarithms. All graphs considered in the thesis are finite, loopless and without multiple edges.

This thesis is organized as follows:

- in this chapter we introduce the conjecture and several definitions used later, we also show some previous results concerning the conjecture,
- in Chapter 2 we give a complete structural characterization of all tournaments satisfying the conjecture in almost linear sense,
- in Chapter 3 we give several definitons and technical lemmas used in Chapters 3 and Chapter 4,
- in Chapter 4 we formally define an infinite family of tournaments called constellations and prove the conjecture for tournaments from this family,

- in Chapter 5 we prove the conjecture for all tournaments on at most 5 vertices,
- in Chapter 6 we give upper bounds on EH-suprema for several classes of tournaments, in particular we present upper bounds for almost all tournaments and tournaments with small nontrivial homogeneous sets,
- in Chapter 7 we summarize all our results and mention some open problems related to the conjecture.

1.2 The conjecture

A celebrated unresolved Conjecture of Erdös and Hajnal ([18]) states that:

1.2.1 *For every undirected graph H there exists $\epsilon(H) > 0$ such that every n-vertex undirected graph that does not contain H as an induced subgraph contains a clique or a stable of size at least $n^{\epsilon(H)}$.*

In 2001 Alon, Pach and Solymosi proved ([2]) that Conjecture 1.2.1 has an equivalent directed version, where undirected graphs are replaced by tournaments and cliques and stable sets by transitive subtournaments.

The equivalent directed version ([2]) states that:

1.2.2 *For every tournament H there exists $\epsilon(H) > 0$ such that every n-vertex H-free tournament contains a transitive subtournament of size at least $n^{\epsilon(H)}$.*

We say that an undirected graph H *has the Erdös-Hajnal property* (or equivalently: *the conjecture is true for* H */ is satisfied by* H) if there exists $\epsilon(H) > 0$ such that every H-free n-vertex undirected graph contains a clique or a stable set of size at least $n^{\epsilon(H)}$. Similarly, we say that a tournament H *has the Erdös-Hajnal property* (or equivalently: *the conjecture is true for* H */ is satisfied by* H) if there exists $\epsilon(H) > 0$ such that every H-free n-vertex tournament T contains a transitive subtournament of size at least $n^{\epsilon(H)}$. The Erdös-Hajnal property is a *hereditary property*, i.e. if a graph H has the Erdös-Hajnal property then all its induced subgraphs also have the Erdös-Hajnal property.

Note first that the conjecture is true if and only if for every tournament H its EH-supremum $\xi(H)$ is positive. Equivalently, the conjecture is true if and only if for every undirected graph G its EH-supremum $\xi(G)$ is positive.

We cannot find stable sets or cliques of polynomial size in the undirected setting, or transitive subtournaments of polynomial size in the directed setting for an arbitrary graph if we do not assume anything about its structure. Indeed, if we take a random graph on n vertices, where for every pair of different vertices v,w we make v adjacent to w independently and with probability $\frac{1}{2}$ then with probability tending to 1 as $n \rightarrow \infty$ its biggest cliques and and stable sets are only of logarithmic order. Similarly, for an n-vertex tournament, in which for every pair of different vertices v, w we make v adjacent to w with probability $\frac{1}{2}$ independently for every pair, with probability tending to 1 as $n \to \infty$ its biggest transitive subtournaments are of logarithmic order. Therefore the conjecture states that if a graph H is fixed and we consider only the family of H-free graphs then the logarithmic sizes of the mentioned substructures may be replaced by polynomial sizes. If true, the conjecture says that the local condition of not having some forbidden subgraph H implies a global structural property of having some very simple large substructure. The emergence of a clique or a stable set of size $\omega(\log(n))$ has been already proven by Erdös and Hajnal in the same paper where the conjecture was stated ([18]).

1.2.3 *For any undirected graph H there exists $\epsilon(H) > 0$ such that every H-free n-vertex undirected graph contains a clique or a stable set of size at least $e^{\epsilon(H)\sqrt{\log(n)}}$.*

A similar theorem can be proven for tournaments. Thus the conjecture states that in the theorem above we can replace $\sqrt{\log(n)}$ factor by $\log(n)$ factor.

In the directed version of the conjecture one needs to prove the existence of one specific substructure - a transitive subtournament. In the undirected version we have substructures of two types: cliques and stable sets. However there is also a formulation of the undirected version of the conjecture that involves only one type of an induced substructure - a *perfect graph*. An undirected graph G is perfect if every induced subgraph H of G satisfies $\omega(H) = \chi(H)$.

The equivalent undirected version of the conjecture involving perfect graphs states that:

1.2.4 *For any undirected graph H there exists $\epsilon(H) > 0$ such that every H-free n-vertex undirected graph contains perfect induced subgraph of size at least $n^{\epsilon(H)}$.*

The equivalence follows directly from the following basic but very useful property of perfect graphs:

1.2.5 *If G is perfect then either $\omega(G) \geq \sqrt{|G|}$ or $\alpha(G) \geq \sqrt{|G|}$.*

The formulation 1.2.4 has an advantage over the original one since instead of dealing with two graph objects (cliques and stable sets) it deals with one - perfect induced subgraph.

There is a characterization of all perfect graphs that uses forbidden induced subgraphs (the *Strong Perfect Graph Theorem*, see [**13**]):

1.2.6 *A graph G is perfect if and only if no induced subgraph of G or G^c is an odd cycle of length at least five.*

1.3 Graphs with the Erdös-Hajnal property

1.3.1 The substitution procedure

For a long time the Erdös-Hajnal Conjecture was known to be true only for some graphs on at most 5 vertices and graphs obtained from them by the so-called *substitution procedure*, proposed by Alon, Pach and Solymosi in [2]. We define it now. Let H_1 and H_2 be either two undirected graphs or two tournaments with disjoint sets of vertices. Assume furthermore that $|V(H_1)|, |V(H_2)| \geq 2$. For a given $v \in V(H_1)$ we say that graph H (H is a tournament in the directed setting) *is obtained from* H_1 *by substituting* H_2 *for* v (or *is obtained from* H_1 *and* H_2 when we do not take care of details) if the following conditions are satisfied:

- $V(H) = (V(H_1) \bigcup V(H_2)) \backslash v$
- $H|(V(H_1) \backslash v) = H_1 \backslash v$
- $H|V(H_2) = H_2$
- vertex $u \in V(H_1)$ is adjacent in H to a vertex $w \in V(H_2)$ if and only if u is adjacent to v in H_1

In [2] Alon, Pach and Solymosi proved that the Erdös-Hajnal property is preserved under substitution:

1.3.1 *Let H_1 and H_2 be two graphs with the Erdös-Hajnal property (we assume that either both H_1 and H_2 are undirected or they are both tournaments). If H is obtained from H_1 and H_2 by a substitution procedure then H has the Erdös-Hajnal property.*

Let us remind that a graph H is prime if it does not have nontrivial homogeneous sets. Note that H is prime if and only if it is not obtained from smaller graphs by substitution. The substitution procedure was the only known procedure that allowed us to obtain infinitely many tournaments satisfying the conjecture. From Theorem 1.3.1 we know that if the conjecture is not true then the smallest counterexample is prime. That is why prime graphs are of special interests. Until very recently not much was known about the conjecture for prime graphs. The conjecture has been proven for most prime graphs on at most 5 vertices but it was open for all prime graphs on at least 6 vertices. Thus, in particular the question whether there are infinitely many prime tournaments satisfying the conjecture was open. We answer this question in this thesis in Chapter 4 by proving the conjecture for a new infinite family of tournaments containing infinitely many prime tournaments. Those results are based on [10].

1.3.2 Prime graphs with the Erdös-Hajnal property

We describe now what was known about the conjecture for prime graphs. We consider first the undirected scenario. All undirected graphs on at most three vertices trivially satisfy the conjecture. It turns out that the only prime undirected graphs on 4 vertices is the three-edge path. Theorem 1.2.6 implies that every undirected graph G that does not contain a three-edge path as an induced subgraph is perfect and thus, by 1.2.5, a three-edge path has the Erdös-Hajnal property. The other way to prove that the conjecture is satisfied by the three-edge path is by an induction and the result below:

1.3.2 *If G is an undirected graph with $|G| \geq 2$, H is a three-edge path and G is H-free then either G or G^c is not connected.*

The proof of this theorem may be found [**29**].

The prime undirected graphs on 5 vertices are:

- the cycle of length 5
- the four-edge path
- the complement of the four-edge path
- the graph with vertex set $\{v_1, v_2, v_3, w_1, w_2\}$ and edge set $\{v_1v_2, v_1v_3, v_2v_3, v_1w_1, v_2w_2\}$ (sometimes called the *bull*).

Thus, according to Theorem 1.3.1, all other graphs on at most 5 vertices have the Erdös-Hajnal property. The Conjecture is still open for the cycle of length 5, the four-edge path (thus also for every path of at least four edges) and the complement of the four-edge path. However it is known to be true for the bull.

In [**14**] it has been proven that:

1.3.3 *Every bull-free undirected graph G contains a clique or a stable set of size at least $|G|^{\frac{1}{4}}$.*

We note here that the coefficient $\frac{1}{4}$ in the theorem above cannot be improved. Indeed, take a graph B which is triangle-free and does not contain stable sets of size larger than $\sqrt{|B|\log(|B|)}$ (for a construction look here: [**23**]). If G is obtained from B by substituting a copy of B^c for every vertex of B then G is triangle-free and contains no clique or stable set of size larger than $2\sqrt{|B|\log(|B|)}$.

We will now introduce the notion of α-narrowness which is helpful while working on the undirected version of the conjecture for prime graphs. A function $f : V(G) \rightarrow [0,1]$ is *good* if the following holds for every perfect induced subgraph P of the undirected graph G:

$$\sum_{v \in V(P)} f(v) \leq 1. \tag{1.1}$$

We say that an undirected graph G is α-narrow if for every good function f the following holds:

$$\sum_{v \in V(G)} f(v)^{\alpha} \leq 1. \tag{1.2}$$

Note that perfect graphs are 1-narrow. We already know, by 1.2.5, that perfect graphs have cliques or stable sets of polynomial size. It is easy to see that for every $\alpha \geq 1$ every α-narrow graph has a clique or a stable set of a polynomial size. Let $M = \max |V(P)|$, where the maximum is taken over all perfect induced subgraphs P of G. Therefore if we take $f = \frac{1}{M}$, then 1.1 is trivially satisfied, so f is good. Since G is α-narrow, by 1.2, we get $M \geq |G|^{\frac{1}{\alpha}}$. Thus, by Theorem 1.2.5, we conclude that G has a clique or a stable set of size at least $|G|^{\frac{1}{2\alpha}}$. Therefore if one can prove that for a given undirected graph H every H-free graph is α-narrow for some $\alpha \geq 1$, that immediately implies the Erdös-Hajnal property for H. The following conjecture was stated in [**16**]:

1.3.4 *For every undirected graph H there exists $\alpha(H) \geq 1$ such that every H-free graph is $\alpha(H)$-narrow.*

This conjecture has been proven to be equivalent to the Erdös-Hajnal Conjecture by Fox ([**22**]). The approach that uses α-narrowness turned out to be useful while working on particular special cases of the Erdös-Hajnal Conjecture. Theorem 1.3.3 has been proven by showing that every bull-free graph is 2-narrow. We have already mentioned that the Erdös-Hajnal property is preserved by the substitution procedure. Note that α-narrowness is also preserved by this procedure. In [14] it was shown that:

1.3.5 *If H_1, H_2 are undirected α-narrow graphs for some $\alpha \geq 1$ and H is obtained from H_1 and H_2 by substitution, then H is α-narrow.*

Let us describe now what was known for prime tournaments. The conjecture is trivially true if we consider tournaments on at most three vertices. There are no prime tournaments on four vertices. Thus, the conjecture was known to be true for all tournaments on at most four vertices. However there exist prime tournaments on five vertices. Some of them were proven to be heroes in [7] therefore they satisfy the conjecture in the strongest - linear sense. However for other prime five-vertex tournaments the conjecture was open. In particular, it was open for C_5 - a unique tournament on five vertices, where each vertex has outdegree two. One of the results of this thesis is that all tournaments on at most 5 vertices satisfy the conjecture.

1.4 Excluding families of graphs

For a set $\mathcal{S}$ of graphs (either all undirected or all tournaments) we say that a graph G (where G is undirected in the undirected scenario and is a tournament in the directed one) is $\mathcal{S}$-*free* if it is H-free for every $H \in \mathcal{S}$. Instead of analyzing the biggest cliques/stable sets in the undirected scenario or transitive subtournaments (in the directed one) of H-free graphs one may try to do the same for $\mathcal{S}$-free graphs, where $\mathcal{S}$ is some set of graphs (undirected graphs in the undirected scenario and tournaments in the directed scenario). We say that a set of undirected graphs $\mathcal{S}$ has the Erdös-Hajnal property if there exists $\epsilon(\mathcal{S}) > 0$ such that every n-vertex $\mathcal{S}$-free undirected graph contains a clique or a stable set of size at least $n^{\epsilon(\mathcal{S})}$. Similarly, we say that a set of tournaments $\mathcal{S}$ has the Erdös-Hajnal property if there exists $\epsilon(\mathcal{S}) > 0$ such that every n-vertex $\mathcal{S}$-free tournament contains a transitive subtournament of size at least $n^{\epsilon(\mathcal{S})}$.

In particular, some previous results concern excluding pairs of graphs. One can propose the following conjecture which is a weaker version of the Erdös-Hajnal Conjecture:

1.4.1 *For every undirected graph H the two-element family $\{H, H^c\}$ has the Erdös-Hajnal property.*

This conjecture is still open too. However strengthening the condition put on the graph G by excluding several graphs rather than just one allows us to use methods that are not useful when $|\mathcal{S}| = 1$. We state here few results that deal with excluding pairs of undirected graphs. In [16] it was shown that:

1.4.2 *If H_1 is a four-edge path and H_2 is a complement of the five-edge path then* *an $\{H_1, H_2\}$-free undirected n-vertex graph contains a clique or a stable set of size at least $n^{\frac{1}{6}}$.*

Chudnovsky and Seymour proved in [**15**] the following:

1.4.3 *Let H be a five-edge path. Then the family $\{H, H^c\}$ has the Erdös-Hajnal property.*

Thus, a five-edge path satisfies Conjecture 1.4.1.

In the same paper they proved also the following related result:

1.4.4 *Let H_1 be a six-edge path and H_2 be a four-edge path. Then the family $\{H_1, H_2^c\}$ has the Erdös-Hajnal property.*

In Chapter 7 we propose an open problem in which a family of tournaments is excluded.

1.5 Polynomial-size cliques or stable sets in geometric graphs and some general approximate results

Even though the conjecture is still open, much work was done to prove related approximate results. Some of those results concern finding in H-free graphs polynomial-size substructures that are somehow close to cliques and stable sets. In others, special families of undirected graphs with a clique or a stable set of polynomial size are considered. In this section we focus on these types of problems. Most of the problems we consider here are for the undirected setting. However we will also state an open problem concerning tournaments. Its analogous undirected version was solved.

Erdös, Hajnal and Pach proved in [19] that for every undirectd graph H there exists $\epsilon(H) > 0$ such that every H-free n-vertex undirected graph G satisfies the following: either G or G^c contains a bi-clique of at least $n^{\epsilon(H)}$ vertices (similar result is true for tournaments). However this result does not say anything about graphs induced by color classes of the bi-clique. The result was strengthened by Fox and Sudakov who proved in [21] that:

1.5.1 *For every undirected graph H there exists $\epsilon(H) > 0$ such that every H-free n-vertex undirected graph G contains either a stable set of size at least $n^{\epsilon(H)}$ or a bi-clique of at least $n^{\epsilon(H)}$ vertices.*

We say that a hereditary family of undirected graphs $\mathcal{F}$ is *good* if there exists a constant $\epsilon > 0$ such that every n-vertex graph $F \in \mathcal{F}$ contains either a clique or a stable set of size at least n^{ϵ}. The Erdös-Hajnal Conjecture states that for every undirected graph H the family of all H-free undirected graphs is good (an analogous version using goodness property can be formulated for tournaments).

Checking the goodness property for a class of H-free graphs (for some fixed H) is very difficult in general. However some research was done to prove the goodness property for other families of undirected graphs such that graphs that may be described as intersection graphs of some geometric objects.

Before showing some of those results we need to introduce one more definition. We say that a hereditary family $\mathcal{F}$ of undirected graphs is *strongly good* if the following holds for some $\epsilon > 0$: for every n-vertex graph $F \in \mathcal{F}$ either F or F^c contains a bi-clique of at least ϵn vertices. It was observed in [1] that:

1.5.2 *If a hereditary family $\mathcal{F}$ of undirected graphs is strongly good then it is also good.*

Thus to prove the Erdös-Hajnal Conjecture it suffices to prove that for every undirected graph H the family of H-free graphs is strongly good. Unfortunately this statement is not true. Indeed, for instance the family of triangle-free graphs is not strongly good. However some families of geometric graphs do have this property. We will focus here on the family of intersection graphs of some two-dimensional objects. For a finite family $\mathcal{O}$ of two-dimensional objects we define its *intersection graph* as an undirected graph G such that $V(G) = \mathcal{O}$ and two vertices of G are adjacent iff they intersect as geometric objects. Consider a plane with given coordinate system.We say that a planar connected set is *vertically convex* if its intersection with any vertical line is an interval. Larman *et al.* proved in [24] that the hereditary family of intersection graphs of convex planar sets is good. They showed that:

1.5.3 *Any family of n vertically convex sets in the plane contains at least $n^{\frac{1}{5}}$ members that are either pairwise disjoint or pairwise intersecting.*

We say that a continuous curve is *x-monotone* if it intersects every vertical line in at most one point. Clearly, every x-monotone curve is vertically convex. However for the hereditary family of intersection graphs of x-monotone curves other results can be proven. In [20] it was shown that:

1.5.4 *There exists a constant $c > 0$ with the property that the intersection graph G of any collecton of n x-monotone curves in the plane satisfies at least one of the following conditions:*

- *G contains a bi-clique of size at least $\frac{cn}{\log(n)}$; or*
- *G^c contains a bi-clique of size at least cn.*

The family of intersection graphs of convex bodies in R^d for $d \geq 3$ is not good. However some special families of intersection graphs of objects taken from high-dimensional space are good. An example is the family of so-called *K-fat* sets. For a given $d \geq 1$ we say that a set $S \subseteq R^d$ is *K-fat* if there exist d-dimensional balls

B_1, B_2 with radii $R_1, R_2 > 0$ such that $B_1 \subseteq S \subseteq B_2$ and $\frac{R_2}{R_1} \leq K$. In [27] it has been proven that:

1.5.5 *For any constant $K \geq 1$ and for any positive integer d, the family of intersection graphs of K-fat convex bodies in R^d is strongly good.*

We saw above that certain hereditary families of graphs have the Erdös-Hajnal property. The problem whether for every undirected graph H the family $\mathcal{F}_H$ of H-free graphs is good is still open. However if we do not consider all H-free graphs but *almost all* then the goodness property can be proven. For any $0 < \epsilon < 1$ and an undirected graph H let $\mathcal{F}_H^{\epsilon} \subseteq \mathcal{F}_H$ be the family of those H-free undirected graphs G that contain a clique or a stable set of size at least $|G|^{\epsilon}$. If the Erdös-Hajnal Conjecture is true then there exists $\epsilon > 0$ such that $\mathcal{F}_H = \mathcal{F}_H^{\epsilon}$. It has been proven in [25] that:

1.5.6 *For every undirected graph H there exists $\epsilon > 0$ such that* $\lim_{n\to\infty} \frac{|\mathcal{F}_H^{\epsilon}|}{|\mathcal{F}_H|} = 1$.

We may consider an analogous problem for tournaments. Let H be a tournament and let $\mathcal{F}_H$ be the family of H-free tournaments. For a parameter $0 < \epsilon < 1$ let $\mathcal{F}_H^{\epsilon} \subseteq \mathcal{F}_H$ be the family of those H-free tournaments T that contain a transitive subtournament of size at least $|T|^{\epsilon}$. The following is still open:

1.5.7 *For every tournament H there exists $\epsilon > 0$ such that* $\lim_{n\to\infty} \frac{|\mathcal{F}_H^{\epsilon}|}{|\mathcal{F}_H|} = 1$.

2

Pseudo-celebrities

2.1 Introduction

There are some tournaments H with the property that every H-free tournament has chromatic number at most a constant (depending on H). These are called *heroes*, and they were all explicitly described in [7]. In this chapter we describe the most heroic non-heroes. All results of this chapter is a joint work with Maria Chudnovsky and Paul Seymour ([12]). It turns out that for some non-heroes H, the chromatic number of every H-free tournament G is at most a polylog function of the number of vertices of G, and all the others give nothing better than a polynomial bound. We prove in this chapter that:

2.1.1 *Every tournament has exactly one of the following properties:*

- *for some c, every H-free tournament has chromatic number at most c (the heroes)*
- *for some c, d, every H-free tournament G with $|G| > 1$ has chromatic number at most $c(\log(|G|))^d$, and for all c, there are H-free tournaments G with $|G| > 1$ and with chromatic number at least $c(\log(|G|))^{1/3}$*
- *for all c, there are H-free tournaments G with $|G| > 1$ and with chromatic number at least $c|G|^{1/6}$.*

We also give an explicit construction for all tournaments of the second type, which we call *pseudo-heroes*.

Consider the following problem closely related to the Erdös-Hajnal Conjecture: for which tournaments is some given $\epsilon > 0$ an EH-coefficient ? In [7], this question was completely answered for $\epsilon = 1$; and our goal in this chapter is a similar result for $\epsilon > 5/6$.

Before we go on, let us state the result of [7] properly; and to do so we need some more definitions and denotations. In this chapter we denote by T_k the transitive tournament with k vertices. If G is a tournament and X, Y are disjoint subsets of $V(G)$ such that X is complete to Y, we write $X \Rightarrow Y$. We write $v \Rightarrow Y$ for $\{v\} \Rightarrow Y$, and $X \Rightarrow v$ for $X \Rightarrow \{v\}$. If G is a tournament and (X, Y, Z) is a partition of $V(G)$

into nonempty sets satisfying $X \Rightarrow Y$, $Y \Rightarrow Z$, and $Z \Rightarrow X$, we call (X, Y, Z) a *trisection* of G. If A, B, C, G are tournaments, and there is a trisection (X, Y, Z) of G such that $G|X, G|Y, G|Z$ are isomorphic to A, B, C respectively, we write $G = \Delta(A, B, C)$. It is convenient to write k for T_k here, so for instance $\Delta(1, 1, 1)$ means $\Delta(T_1, T_1, T_1)$, and $\Delta(H, 1, k)$ means $\Delta(H, T_1, T_k)$.

A tournament is a *celebrity* if 1 is an EH-coefficient for it; that is, for some $c > 0$, every H-free tournament G satisfies $\alpha(G) \geq c|G|$. The main result of [7] is:

2.1.2 *The following hold:*

- *A tournament is a hero if and only if it is a celebrity.*
- *A tournament is a hero if and only if all its strong components are heroes.*
- *A strongly-connected tournament with more than one vertex is a hero if and only if it equals $\Delta(1, H, k)$ or $\Delta(1, k, H)$ for some hero H and some integer $k > 0$.*

In this chapter, we study the tournaments H which are "almost" heroes, in the sense that all H-free tournaments have chromatic number at most a polylog function of their order. More precisely, we say a tournament H is

- a *pseudo-hero* if there exist $c, d \geq 0$ such that every H-free tournament G with $|G| > 1$ satisfies $\chi(G) \leq c(\log(|G|))^d$

- a *pseudo-celebrity* if there exist $c > 0$ and $d \geq 0$ such that every H-free tournament G with $|G| > 1$ satisfies $\alpha(G) \geq c\frac{|G|}{(\log(|G|))^d}$.

Logarithms are of base two throughout this chapter. The conditions $|G| > 1$ are included just to ensure that $\log(|G|) > 0$.) The next result is an analogue of 5.2.1:

2.1.3 *The following hold:*

- *A tournament is a pseudo-hero if and only if it is a pseudo-celebrity.*
- *A tournament is a pseudo-hero if and only if all its strong components are pseudo-heroes.*
- *A strongly-connected tournament with more than one vertex is a pseudo-hero if and only if either*
  - *it equals $\Delta(2, k, l)$ for some $k, l \geq 2$, or*
  - *it equals $\Delta(1, H, k)$ or $\Delta(1, k, H)$ for some pseudo-hero H and some integer $k > 0$.*

More generally, let $0 \leq \epsilon \leq 1$; we say that a tournament H is

- an *ϵ-hero* if there exist $c, d \geq 0$ such that every H-free tournament G with $|G| > 1$ satisfies $\chi(G) \leq c|G|^{1-\epsilon}\log(|G|)^d$; and

- an ϵ-*celebrity* if there exist $c > 0$ and $d \geq 0$ such that every H-free tournament G with $|G| > 1$ satisfies $\alpha(G) \geq c^{-1}|G|^{\epsilon} \log(|G|)^{-d}$.

Thus, a 1-hero is the same thing as a pseudo-hero, and a 1-celebrity is the same as a pseudo-celebrity. We will prove:

2.1.4 *For all ϵ with $0 \leq \epsilon \leq 1$:*

- *a tournament is an ϵ-hero if and only if it is an ϵ-celebrity*
- *a tournament is an ϵ-celebrity if and only if its strong components are ϵ-celebrities*
- *if H is an ϵ-celebrity and $k \geq 1$, then $\Delta(1, H, k)$ and $\Delta(1, k, H)$ are ϵ-celebrities.*

(Much of 2.1.3 is implied by setting $\epsilon = 1$ in 2.1.4.) In addition, we will prove the following theorem:

2.1.5 *Every tournament H with $\xi(H) > 5/6$ is a pseudo-hero and hence satisfies $\xi(H) = 1$.*

Thus, if $\xi(H) > 5/6$ then every H-free tournament has chromatic number at most a polylog function of its order. We do not know if 5/6 is best possible; but the polylog behaviour is best possible, in the following sense:

2.1.6 *For every real d with $0 \leq d < \frac{1}{3}$ and all sufficiently large integers n (depending on d), there is a tournament G with n vertices such that*

- $\alpha(G) \leq n(\log(n))^{-d}$, *and*
- *every pseudo-hero contained in G is a hero.*

This last is a corollary of a result of [7]; let us see that now. Since every pseudo-hero that is not a hero contains $\Delta(2,2,2)$, by 5.2.1 and 2.1.3, it follows that 2.1.6 is implied by the following result of [7]:

2.1.7 *For every real d with $0 \leq d < \frac{1}{3}$, and all sufficiently large integers n (depending on d), there is a tournament G with n vertices, not containing $\Delta(2,2,2)$, such that*

$$\alpha(G) \leq \frac{n}{(\log(n))^d}.$$

(More precisely, the result of [7] asserts this with $\log(n)$ replaced by $\ln(n)$; we leave the reader to check the equivalence.) The chapter is organized as follows:

- in Sections 2.2,2.3 and 2.4 we prove the first, second and third assertion of 2.1.4 respectively;
- in Section 2.5 we prove that for all $k,l \geq 2$, $\Delta(2,k,l)$ is a pseudo-celebrity, and indeed there exists $c > 0$ such that every $\Delta(2,k,l)$-free tournament G with $|G| > 1$ satisfies $\alpha(G) \geq c|G|/\log(|G|)$;

- in Section 2.6 we prove the "only if" part of the third statement of 2.1.3, and thereby finish the proof of 2.1.3; and we also prove 2.1.5.

2.2 ϵ-celebrities are ϵ-heroes

In this section we prove the first statement of 2.1.4. Let us say a function ϕ is *round* if for each integer $n \geq 2$, $\phi(n)$ is a real number, at least 1 and (non-strictly) increasing with n. We need:

2.2.1 *Let ϕ be round. Suppose that G is a tournament with $|G| > 1$, and for all $n > 1$, every n-vertex subtournament of G has a transitive set of cardinality at least $n/\phi(n)$. Then $\chi(G) \leq \phi(|G|)\log(|G|)$.*

Proof. We proceed by induction on $|G|$. Let $n = |G|$. By hypothesis, G has a transitive set X of cardinality x say, where $x \geq n/\phi(n) > 0$. Thus $1 \leq \phi(n)\log(n)$ (since $\phi(n) \geq 1$, and logarithms are to base 2), and so we may assume that $\chi(G) \geq 2$. In particular, $x \leq n-1$, and so $n-1 \geq n/\phi(n)$. Consequently $\phi(n) \geq n/(n-1) \geq 2/\log(n)$, and so $2 \leq \phi(n)\log(n)$. Hence we may assume that $\chi(G) \geq 3$. In particular, $G \setminus X$ has at least two vertices, and therefore we may apply the inductive hypothesis to $G \setminus X$. Since $\chi(G) \leq 1 + \chi(G \setminus X)$, we deduce that

$$\chi(G) \leq 1 + \phi(n-x)\log(n-x) \leq 1 + \phi(n)\log(n-x).$$

Now

$$\log(1 - x/n) \leq \ln(1 - x/n) \leq -x/n \leq -(\phi(n))^{-1},$$

and so $1 + \phi(n)\log(1 - x/n) \leq 0$. Consequently

$$\chi(G) \leq 1 + \phi(n)\log(n - x) = 1 + \phi(n)\log(1 - x/n) + \phi(n)\log(n) \leq \phi(n)\log(n).$$

This proves 2.2.1. ∎

Sometimes the previous result can be improved:

2.2.2 *Let G be a tournament with $|G| > 0$, and for each integer n with $1 \leq n \leq |G|$, let $\phi(n)$ be a positive real number, and let ϵ be a real number with $0 < \epsilon \leq 1$, such that*

- *every subtournament H of G with $|H| > 0$ has a transitive set of cardinality at least $|H|/\phi(|H|)$, and*
- *$\phi(n)/\phi(m) \geq (n/m)^\epsilon$ for all m, n with $1 \leq m \leq n \leq |G|$.*

Let $c = 2^\epsilon - 1$. Then $\chi(G) \leq c^{-1}\phi(|G|)$.

Proof. We proceed by induction on $|G|$. Let $n = |G|$. From the hypothesis, there is a transitive subset with cardinality at least $n/\phi(n) \geq 2^{\epsilon-1}n/\phi(n)$. Let us choose $X_1, \ldots, X_k \subseteq V(G)$, pairwise disjoint and each transitive with cardinality at least

$2^{\epsilon-1}n/\phi(n)$, with k maximal; it follows that $k \geq 1$. Let $X_1 \cup \cdots \cup X_k = W$, and let $G \setminus W = G'$, and $|G'| = n'$. Let $x = n'/n$. Now W includes k disjoint subsets of cardinality at least $2^{\epsilon-1}n/\phi(n)$, and so

$$n - n' = |W| \geq k2^{\epsilon-1}n/\phi(n),$$

that is, $k \leq (1-x)\phi(n)2^{1-\epsilon}$. If $n' = 0$, then

$$\chi(G) \leq k \leq \phi(n)2^{1-\epsilon} \leq c^{-1}\phi(|G|),$$

as required. Thus we may assume that $n' > 0$. Now G' has no transitive set of cardinality at least $2^{\epsilon-1}n/\phi(n)$ by the maximality of k, and yet by hypothesis, it has a transitive set of cardinality at least $n'/\phi(n')$. It follows that $n'/\phi(n') < 2^{\epsilon-1}n/\phi(n)$, that is,

$$\phi(n')/\phi(n) > 2^{1-\epsilon}x.$$

By hypothesis, $\phi(n')/\phi(n) \leq x^{\epsilon}$, and so $2^{1-\epsilon}x < x^{\epsilon}$, that is, $x < 1/2$. From the inductive hypothesis, $\chi(G') \leq c^{-1}\phi(n')$. Since $\chi(G) \leq \chi(G') + k$, and $k \leq (1-x)\phi(n)2^{1-\epsilon}$, we deduce that

$$\chi(G) \leq c^{-1}\phi(n') + (1-x)\phi(n)2^{1-\epsilon}.$$

Since $\phi(n') \leq \phi(n)x^\epsilon$, it follows that

$$c\chi(G)/\phi(G) \leq x^\epsilon + (1-x)2^{1-\epsilon}c.$$

Now the function $(1-x^\epsilon)/(1-x)$ is minimized for $0 \leq x \leq 1/2$ when $x = 1/2$, and its value then is $2^{1-\epsilon}c$; and so $(1-x^\epsilon)/(1-x) \geq 2^{1-\epsilon}c$, that is,

$$x^\epsilon + (1-x)2^{1-\epsilon}c \leq 1.$$

It follws that $c\chi(G)/\phi(G) \leq 1$. as required. This proves 2.2.2. ∎

Thus if ϕ grows sufficiently quickly then we can avoid the extra log factor introduced by 2.2.1. Curiously, it has been proven in [7] that the same is true when ϕ is constant. We do not know whether it is also true in the cases in between, when ϕ is not constant but only grows slowly. Unfortunately, these are the cases of most interest to us in this chapter, and for them we have to make do with 2.2.1.

We deduce the first statement of 2.1.4, namely:

2.2.3 *For $0 \leq \epsilon \leq 1$, a tournament is an ϵ-hero if and only if it is an ϵ-celebrity.*

Proof. Let H be an ϵ-celebrity, and choose $c > 0$ and $d \geq 0$ such that every H-free tournament G with $|G| > 1$ satisfies $\alpha(G) \geq c^{-1}|G|^\epsilon \log(|G|)^{-d}$. We may assume that $c \geq 1$. Define $\phi(n) = cn^{1-\epsilon}(\log(n))^d$ for $n \geq 2$. Thus ϕ is round, and every

H-free tournament G with $|G| > 1$ satisfies $\alpha(G) \geq |G|/\phi(|G|)$. Then if G is H-free and $|G| > 1$, the hypotheses of 2.2.1 are satisfied, and so

$$\chi(G) \leq \phi(|G|)\log(|G|) \leq c|G|^{1-\epsilon}(\log(|G|))^{d+1},$$

and therefore H is an ϵ-hero. (Note that, if $\epsilon < 1$, we could apply 2.2.2 here instead, and avoid the extra log factor.)

For the converse, let H be an ϵ-hero. Thus there exist $c, d \geq 0$ such that every H-free tournament G with $|G| > 1$ satisfies $\chi(G) \leq c|G|^{1-\epsilon}(\log(|G|))^{d}$. But every non-null tournament G has a transitive set of cardinality at least $|G|/\chi(G)$ (take the largest set of the partition given by the colouring). Consequently, every H-free tournament G with $|G| > 1$ has a transitive set of cardinality at least $c^{-1}|G|^{\epsilon}(\log(|G|))^{-d}$. It follows that H is an ϵ-celebrity. This proves 6.3.6. ∎

2.3 ϵ-celebrities that are not strongly connected

In this section we prove the second statement of 2.1.4, the following.

2.3.1 *For $0 \leq \epsilon \leq 1$, a tournament is an ϵ-celebrity if and only if all its strong components are ϵ-celebrities.*

We need the following theorem of [6].

2.3.2 *For every tournament H and every real $\lambda > 0$ there exists a real $c > 0$ with the following property. For every H-free tournament G there exist disjoint subsets $X, Y \subseteq V(G)$ with $|X|, |Y| = \lceil c|V(G)|\rceil$, such that $d(X, Y) < \lambda$.*

Let H_1, H_2 be tournaments. Let G be a tournament such that there is a partition (V_1, V_2) of $V(G)$ with $V_1 \Rightarrow V_2$, where for $i = 1, 2$, the subtournament of G with vertex set V_i is isomorphic to H_i. We denote such a tournament G by $H_1 \Rightarrow H_2$. For two sets of tournaments $\mathcal{F}_1$ and $\mathcal{F}_2$, we denote by $\mathcal{F}_1 \Rightarrow \mathcal{F}_2$ the set consisting of all tournaments (up to isomorphism) of the form $H_1 \Rightarrow H_2$ for some $H_1 \in \mathcal{F}_1$ and $H_2 \in \mathcal{F}_2$. For a set $\mathcal{F}$ of tournaments, we say that a tournament T is *$\mathcal{F}$-free* if no subtournament of T is isomorphic to a member of $\mathcal{F}$. We need the following lemma.

2.3.3 *Let $h \geq 1$ be an integer, and let $\mathcal{F}_1$ and $\mathcal{F}_2$ be two sets of tournaments, where each tournament in $\mathcal{F}_1 \cup \mathcal{F}_2$ has at most h vertices. Then there exists $C > 0$ with the following property. Let ϕ be round, such that for $i = 1, 2$, every $\mathcal{F}_i$-free tournament T of order $n > 1$ satisfies $\alpha(T) \geq n/\phi(n)$. Then every ($\mathcal{F}_1 \Rightarrow \mathcal{F}_2$)-free tournament T of order $n > 1$ satisfies $\alpha(T) \geq Cn/\phi(n)$.*

Proof. If one of $\mathcal{F}_1$ and $\mathcal{F}_2$ is empty, the result is trivial, so we assume both are non-empty, and hence $\mathcal{F}_1 \Rightarrow \mathcal{F}_2$ is nonempty. Choose one of its members, H_0 say. Choose $c > 0$ satisfying 2.3.2, taking $H = H_0$ and $\lambda = (4h)^{-1}$. Let $C = c/2$. We will show that C satisfies the theorem.

Let T be an $(\mathcal{F}_1 \Rightarrow \mathcal{F}_2)$-free tournament with $n > 1$ vertices. By 2.3.2, there exist disjoint $V_1, V_2 \subseteq V(G)$ with $|V_1|, |V_2| \geq c|V(T)|$ such that $d(V_2, V_1) < (4h)^{-1}$. Let X be the set of all vertices in V_1 with at least $(1-(2h)^{-1})|V_2|$ out-neighbours in V_2. Every vertex in $V_1 \setminus X$ is adjacent from at least $(2h)^{-1}|V_2|$ members of V_2, and so

$$|V_1 \setminus X|(2h)^{-1}|V_2| \leq (4h)^{-1}|V_1||V_2|,$$

that is, $|X| \geq |V_1|/2$.

Now $|V_1| \geq cn$. Suppose that $T|X$ is $\mathcal{F}_1$-free. From the hypothesis, X includes a transitive subset of cardinality at least $|X|/\phi(|X|)$; but $\phi(|X|) \leq \phi(n)$, and $|X| \geq cn/2$, and so $\alpha(T) \geq Cn/\phi(n)$ as required. Thus we may assume that there exists $X' \subseteq X$ such that $T|X'$ is isomorphic to some member H_1 of $\mathcal{F}_1$. For each $x \in X'$, at most $(2h)^{-1}|V_2|$ vertices in V_2 are adjacent to x, since $x \in X$; and since $|X'| \leq h$, it follows that at most $|V_2|/2$ vertices in V_2 are adjacent to a vertex in X'. Let Y be the set of all $y \in V_2$ that are adjacent from every vertex in X'; then $|Y| \geq |V_2|/2$. Since T is $(\mathcal{F}_1 \Rightarrow \mathcal{F}_2)$-free, it follows that $T|Y$ is $\mathcal{F}_2$-free; and so from the hypothesis, Y includes a transitive subset of cardinality at least $|Y|/\phi(|Y|)$. But $\phi(|Y|) \leq \phi(n)$, and

$$|Y| \geq |V_2|/2 \geq cn/2 = Cn,$$

and so $\alpha(G) \geq Cn/\phi(n)$. This proves 2.3.3. ∎

Proof of 2.3.1. Since every subtournament of an ϵ-celebrity is an ϵ-celebrity, the "only if" part of 2.3.1 is clear. The "if" part is implied by 2.3.3, taking $\phi(n) = cn^{1-\epsilon}(\log(n))^d$ for appropriate c, d. This proves 2.3.1. ∎

2.4 Adding handles

To complete the proof of 2.1.4, we need to show the following, which is proved in this section:

2.4.1 *For $0 \leq \epsilon \leq 1$, let H be an ϵ-hero, and let $k \geq 1$ be an integer. Then $\Delta(H, 1, k)$ and $\Delta(k, 1, H)$ are ϵ-heroes.*

We prove, more generally:

2.4.2 *Let H be a tournament, and let ϕ be round, such that every H-free tournament G satisfies $\chi(G) \leq \phi(|G|)$. Let $k \geq 1$ be an integer. Then there exists $c \geq 0$ such that every $\Delta(H, 1, k)$-free tournament G satisfies $\chi(G) \leq c\phi(G)\log(|G|)$, and the same for $\Delta(k, 1, H)$.*

We remark that if ϕ grows sufficiently quickly to satisfy the hypotheses of 2.2.2 we could use the latter to avoid the extra log factor.

Let H, K be tournaments, and let $a \geq 1$ be an integer. An (a, H, K)-*jewel* in a tournament G is a subset $X \subseteq V(G)$ such that $|X| = a$, and for every partition (A, B) of X, either $G|A$ contains H or $G|B$ contains K. An (a, H, K)-*jewel-chain* of length t is a sequence $Y_1, \ldots, Y_t$ of (a, H, K)-jewels, pairwise disjoint, such that $Y_i \Rightarrow Y_{i+1}$ for $1 \leq i < t$. We need the following lemma, proved in [7]:

2.4.3 *Let H, K be tournaments, and let $a \geq 1$ be an integer. Then there are integers $\lambda_1, \lambda_2 \geq 0$ with the following property. For every $\Delta(H, 1, K)$-free tournament G, if*

- *c_1 is such that every H-free subtournament of G has chromatic number at most c_1, and every K-free subtournament of G has chromatic number at most c_1, and*
- *c_2 is such that every subtournament of G containing no (a, H, K)-jewel-chain of length four has chromatic number at most c_2,*

then G has chromatic number at most $\lambda_1 c_1 + \lambda_2 c_2$.

Proof of 2.4.2, 2.4.1 and 2.1.4. Let K be a transitive tournament with k vertices; from the symmetry, it suffices to show the result for $\Delta(H, 1, K)$. Let ϕ be as in the hypothesis of the theorem. We may assume that $\phi(2) \geq 2^k$, by scaling ϕ. Let $a = 2^k|V(H)|$, and let $\lambda_1, \lambda_2 \geq 0$ be as in 2.4.3.

(1) *If G is a tournament with $|G| > 1$, not containing an (a, H, K)-jewel, then $\chi(G) \leq a\phi(|G|)$.*

Choose pairwise vertex-disjoint subtournaments $H_1, \ldots, H_t$ of G, each isomorphic to H, with t maximum, and let the union of their vertex sets be W. If $t \geq 2^k$, then since every tournament with at least 2^k vertices has a transitive subset of cardinality k, it follows that $V(H_1) \cup \cdots \cup V(H_{2^k})$ is an (a, H, K)-jewel, a contradiction. Thus $t < 2^k$. Consequently $\chi(G|W) \leq |W| \leq a$, and $\chi(G \setminus W) \leq \phi(|G| - |W|) \leq \phi(|G|)$ since $G \setminus W$ is H-free. It follows that $\chi(G) \leq a + \phi(|G|) \leq a\phi(|G|)$ since $a, \phi(|G|) \geq 2$. This proves (1).

(2) *There exists $C \geq 0$ such that if G is a tournament with $|G| > 1$, not containing an (a, H, K)-jewel-chain of length four, then $\chi(G) \leq C\phi(G)\log(|G|)$.*

By (1), if G is a tournament with $n > 1$ vertices, not containing an (a, H, K)-jewel, then $\alpha(G) \geq a^{-1}n/\phi(n)$. By 2.3.3 applied twice, there exists $C > 0$ such that every tournament G of order $n > 1$ containing no (a, H, K)-jewel-chain of length four satisfies $\alpha(G) \geq C^{-1}n/\phi(n)$. By 2.2.1, every such G satisfies $\chi(G) \leq C\phi(n)\log(n)$. This proves (2).

Let $c = \lambda_1 + \lambda_2 C$; we claim that c satisfies the theorem. For let G be a $\Delta(H, 1, K)$-free tournament, with $n > 1$ vertices. Let $c_1 = \phi(n)$. Then every H-free subtournament of G has chromatic number at most c_1; and so does every K-free subtournament of G, since every K-free tournament has at most 2^k vertices and hence has chromatic number at most $2^k \leq \phi(2) \leq \phi(n) = c_1$. Let $c_2 = C\phi(n)\log(n)$; then every subtournament of G not containing an (a, H, K)-jewel-chain of length four has chromatic number at most c_2, by (2). By 2.4.3,

$$\chi(G) \leq \lambda_1 c_1 + \lambda_2 c_2 = \lambda_1 \phi(n) + \lambda_2 C\phi(n)\log(n) \leq (\lambda_1 + \lambda_2 C)\phi(n)\log(n).$$

This proves 2.4.2, and hence 2.4.1, and therefore finishes the proof of 2.1.4. ∎

That completes all we have to say about ϵ-heroes in general.

2.5 Excluding $\Delta(2, k, l)$

Now we return to the case $\epsilon = 1$ and the proof of 2.1.3. So far we have proved the first two statements of 2.1.3, and part of the "if" half of the third statement, all as corollaries of 2.1.4. In this section we complete the proof of the "if" half of the third statement of 2.1.3, by proving the following.

2.5.1 *For all $k, l \geq 2$, there exists $c > 0$ such that every $\Delta(2, k, l)$-free tournament*

G with $|G| > 1$ satisfies $\alpha(G) \geq c|G|/\log(|G|)$.

This follows immediately from 2.5.3 and 2.5.4, proved below. We need the "bipartite Ramsey theorem", proved by Beineke and Schwenk [5], the following. If X, Y are disjoint subsets of the vertex set of a graph G, we say X is *complete to* Y if every vertex in X is adjacent to every vertex in Y, and X is *anticomplete to* Y if there are no edges between X and Y.

2.5.2 *For all integers $l \geq 0$ there exists $K \geq 0$, such that for every graph with bipartition (A, B) where $|A|, |B| \geq K$, there exist $X \subseteq A$ and $Y \subseteq B$ with $|X| = |Y| = l$, such that either X is complete to Y or X is anticomplete to Y.*

The smallest K satisfying the statement of 2.5.2 will be denoted by $K(l)$.

If G is a tournament and uv is an edge, we say that u is *adjacent to* v and v is *adjacent from* u. Let $(v_1, ..., v_n)$ be an enumeration of the vertex set of a tournament G (thus, with $n = |V(G)|$). We say that an edge $v_i v_j$ of G is a *backedge* under this enumeration if $i > j$. If $t \geq 0$ is an integer, an enumeration $(v_1, \dots, v_n)$ of $V(G)$ is said to be *t-forward* if for every two sets $X, Y \subseteq V(G)$ with $|X| = |Y| = t$, there exist $v_i \in X$ and $v_j \in Y$ such that either $i \geq j$, or $v_i v_j$ is an edge of G.

2.5.3 *For all integers $k \geq 2$, there exists $c > 0$ such that, if G is a $\Delta(2, k, k)$-free tournament with $|G| > 1$ that admits a 2^k-forward enumeration, then $\alpha(G) \geq c|G|/\log(|G|)$.*

Proof. Let $M = 2^k K(2^k)$ and $c = 1/(4M)$. We will show that c satisfies the theorem. For let G be a $\Delta(2, k, k)$-free tournament with $|G| > 1$, and let $(v_1, \ldots, v_n)$ be a 2^k-forward enumeration of $V(G)$. For $1 \leq i \leq n$, we define $\phi(v_i) = i$. A backedge vu of G is *left-active* if there is no set $A \subseteq V(G)$ such that:

- $|A| = K(2^k)$
- for each $a \in A$, $\phi(u) < \phi(a) < (\phi(u) + \phi(v))/2$
- each $a \in A$ is adjacent from u and from v.

Similarly, a backedge vu is *right-active* if there is no set $B \subseteq V(G)$ such that:

- $|B| = K(2^k)$
- for each $b \in B$, $(\phi(u) + \phi(v))/2 < \phi(b) < \phi(v)$
- each $b \in B$ is adjacent to u and to v.

(1) *Every backedge vu is either left-active or right-active.*

For suppose that vu is a backedge that is neither left-active nor right-active. Thus there exists sets A and B as above. Let J be the graph with bipartition (A, B), in which $a \in A$ and $b \in B$ are adjacent if ba is an edge (and hence a backedge) of G. By 2.5.2, there exist $X \subseteq A$ and $Y \subseteq B$ such that $|X| = |Y| = 2^k$, and X is either complete or anticomplete to Y in J. Since the enumeration is 2^k-forward, and $\phi(x) < (\phi(u)+\phi(v))/2 < \phi(y)$ for all $x \in X$ and $y \in Y$, it follows that there exists $x \in X$ and $y \in Y$ such that yx is not a backedge of G, and thus x, y are not adjacent in J; and consequently X is anticomplete to Y in J, and so every vertex in y is adjacent in G from every vertex in X. Since $|X| = |Y| = 2^k$, there are transitive subsets X' of X and Y' of Y, both of cardinality k (by a theorem of [**30**]). But then the subtournament of G with vertex set $X' \cup Y' \cup \{u, v\}$ is isomorphic to $\Delta(2, k, k)$, a contradiction. This proves (1).

For a backedge vu, we call $\phi(v) - \phi(u)$ its *length*.

(2) *There do not exist $M \log(n)$ left-active edges in G with the same tail v.*

Suppose there do exist such edges. Since their lengths are all between 1 and $n-1$, it follows that for some integer t with $0 \leq t \leq \log(n)$, there are M left-active edges all with tail v and all with length between 2^t and $2^{t+1}-1$. Let them be vu_i ($1 \leq i \leq M$), numbered such that $\phi(u_i) < \phi(u_j)$ for $1 \leq i < j \leq M$. For $1 \leq i < j \leq M$, since

$$\phi(v) - \phi(u_j) \geq 2^t > (\phi(v) - \phi(u_i))/2,$$

it follows that $\phi(u_i) < \phi(u_j) < (\phi(u_i) + \phi(v))/2$. Let $X = \{u_i : 1 \leq i \leq 2^k\}$, and $Y = \{u_i : 2^k < i \leq M\}$. For each $u_i \in X$, vu_i is left-active, and so u_i is adjacent in G to at most $(K(2^k) - 1)$ members of Y. Consequently there are at least $|Y| - |X|(K(2^k) - 1) \geq 2^k$ members of Y that are adjacent in G to each member of X, contradicting that the enumeration is 2^k-forward. This proves (2).

By (2) there are at most $Mn\log(n)$ left-active edges in G, and similarly at most $Mn\log(n)$ right-active. By (1), it follows that there are at most $2Mn\log(n) = (2c)^{-1}n\log(n)$ backedges. Let J be the graph with vertex set $V(G)$ in which u, v are adjacent for each backedge vu. Thus $|E(J)| \leq (2c)^{-1}n\log(n)$. By Turan's theorem [**17**], applied to J, we deduce that J has a stable set of cardinality at least $cn/\log(n)$, and so $\alpha(G) \geq cn/\log(n)$. This proves 2.5.3. ∎

2.5.4 *For all integers $k \geq 2$ there exists $c > 0$ such that every $\Delta(2, k, k)$-free tournament G has a subtournament with at least $c|G|$ vertices that admits a 2^k-forward enumeration.*

Proof. Let $b = 2k+1$, and $d = (12k-1)b$. Let $c > 0$ be the real number satisfying

$$\log(c) = -240b^2 2^{7bd}.$$

We will show that c satisfies the theorem.

Let G be a $\Delta(2,k,k)$-free tournament. Let us say a *chain* is a sequence $A_1,\ldots,A_m$ of subsets of $V(G)$ with the following properties:

- $A_1,\ldots,A_m$ are pairwise disjoint
- for $1 \le i \le m$, $|A_i| = bd$ and A_i is transitive
- for $1 \le i < j \le m$, each vertex in A_j is adjacent to at most d vertices in A_i, and each vertex in A_i is adjacent from at most d vertices in A_j.

(1) *We may assume that G admits a chain $A_1,\ldots,A_m$ with $m \ge 4$.*

For if $n < 2^{4bd}$ then the theorem holds, since $c < 2^{-4bd}$ and so any one-vertex subtournament of G satisfies the theorem (and if G is null then G itself satisfies the theorem). Thus we assume that $n \ge 2^{4bd}$, and so G contains a transitive set of cardinality $4bd$. But then there is a chain A_1, A_2, A_3, A_4. This proves (1).

Let $A_1,\ldots,A_m$ be a chain with m maximum. Define $A = A_1 \cup \cdots \cup A_m$. For $1 \le i < m$, let B_i be the set of all $v \in V(G) \setminus A$ such that there exists $Y \subseteq A_i$ and $Z \subseteq A_{i+1}$ with $|Y| = |Z| = k$ and $\{v\} \Rightarrow Y \Rightarrow Z \Rightarrow \{v\}$. Let $B = B_1 \cup \cdots \cup B_{m-1}$, and $C = V(G) \setminus (A \cup B)$.

(2) $|B| \le m(bd)^{2k}$.

For suppose not. Then $|B_i| > (bd)^{2k}$ for some i with $1 \leq i < m$. For each $v \in B_i$, choose $Y_v \subseteq A_i$ and $Z_v \subseteq Y_i$ such that $|Y_v| = |Z_v| = k$ and $\{v\} \Rightarrow Y_v \Rightarrow Z_v \Rightarrow \{v\}$. Since there are at most $(bd)^{2k}$ possibilities for the pair (Y_v, Z_v), there exist distinct u, v with $Y_u = Y_v$ and $Z_u = Z_v$. But then the subtournament of G with vertex set $\{u, v\} \cup Y_v \cup Z_v$ is isomorphic to $\Delta(2, k, k)$, a contradiction.

(3) *For each $v \in C$, there is no i with $1 \leq i < m$ such that v has at least k out-neighbours in A_i and at least $(d+1)k$ in-neighbours in A_{i+1}. Also, there is no i with $1 \leq i < m$ such that v has at least $(d+1)k$ out-neighbours in A_i and at least k in-neighbours in A_{i+1}. In particular, there is no i with $1 \leq i < m$ such that v has at least $bd/2$ out-neighbours in A_i and at least $bd/2$ in-neighbours in A_{i+1}.*

For the first claim, suppose that $Y \subseteq A_i$ and $Z \subseteq A_{i+1}$ with $|Y| = k$ and $|Z| \geq (d+1)k$, and v is adjacent to every vertex in Y and adjacent from every vertex in Z. Now each vertex in Y has at most d in-neighbours in Z, and so at most dk vertices in Z have an out-neighbour in Y. Consequently, there exists $Z' \subseteq Z$ with $|Z'| = k$, such that $Y \Rightarrow Z'$. But then Y, Z' show that $v \in B_i \subseteq B$, a contradiction. This proves the first claim, and the second follows from the symmetry. The third follows since $bd/2 \geq k$ and $bd/2 \geq (d+1)k$. This proves (3).

For $1 \leq i < m$ let C_i be the set of all vertices $v \in C$ such that v has at least $bd/2$ in-neighbours in A_i and at least $bd/2$ out-neighbours in A_{i+1}. (Note that bd is odd, so equality is not possible here.) Let C_0 be the set of all $v \in C$ with at least $bd/2$ out-neighbours in A_1, and let C_m be the set of all $v \in C$ with at least $bd/2$ in-neighbours in A_m. By (3), it follows that $C_0, C_1, \dots, C_m$ are pairwise disjoint and have union C.

(4) *Let $0 \leq i \leq m$ and let $v \in C_i$. Then for $1 \leq h < i$, v has at most $k-1$ out-neighbours in A_h; and for $i+1 < j \leq m$, v has at most $k-1$ in-neighbours in A_j.*

For v has at least $bd/2$ in-neighbours in A_i, and since $v \notin B$, it follows from (3) that v has at least $bd/2$ in-neighbours in each of $A_1, \dots, A_i$. In particular, v has at least $bd/2$ in-neighbours in A_{h+1}. By (3), v has at most $k-1$ out-neighbours in A_h. This proves the first assertion. The second follows by the symmetry. This proves (4).

For $2 \leq i \leq m$ let $L_i = A_1 \cup \cdots \cup A_{i-2}$, and for $0 \leq i \leq m-2$ let $R_i = A_{i+3} \cup \cdots \cup A_m$. Let L_0, L_1, R_{m-1}, R_m all be the null set.

(5) *Let $0 \leq i \leq m$, and let $u, v \in L_i$ be distinct. Then there is no transitive set $Z \subseteq C_i$ with $|Z| = k$ such that $Z \rightarrow \{u, v\}$, and consequently there are at most*

2^k vertices in C_i that are adjacent to both u and v. Similarly, for $0 \le i \le m$, if $u, v \in R_i$ then there is no transitive set $Z \subseteq C_i$ with $|Z| = k$ such that $\{u, v\} \Rightarrow Z$, and hence there are at most 2^k vertices in C_i that are adjacent from both u and v.

For let $0 \le i \le m$, and let $u, v \in L_i$ (thus $i \ge 3$), and suppose that there exists a transitive set $Z \subseteq C_i$ with $|Z| = k$ such that every vertex in Z is adjacent to both u, v. By (4), each member of Z has at most $k-1$ out-neighbours in A_{i-1}. Also, u, v each have at most at in-neighbours in A_{i-1}. Consequently there is a subset Y of A_{i-1} with $|Y| = k$ such that $\{u, v\} \Rightarrow Y \Rightarrow Z$, since $bd - (k-1)k - 2d \ge k$. But then the subtournament of G with vertex set $\{u, v\} \cup Y \cup Z$ is isomorphic to $\Delta(2, k, k)$, a contradiction. This proves the first assertion, and the second follows by symmetry. This proves (5).

(6) *For $0 \le i \le m$, and all $u \in L_i$ and $v \in R_i$, there are fewer than 2^{7bd} vertices in C_i that are adjacent to u and from v.*

For since $L_i, R_i \ne \emptyset$, it follows that $3 \le i \le m-3$. Suppose that there are at least 2^{7bd} vertices in C_i adjacent to u and from v; then they include a transitive set Y of cardinality $7bd$. Choose a chain $Y_1, \dots, Y_7$ of subsets of Y such that $Y_h \Rightarrow Y_j$ for all h, j with $1 \le h < j \le 7$. By (5), every vertex in $L_i \setminus \{u\}$ has at most $k-1 \le d$ in-neighbours in Y, and every vertex in $R_i \setminus \{v\}$ has at most d out-neighbours in Y.

Also, each vertex in Y has at most $k-1 \leq d$ out-neighbours in A_h for $1 \leq h \leq i-2$, and at most d in-neighbours in A_j for $i+2 \leq j \leq m$, by (4). Choose h, j with $u \in A_h$ and $v \in A_j$. Then

$$A_1, \ldots, A_{h-1}, A_{h+1}, \ldots, A_{i-2}, Y_1, Y_2, \ldots, Y_7, A_{i+3}, \ldots, A_{j-1}, A_{j+1}, \ldots, A_m$$

is a chain with $m+1$ terms, contrary to the maximality of m. This proves (6).

(7) *Let $0 \leq i \leq m$, and let $Z \subseteq C_i$ be transitive. Let p be an integer such that $|Z| \leq bdp$ and $2b(k-1)p < d$. Then there are fewer than $2bp$ vertices in L_i that are adjacent from at least d members of Z.*

For suppose that there exists $W \subseteq L_i$ with $|W| = 2bp$ such that each member of W is adjacent from at least d members of Z. Each member of W has at least d in-neighbours in Z, and yet every two distinct members of W have at most $k-1$ common in-neighbours in Z, by (5). Hence $|Z| \geq d|W| - (k-1)|W|^2/2$. Since $|Z| \leq bdp$ and $|W| = 2bp$, it follows that $2(k-1)bp \geq d$, a contradiction. Thus there is no such W. This proves (7).

(8) *For $0 \leq i \leq m$ and all $v \in R_i$, if $Y \subseteq C_i$ is transitive and $v \Rightarrow Y$ then $|Y| < 12b \cdot 2^{7bd}$.*

It follows that $i \leq m-3$. Choose a maximal subset Z of Y such that every vertex in L_i is adjacent from at most d members of Z. Suppose that $|Z| \geq 6bd$, and choose a chain $Z_1, \ldots, Z_6$ of subsets of Z such that $Z_h \Rightarrow Z_j$ for $1 \leq h < j \leq 6$. By (2), every vertex of R_i different from v is adjacent to at most $k-1 \leq at$ members of Y. Let $v \in A_j$. By (4), if $i \geq 2$ then

$$A_1, \ldots, A_{i-2}, Z_1, \ldots, Z_6, A_{i+3}, \ldots, A_{j-1}, A_{j+1}, \ldots, A_m$$

is a chain with $m+1$ terms, contrary to the maximality of m; while if $i \leq 1$ then the chain

$$Z_1, \ldots, Z_6, A_{i+3}, \ldots, A_{j-1}, A_{j+1}, \ldots, A_m$$

gives a contradiction similarly. Thus $|Z| < 6bd$.

We say $u \in L_i$ is *saturated* if u is adjacent from exactly d members of Z. Since $|Z| < 6bd$ and $12(k-1)b < d$, it follows from (7) with $p = 6$ that there are fewer than $12b$ saturated vertices in L_i. But every vertex in $Y \setminus Z$ is adjacent to a saturated vertex in L_i, from the maximality of Z. Since every saturated vertex in L_i is adjacent from at most 2^{7bd} members of Y, by (6), and hence from at most

$2^{7bd} - d$ members of $Y \setminus Z$, it follows that $|Y \setminus Z| \leq 12b(2^{7bd} - d)$, and so

$$|Y| \leq 12b(2^{7bd} - d) + 6bd < 12b \cdot 2^{7bd}.$$

This proves (8).

(9) *For $0 \leq i \leq m$, there is no transitive subset Y of C_i with $|Y| \geq 240b^2 2^{7bd}$.*

Let $Y \subseteq C_i$ be transitive. Choose a maximal subset Z of Y such that every vertex of L_i is adjacent from at most d members of Z, and every vertex in R_i is adjacent to at most d members of Z. Suppose that $|Z| \geq 5bd$, and choose a chain $Z_1, \ldots, Z_5$ of subsets of Z such that $Z_h \Rightarrow Z_j$ for $1 \leq h < j \leq 5$. If $2 \leq i \leq m-2$ then by (4),

$$A_1, \ldots, A_{i-2}, Z_1, \ldots, Z_5, A_{i+3}, \ldots, A_m$$

is a chain with $m+1$ terms, a contradiction; while if $i \leq 1$ then

$$Z_1, \ldots, Z_5, A_{i+3}, \ldots, A_m$$

gives a contradiction, and if $i \geq m-1$ then

$$A_1, \ldots, A_{i-2}, Z_1, \ldots, Z_5$$

gives a contradiction. Thus $|Z| < 5bd$.

We say $u \in L_i$ is *saturated* if it is adjacent from exactly d members of Z; and $v \in R_i$ is *saturated* if it is adjacent to exactly d members of Z. Since $|Z| \leq 5t$, and $10(k-1)b < d$, it follows from (7) with $p = 5$ that there are at most $10b$ saturated vertices in L_i, and similarly at most $10b$ saturated vertices in R_i. From the maximality of Z, every vertex of $Y \backslash Z$ is adjacent to at least one of the saturated vertices in L_i or from at least one of the saturated vertices in R_i. But by (8), each saturated vertex in L_i is adjacent from at most $12b2^{7bd}$ members of Y and hence from at most $12b2^{7bd} - d$ members of $Y \setminus Z$, and similarly every saturated vertex in R_i is adjacent to at most $12b2^{7bd} - d$ members of $Y \setminus Z$. We deduce that

$$|Y| < 20b(12b2^{7bd} - d) + 5bd \leq 240b^2 2^{7bd}.$$

This proves (9).

(10) $|A| \geq 2c|G|$ *where* c *is as defined in the statement of the theorem.*

From (9), each C_i has cardinality at most $2^{240b^2 2^{7bd}-1}$, and so $|C| \leq (m+1)2^{240b^2 2^{7bd}-1}$. Since $m \geq 2$ (and hence $m+1 \leq 2m$), and $|B| \leq m(bd)^{2k}$ by (2), and $|A| = mbd$, we deduce that

$$|G| \leq (2^{240b^2 2^{7bd}} + (bd)^{2k} + bd)m \leq (2^{240b^2 2^{7bd}} + (bd)^{2k} + bd)|A|/(bd).$$

It follows that $|A| \geq 2c|G|$ where c is as defined in the statement of the theorem. This proves (10).

Let V be the union of all A_i with $1 \leq i \leq m$ and i odd. Then $|V| \geq |A|/2 \geq c|G|$. Number the members of V as $\{v_1, \ldots, v_t\}$ say, where for $1 \leq r < s \leq t$, if $x_r \in A_i$ and $x_s \in A_j$ then $i \leq j$, and either $i < j$ or x_r is adjacent to x_s. (This is possible since each A_i is transitive.) We claim that this order is 2^k-forward. For let Y, Z be disjoint subsets of V with $|Y| = |Z| = 2^k$, such that for $1 \leq r, s \leq t$, if $x_r \in Y$ and $x_s \in Z$ then $r < s$. We must show that there exists $y \in Y$ and $z \in Z$ such that y is adjacent to z. Suppose not. Choose i with $1 \leq i \leq m$ and i odd, maximum such that $A_i \cap Y \neq \emptyset$. It follows that $A_h \cap Z = \emptyset$ for all $h < i$. If $Z \cap A_i \neq \emptyset$, let $v_r \in A_i \cap Y$ and $v_s \in A_i \cap Z$; it follows that $r < s$ from the choice of the numbering, and so v_r is adjacent to v_s, a contradiction. Thus $Z \cap A_i = \emptyset$. It follows that $j \geq i + 2$ for each j with $1 \leq j \leq m$ such that $z \cap A_j \neq \emptyset$. Since $|Y| = 2^k$, there exists $Y' \subseteq Y$ with $|Y| = k$ such that Y is transitive, and similarly there exists a transitive $Z' \subseteq Z$ with $|Z'| = k$. Now each member of Y' is adjacent from at most d members of A_{i+1}, and so there are at most dk vertices in A_{i+1} adjacent to some member of Y'; and similarly at most dk are adjacent from some member of Z'. Since $bd \geq 2dk + 2$, there are two vertices $u, v \in A_{i+1}$ such that $Y' \Rightarrow \{u, v\}$ and $\{u, v\} \Rightarrow Z'$. But then the subtournament of G with vertex set $\{u, v\} \cup Y' \cup Z'$ is isomorphic to $\Delta(2, k, k)$, a contradiction. This proves that the order is 2^k-forward, and so completes the proof of 2.5.4. ∎

Proof of 2.5.1. This follows immediately from 2.5.3 and 2.5.4. ∎

2.6 Strongly-connected pseudo-heroes

In this section we complete the proof of 2.1.3, and also prove 2.1.5. As a biproduct of the remainder of the proof of 2.1.3, we are able to identify all the minimal tournaments that are not pseudo-heroes (there are six). Here they are:

- Let H_1 be the tournament with five vertices $v_1, \dots, v_5$, in which v_i is adjacent to v_{i+1} and v_{i+2} for $1 \leq i \leq 5$, reading subscripts modulo 5 (the tournament C_5).
- Let H_2 be the tournament obtained from H_1 by replacing the edge v_5v_1 by an edge v_1v_5.
- Let H_3 be the tournament with five vertices $v_1, \dots, v_5$ in which v_i is adjacent to v_j for all i, j with $1 \leq i < j \leq 4$, and v_5 is adjacent to v_1, v_3 and adjacent

from v_2, v_4.

- Let H_4 be the tournament $\Delta(1, \Delta(1,1,1), \Delta(1,1,1))$
- Let H_5 be the tournament $\Delta(2, 2, \Delta(1,1,1))$
- Let H_6 be the tournament $\Delta(3,3,3)$.

First, we prove they are not pseudo-heroes, but also it is helpful to give the best upper bounds on their ξ-values that we can. We begin with:

2.6.1 *If H is a strongly-connected tournament with more than one vertex that does not admit a trisection, then $\xi(H) \leq 1/\log(3)$. In particular, $\xi(H_i) \leq 1/\log(3)$ for $i = 1, 2, 3$, and so H_1, H_2, H_3 are not pseudo-heroes.*

Proof. Let D_0 be the one-vertex tournament, and for $i \geq 1$ let $D_i = \Delta(D_{i-1}, D_{i-1}, D_{i-1})$. Thus $|D_i| = 3^i$. For $i > 0$, no transitive subtournament of D_i intersects all three parts of the trisection of D_i, so $\alpha(D_i) = 2\alpha(D_{i-1})$; and consequently $\alpha(D_i) = 2^i = |D_i|^{1/\log(3)}$. We claim that for all $i \geq 0$, D_i does not contain H; for suppose D_i contains H for some value of i, and choose the smallest. Then $i \geq 1$ since $|V(H)| \geq 2$, and so D_i admits a trisection (A, B, C) where $D_i|A, D_i|B, D_i|C$ are all isomorphic to D_{i-1}. Choose a subtournament T of D_i isomorphic to H. From the minimality of i, $V(T)$ is not a subset of any of A, B, C, and therefore has nonempty intersection with at least two of them; and since H is strongly-connected, $V(T)$ has nonempty intersection with all three of A, B, C. But then T admits a trisection, a contradiction.

This proves that no D_i contains H. Let ϵ be an EH-coefficient for H, and choose $c > 0$ such that every H-free tournament G satisfies $\alpha(G) \geq c|G|^{\epsilon}$. In particular, taking $G = D_i$ implies that

$$|D_i|^{1/\log(3)} = \alpha(D_i) \geq c|D_i|^{\epsilon},$$

for all $i \geq 0$. It follows that $1/\log(3) \geq \epsilon$. Since this holds for all EH-coefficients ϵ, it follows that $\xi(H) \leq 1/\log(3)$. This proves 2.6.1. ∎

2.6.2 *$\xi(H_4) \leq 1/2$, and hence H_4 is not a pseudo-hero.*

Proof. For $k \geq 1$, let D_k be the tournament with k^2 vertices $v_1, \dots, v_{k^2}$, in which for $1 \leq i < j \leq k^2$, v_i is adjacent to v_j if k does not divide $j - i$, and otherwise v_j is adjacent to v_i. (This construction is due to Gaku Liu, in private communication.) For $1 \leq i \leq k$, let $C_i = \{v_i, v_{i+k}, v_{i+2k}, \dots, v_{i+(k-1)k}\}$. Then $C_1, \dots, C_k$ are disjoint and have union $V(D_k)$.

(1) $\alpha(D_k) \leq 2k - 1$.

Let $X \subseteq V(D_k)$ induce a transitive tournament. For $1 \leq i \leq k$, if $X \cap C_i \neq \emptyset$, let

p_i be the smallest value of j such that $v_j \in X \cap C_i$, and q_i the largest; and let I_i be $\{v_j : p_i \le j \le q_i\}$. If $X \cap C_i = \emptyset$, let $I_i = \emptyset$. Note that if $v_j \in X \cap I_i$ then $j \in C_i$; because otherwise $\{v_{p_i}, v_{q_i}, v_j\}$ would induce a cyclic triangle, contradicting that X is transitive. This has two consequences:

- For each $i \in \{1, \ldots, k\}$, $|X \cap I_i| \le 1 + (|I_i| - 1)/k$, since between any two members of X in I_i there are $k-1$ members of $C_i \setminus X$. Summing over i, we deduce that $|X| \le k - 1 + \sum_i |I_i|/k$.
- The sets I_i $(1 \le i \le k)$ are pairwise disjoint, and so $\sum_i |I_i| \le k^2$.

Combining these, we deduce that $|X| \le 2k - 1$. This proves (1).

(2) *D_k does not contain H_4.*

For $1 \le j \le k^2$, let $\phi(v_j)$ be the value of $i \in \{1, \ldots, k\}$ with $v_j \in C_i$. Thus, let $a, b, c \in V(D_k)$ be distinct:

(P) if $\{a, b, c\}$ induces a cyclic triangle in D_k then $|\{\phi(a), \phi(b), \phi(c)\}| = 2$; and

(Q) if ab, ac, bc are edges and $\phi(a) = \phi(c)$ then $\phi(b) = \phi(a)$.

(R) if $\{a, b, c\}$ induces a cyclic triangle and d is some other vertex such that $d \Rightarrow \{a, b, c\}$ or $d \Leftarrow \{a, b, c\}$ then $\phi(d) \ne \phi(a), \phi(b), \phi(c)$.

(The third condition above follows easily from the other two, but we use it enough to give it a separate name.) For $X \subseteq V(D_k)$, $\phi(X)$ denotes $\{\phi(v)\, v \in X)$. Suppose that D_k contains H_4, and let A, B, C be the trisection of H_4 with $|A| = |B| = 3$; let $A = \{a_1, a_2, a_3\}$, and $B = \{b_1, b_2, b_3\}$, and $C = \{c\}$. Thus from property P applied to A, $|\phi(A)| = 2$, and similarly $|\phi(B)| = 2$; by property R applied to A and each member of B, $\phi(A)$ and $\phi(B)$ are disjoint; and by property R applied to A and c, $\phi(c) \notin \phi(A)$ and similarly $\phi(c) \notin \phi(B)$. Choose $a \in A$ and $b \in B$; then $\phi(a), \phi(b), \phi(c)$ are all distinct, contrary to property P. This proves (2).

Let ϵ be an EH-coefficient for H_4, and choose $c > 0$ such that every H_4-free tournament G satisfies $\alpha(G) \geq c|G|^{\epsilon}$. In particular, for each $k \geq 1$, $\alpha(D_k) \geq c|D_k|^{\epsilon}$, and so from (1), $2k - 1 \geq ck^{2\epsilon}$. Since this holds for all $k \geq 1$, we deduce that $\epsilon \leq 1/2$, and so $\xi(H_4) \leq 1/2$. This proves 2.6.2. ∎

The above is not the easiest way to prove that H_4 is not a pseudo-hero, but it gives the best bound on $\xi(H_4)$.

Next we need a lemma proved in [**11**] (it will be also given in Chapter 6), the following:

2.6.3 *The vertex set of every tournament H can be ordered such that the set of backward edges of every non-null subtournament S of H has cardinality at most* $(|S| - 1)(\xi(H))^{-1}$.

We deduce

2.6.4 *$\xi(H_5) \leq 5/6$, and so H_5 is not a pseudo-hero.*

Proof. Let $H = H_5$, and let $V(H) = A \cup B \cup C$, where

- $A = \{a_1, a_2\}, B = \{b_1, b_2\}$, and $C = \{c_1, c_2, c_3\}$
- $A \Rightarrow B \Rightarrow C \Rightarrow A$
- c_1-c_2-c_3-c_1 is a directed cycle.

Suppose there is an ordering of $V(H)$ such that no cycle of the backedge graph has length at most six; let X be the set of backedges in this ordering, and let $Y = E(H) \setminus X$. We have two properties:

(P) For every directed cycle of H, at least one of its edges in in X.

(Q) For every undirected cycle of H of length at most six, at least one of its edges is in Y.

Since every undirected graph with seven vertices and eight edges has a cycle of length at most six (indeed, at most five), it follows that $|X| \leq 7$. Suppose first that $a_1b_1, a_2b_2 \in Y$. From property P applied to the directed cycle c_i-a_j-b_j-c_i, at least one of c_ia_j, b_jc_i is in X, for $i = 1, 2, 3$ and $j = 1, 2$. Thus there are at least six edges in X between $A \cup B$ and C. By property P applied to $H|C$, some edge of X has both ends in C. Since $|X| \leq 7$, it follows that all edges from A to B belong to Y; and so by property P, for $i = 1, 2, 3$ either $c_i a_1, c_i a_2 \in X$, or $b_1 c_i, b_2 c_i \in X$. Thus from the symmetry we may assume that $c_1 a_1, c_1 a_2, c_2 a_1, c_2 a_2 \in X$. But these four edges form a cycle contrary to property Q.

Thus not both $a_1 b_1, a_2 b_2 \in Y$, and similarly not both $a_1 b_2, a_2 b_1 \in Y$. Suppose next that $a_1 b_1, a_1 b_2 \in Y$. Thus $a_2 b_1, a_2 b_2 \in X$. By property Q applied to the cycle a_2-b_1-c_i-b_2-a_2, for $i = 1, 2, 3$ not both $b_1 c_i, b_2 c_i \in X$. By property P applied to the directed cycles c_i-a_1-b_1-c_i and c_i-a_1-b_2-c_i it follows that $c_i a_1 \in X$, for $i = 1, 2, 3$. But some edge of X has both ends in C, contrary to property Q.

It follows that not both $a_1 b_1, a_2 b_2 \in Y$, and so from the symmetry, at most one edge from A to B belongs to Y. By property Q, not all four of these edges are in X, so we may assume that $a_1 b_1 \in Y$, and $a_2 b_1, a_1 b_2, a_2 b_2 \in X$. From property P, some edge of $H|C$ belongs to X, say $c_1 c_2$. Now by property P again, for $i = 1, 2$ at least one of $c_i a_1, b_1 c_i \in X$. But then there are six edges in X each with both ends in $V(H) \setminus \{c_3\}$, contrary to property Q.

It follows that in every ordering of $V(H)$, some cycle of the backedge graph has length at most six. From 2.6.3, we deduce that $\xi(H) \leq 5/6$. This proves the first assertion of the theorem, and the second follows. ∎

Finally:

2.6.5 $\xi(H_6) \leq 3/4$, *and so* H_6 *is not a pseudo-hero.*

Proof. Let $H = H_6$, and let $V(H) = A \cup B \cup C$, where

- $A = \{a_1, a_2, a_3\}, B = \{b_1, b_2, b_3\}$, and $C = \{c_1, c_2, c_3\}$
- $A \Rightarrow B \Rightarrow C \Rightarrow A$
- A, B, C are all transitive.

Suppose there is an ordering of $V(H)$ such that no cycle of the backedge graph has length at most four; let X be the set of backedges in this ordering, and let $Y = E(H) \setminus X$. We have two properties:

(P) For every directed cycle of H, at least one of its edges in in X.

(Q) For every undirected cycle of H of length at most four, at least one of its edges is in Y.

If there is a three-edge matching of members of Y between A, B, and also between B, C and between C, A, then the union of these three matchings uncludes a directed cycle of H, contrary to property P. So we may assume there is no three-edge matching of members of Y between A and B. By Hall's theorem, there are two vertices $x, y \in A \cup B$ such that every edge in Y between A and b is incident with one of x, y. If $x \in A$ and $y \in B$, and $x = a_3, y = b_3$ say, then $a_1b_1, a_1b_2, a_2b_1, a_2b_2$ are all in

X, contrary to property Q. Thus we may assume that $x, y \in A$; say $x = a_1, y = a_2$. Hence $a_3b_1, a_3b_2, a_3b_3 \in X$. Let $1 \leq k \leq 3$. We claim that $c_ka_1, c_ka_2 \in X$. For suppose that $c_ka_1 \in Y$ say. From property Q at most one of the edges a_1b_1, a_1b_2, a_1b_3 is in X (otherwise there is a cycle of edges in X of length four passing through a_3); say $a_1b_1, a_1b_2 \in Y$. Now from property P applied to a_1-b_j-c_k-a_1, it follows that $b_jc_k \in X$ for $j = 1, 2$, contrary to property Q. This proves that $c_ka_1, c_ka_2 \in X$, for $k = 1, 2, 3$; but again this contradicts property Q. This proves 2.6.5. ∎

Now we complete the proof of 2.1.3; all that remains is to prove the "only if" half of the third statement of 2.1.3, which is the equivalence of the first two statements of the following.

2.6.6 *Let H be a strongly-connected tournament with more than one vertex. Then the following are equivalent:*

- *H is a pseudo-hero*
- *every strong component of H is isomorphic to $\Delta(2, k, l)$ for some $k, l \geq 2$, or to $\Delta(1, P, T)$ or $\Delta(1, T, P)$ for some pseudo-hero P and some nonempty transitive tournament T*
- *H contains none of $H_1, \ldots, H_6$.*

Proof. The first statement implies the third, by 2.6.1, 2.6.2, 2.6.4 and 2.6.5, since every subtournament of a pseudo-hero is a pseudo-hero. By 2.5.1 and 2.4.1 with $\epsilon = 1$, and 2.3.1 with $\epsilon = 1$, the second statement implies the first. It remains to show that the third implies the second, and we proceed by induction on $|V(H)|$. Thus, let H contain none of $H_1, \ldots, H_6$. If H is not strongly-connected, then inductively we may assume that all its strong components are pseudo-heroes, and hence so is H, by 2.3.1 with $\epsilon = 1$. If H is strongly-connected, then by a theorem of Gaku Liu, proved in [7], since H contains none of H_1, H_2, H_3, it admits a trisection (A, B, C). We may assume that $|C| \leq |A|, |B|$. If $|C| = 1$ then since H does not contain H_4, it follows that at least one of A, B is transitive, and so $H = \Delta(1, P, T)$ or $H = \Delta(1, T, P)$ for some pseudo-hero P and some nonempty transitive tournament T, and the theorem holds. If $|C| \geq 2$, then since H does not contain H_5 and $|A|, |B| \geq 2$ it follows that A, B, C are all transitive, and therefore $|C| = 2$ since H does not contain H_6; but then $H = \Delta(2, k, l)$ for some $k, l \geq 2$, and the theorem holds. This proves 2.6.6, and hence completes the proof of 2.1.3. ∎

Proof of 2.1.5. If H is not a pseudo-hero then from 2.6.6, H contains one of $H_1, \ldots, H_6$, and so $\xi(H) \leq \max(\xi(H_1), \ldots, \xi(H_6))$. But by 2.6.1, 2.6.2, 2.6.4 and 2.6.5, this maximum is at most $5/6$. This proves 2.1.5. ∎

3

Tools

3.1 The regularity lemma

In this chapter we prove several technical lemmas and give some definitions that will turn out to be very useful in next two chapters. We start by introducing regularity lemma.

The regularity lemma is one of the most fundamental mathematical tools in modern graph theory. We will use its directed version to prove some of the results concerning constellations and tournaments on at most 5 vertices.

Let $X, Y \subseteq V(T)$ be disjoint. Denote by $e_{X,Y}$ the number of directed edges (x, y), where $x \in X$ and $y \in Y$. The *directed density from X to Y* is defined as $d(X, Y) =$

$\frac{e_{X,Y}}{|X||Y|}$.

Given $\epsilon > 0$ we call a pair A, B of disjoint subsets of $V(T)$ *ϵ-regular* if all $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \epsilon|A|$ and $|Y| \geq \epsilon|B|$ satisfy: $|d(X,Y) - d(A,B)| \leq \epsilon$, $|d(Y,X) - d(B,A)| \leq \epsilon$.

Consider a partition $\{V_0, V_1, ..., V_k\}$ of $V(T)$ in which one set V_0 has been singled out as an *exceptional* set. (This exceptional set V_0 may be empty). We call such a partition an ϵ-regular partition of T if it satisfies the following three conditions:

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

The following directed version of the *Regularity Lemma* has been proven in [3]:

3.1.1 *For every $\epsilon > 0$ and every $m \geq 1$ there exists an integer $DM = DM(m, \epsilon)$ such that every tournament of size at least m admits an ϵ-regular partition $\{V_0, V_1, ..., V_k\}$ with $m \leq k \leq DM$.*

We also need the following lemma that will be used together with the regularity lemma:

3.1.2 *For every natural number k and real number $0 < \lambda < 1$ there exists $0 < \eta = \eta(k, \lambda) < 1$ such that for every tournament H with vertex set $\{x_1, ..., x_k\}$ and tournament T with vertex set $V(T) = \bigcup_{i=1}^{k} V_i$, if the V_i's are disjoint sets, each of size at least one, and each pair (V_i, V_j), $1 \leq i < j \leq k$ is η-regular, with $d(V_i, V_j) \geq \lambda$ and $d(V_j, V_i) \geq \lambda$, then there exist vertices $v_i \in V_i$ for $i \in \{1, \ldots, k\}$, such that the map $x_i \rightarrow v_i$ gives an isomorphism between H and the subtournament of T induced by $\{v_1, ..., v_k\}$.*

The proof of the undirected version of Lemma 3.1.2 can be found in [**8**]. We omit the proof of 3.1.2 since it is completely analogous to the undirected version.

3.2 ϵ-critical tournaments

We will prove here several properties of so-called *ϵ-critical* tournaments that play an important role in all our proofs establishing lower bounds on EH-suprema for several families of tournaments.

We denote by $tr(T)$ the size of the largest transitive subtournament of T. We call a tournament T *ϵ-critical* for $\epsilon > 0$ if $tr(T) < |T|^{\epsilon}$ but for every proper subtournament S of T we have: $tr(S) \geq |S|^{\epsilon}$. Below we list some properties of ϵ-critical tournaments.

3.2.1 *For every $N > 0$ there exists $\epsilon(N) > 0$ such that for every $0 < \epsilon < \epsilon(N)$*

every ϵ-critical tournament T satisfies $|T| \geq N$.

Proof. Since every tournament contains a transitive subtournament of size 2, it suffices to take $\epsilon(N) = \log_N(2)$. ∎

3.2.2 *Let T be an ϵ-critical tournament with $|T| = n$ and $\epsilon, c, f > 0$ be constants such that $\epsilon < \log_c(1 - f)$. Then for every $A \subseteq V(T)$ with $|A| \geq cn$ and every transitive subtournament G of T with $|G| \geq ftr(T)$ A is not complete from $V(G)$ and A is not complete to $V(G)$.*

Proof. Assume otherwise. Let A_T be a transitive subtournament in $T|A$ of size $tr(A)$. Then $|A_T| \geq (cn)^{\epsilon}$. Now we can merge A_T with G to obtain a transitive subtournament of size at least $(cn)^{\epsilon} + ftr(T)$. From the definition of $tr(T)$ we have $(cn)^{\epsilon} + ftr(T) \leq tr(T)$. So $c^{\epsilon}n^{\epsilon} \leq (1-f)tr(T)$, and in particular $c^{\epsilon}n^{\epsilon} < (1-f)n^{\epsilon}$. But this contradicts the fact that $\epsilon < \log_c(1 - f)$. ∎

The next lemma is a starting point for all of our constructions. This is also the only step in the proof where we use 3.1.1. Note that in what follows we do not require for the pairs (A_i, A_j) to be regular, and so even we do not need the full strength of 3.1.1.

3.2.3 *Let H be a tournament, $P > 0$ be an integer and $0 < \lambda < \frac{1}{2}$. Then there is an integer N such that for every H-free tournament T with $|T| \geq N$ there exists* *a constant $c > 0$ and P pairwise disjoint subsets $A_1, A_2, ..., A_P$ of vertices of T satisfying:*

- $d(A_i, A_j) \geq 1 - \lambda$ *for* $i, j \in \{1, 2, ..., P\}$, $i < j$
- $|A_i| \geq c|T|$ *for* $i \in \{1, 2, ..., P\}$.

Proof. Write $|T| = n$, $|H| = h$. Let $R(t_1, t_2)$ be the smallest integer such that every graph of size at least $R(t_1, t_2)$ contains either a stable set of size t_1 or a clique of size t_2 (so $R(t_1, t_2)$ is simply a *Ramsey number*, see [**17**]). Write $k = R(2^{P-1}, h)$. Write $\eta = \min\left\{\frac{1}{2(k-1)}, \eta_0(h, \lambda)\right\}$ (where η_0 is as in the statement of 3.1.2). Let $u > 0$ be the smallest integer such that: $\binom{\hat{u}}{2} - \eta\hat{u}^2 > \frac{1}{2}\frac{k-2}{k-1}\hat{u}^2$ holds for all $\hat{u} \geq u$. By 3.1.1 there exists an integer $N > 0$ such that every tournament T with $|T| \geq N$ admits an η-regular partition of at least u parts. Denote by DM the upper bound (from 3.1.1) on the number of parts of this partition. Denote the parts of the partition by: $W_0, W_1, ..., W_r$, where $u \leq r \leq DM$ and W_0 is the exceptional set. We have: $|W_i| \geq \frac{(1-\eta)n}{DM}$. Now consider the graph G with $V(G) = \{W_1, \ldots, W_r\}$, where there is an edge between two vertices if the pair (W_i, W_j) is η-regular. Then, from the definition of u, we have: $|E(G)| \geq \frac{k-2}{2(k-1)}|V(G)|^2$. So by Turan's theorem ([**17**]) it follows that G has a clique of size at least k. That means that there exist k parts of the partition, without loss of generality $W_1, ...W_k$, such that for all $i, j \in \{1, 2, ..., k\}, i \neq j$ the pair (W_i, W_j) is η-regular. We say that a pair (W_i, W_j) for $i, j \in \{1, 2, ..., k\}, i \neq j$ is *good* if $\lambda \leq d(W_i, W_j) \leq 1 - \lambda$. Otherwise we say this pair is *bad*. Now consider the graph $\hat{G}$ with $V(\hat{G}) = \{W_1, ...W_k\}$, where there is an edge between W_i and W_j for $i, j \in \{1, ..., k\}$, $i \neq j$ if (W_i, W_j) is a good pair. From the definition of k we know that $\hat{G}$ contains a clique of size h or a stable set of size 2^{P-1}. In other words, either

- there exist h parts of the partition, without loss of generality denote them by $W_1, ...W_h$, such that every pair (W_i, W_j) is η-regular and $\lambda \leq d(W_i, W_j) \leq 1 - \lambda$ for $i, j \in \{1, 2, ..., k\}, i \neq j$, or
- there exist 2^{P-1} parts of the partition, without loss of generality denote them by $W_1, ...W_{2^{P-1}}$, such that every pair (W_i, W_j) is η-regular and $d(W_i, W_j) > 1 - \lambda$ or $d(W_j, W_i) > 1 - \lambda$ for $i, j \in \{1, 2, ..., 2^{P-1}\}, i \neq j$.

Since T is H-free and $\eta \leq \eta_0$, 3.1.2 implies that the former is impossible.

Now define $\hat{T}$ to be a tournament with $V(\hat{T}) = \{W_1, ..., W_{2^{P-1}}\}$, where an edge is directed from W_i to W_j if $d(W_i, W_j) > 1-\lambda$ and from W_j to W_i otherwise. Using the fact that every tournament of size at least 2^{P-1} contains a transitive subtournament of size at least P ([30]), we conclude that $\hat{T}$ contains a transitive subtournament of size P. That means that there exist P parts of the partition, without loss of generality $W_1, ..., W_P$, such that $d(W_i, W_j) \geq 1 - \lambda$ for $i, j \in \{1, 2, ..., P\}, i < j$. Note that each W_i is of size at least $\frac{(1-\eta)n}{DM}$, so taking $A_i = W_i$ for $i = 1, 2, ...P$ and $c = \frac{1-\eta}{DM}$ completes the proof. ∎

The following is an easy but useful fact.

3.2.4 *Let A_1, A_2 be two disjoint sets such that $d(A_1, A_2) \geq 1 - \lambda$ and let $0 \leq \eta_1, \eta_2 < 1$. Let $\hat{\lambda} = \frac{\lambda}{\eta_1 \eta_2}$. Let $X \subseteq A_1$, $Y \subseteq A_2$ be such that $|X| \geq \eta_1 |A_1|$ and $|Y| \geq \eta_2 |A_2|$. Then $d(X, Y) \geq 1 - \hat{\lambda}$.*

Proof. Denote by B the set of edges directed from A_2 to A_1. We have $|B| \leq \lambda |A_1||A_2|$. On the other hand $|B| \geq (1 - d(X, Y))|X||Y|$. Therefore $d(X, Y) \geq 1 - \lambda \frac{|A_1|}{|X|}\frac{|A_2|}{|Y|}$ and the result follows. ∎

Next we refine 3.2.3 further.

3.2.5 *Let $0 < \lambda < 1$, $c > 0$, $0 < \epsilon < \log_{\frac{c}{2}}(\frac{1}{2})$ be constants and P be a positive integer. Let T be an ϵ-critical tournament with $|T| = n$. Assume that $A_1, A_2, ..., A_P \subseteq V(T)$ are pairwise disjoint sets of vertices such that $d(A_i, A_j) \geq (1 - \lambda)$ for $i, j \in \{1, 2, ..., n\}$, $i < j$ and $|A_i| \geq cn$ for $i \in \{1, 2, ..., P\}$. Let v be a $\{0, 1\}$-vector of length P. Define $I = \{i : v_i = 1\}$. Write $I = \{i_1, i_2, ...i_r\}$, where $i_1 < i_2 < ... < i_r$. Let $\Lambda = (4P)^{|I|}\lambda$. Then there exist transitive tournaments $T_*^{i_1}$, $T_*^{i_2}$,...$T_*^{i_r}$ such that $V(T_*^{i_s}) \subseteq A_{i_s}$, $|V(T_*^{i_s})| \geq \frac{1}{2}tr(T)$ for $s \in \{1, 2, ..., r\}$ and for every $T_*^{i_s}$ we have*

- *if $i < i_s$ and $i \notin I$ then $d(A_i, T_*^{i_s}) \geq 1 - \Lambda$*
- *if $i > i_s$ and $i \notin I$ then $d(A_i, T_*^{i_s}) \leq \Lambda$*

- *if* $i, j \in I$ *and* $i < j$ *then* $d(T_*^i, T_*^j) \geq 1 - \Lambda$

Proof. The proof is by induction on $|I|$. For $|I| = 0$ the statement is obvious. Write $\hat{I} = \{i_1, ..., i_{r-1}\}$. Inductively, we may assume the existence of the sets $T_*^{i_1}, T_*^{i_2}, ..., T_*^{i_{r-1}}$ as in the statement of the lemma. Since T is ϵ-critical, we deduce that $tr(A_{i_r}) \geq |A_{i_r}|^\epsilon \geq (\frac{c}{2})^\epsilon n^\epsilon$, and therefore A_{i_r} contains a transitive subtournament of size $\lceil \frac{1}{2} n^\epsilon \rceil$. Denote this transitive tournament by $T_1^{i_r}$. We have $|T_1^{i_r}| \geq \frac{1}{2} tr(T)$. Similarly there exist an integer w and a family of pairwise disjoint transitive subtournaments: $\mathcal{W} = \{T_1^{i_r}, T_2^{i_r}, ..., T_w^{i_r}\}$ such that $\bigcup_{j=1}^w |T_j^{i_r}| \geq \frac{|A_{i_r}|}{2}$ and for every $j \in \{1, 2, ..., w\}$ $|T_j^{i_r}| \geq \frac{1}{2} tr(T)$. Denote by T^{i_r} a tournament induced by $\bigcup_{j=1}^w V(T_j^{i_r})$. We have $|T^{i_r}| \geq \frac{|A_{i_r}|}{2}$.

Write $\hat{\Lambda} = (4P)^{|I|-1}\lambda$. For $i < i^r$ and $i \notin \hat{I}$ denote by R_i the subset of $\mathcal{W}$ that consists of tournaments $T_x^{i_r}$ for which $d(A_i, T_x^{i^r}) < (1 - 4P\lambda)$. For $i > i^r$ and $i \notin \hat{I}$ denote by R_i the subset of $\mathcal{W}$ that consists of tournaments $T_x^{i_r}$ for which $d(T_x^{i^r}, A_i) < (1 - 4P\lambda)$. For $i < i^r$ and $i \in \hat{I}$ denote by R_i the subset of $\mathcal{W}$ that consists of tournaments $T_x^{i_r}$ for which $d(T_*^i, T_x^{i^r}) < (1 - 4P\hat{\Lambda})$. Finally, for $i > i^r$ and $i \in \hat{I}$ denote by R_i the subset of $\mathcal{W}$ that consists of tournaments $T_x^{i_r}$ for which $d(T_x^{i^r}, T_*^i) < (1 - 4P\hat{\Lambda})$. Since $d(A_i, A_{i^r}) \geq (1 - \lambda)$, by 3.2.4 we have $|R_i| \leq \frac{1}{2P} w$ for all $i \notin \hat{I}$ such that $i \neq i^r$. Similarly, from the induction hypothesis and 3.2.4 we have $|R_i| \leq \frac{1}{2P} w$ for all $i \in \hat{I}$. Write: $\mathcal{R} = \bigcup_{i \neq i_r} R_i$. Note that $\mathcal{R} \subseteq \mathcal{W}$ and $|\mathcal{R}| \leq \frac{1}{2P} w \cdot (P - 1) < w$. Therefore there exists a tournament $T_*^{i_r} \in \mathcal{W} \backslash \mathcal{R}$, and from the definition of $\mathcal{R}$, the following holds for every $i_s \in I$

- if $i < i_s$ and $i \notin I$ then $d(A_i, T_*^{i_s}) \geq 1 - 4P\hat{\Lambda}$
- if $i > i_s$ and $i \notin I$ then $d(A_i, T_*^{i_s}) \leq 4P\hat{\Lambda}$
- if $i < i_s$ and $i \in I$ then $d(T_*^i, T_*^{i_s}) \geq 1 - 4P\hat{\Lambda}$
- if $i > i_s$ and $i \in I$ then $d(T_*^i, T_*^{i_s}) \leq 4P\hat{\Lambda}$.

That completes induction since $4P\hat{\Lambda} = (4P)^{|I|}\lambda = \Lambda$. ∎

We need one more definition. Let $c > 0$, $0 < \lambda < 1$ be constants, and let w be a $\{0,1\}$-vector of length $|w|$. Let T be a tournament with $|T| = n$. A sequence of disjoint subsets $(S_1, S_2, ..., S_{|w|})$ of $V(T)$ is a (c, λ, w)-*structure* if

- whenever $w_i = 0$ we have $|S_i| \geq cn$
- whenever $w_i = 1$ the set $T|S_i$ is transitive and $|S_i| \geq ctr(T)$
- $d(S_i, S_j) \geq 1 - \lambda$ for all $1 \leq i < j \leq |w|$.

We say that (c, λ, w)-*structure* is *strong* if in addition it satisfies the following condition:

- if $w_i = 1$ and $w_j = 1$ for $1 \leq i < j \leq |w|$ then S_i is complete to S_j.

We now use 3.2.3 and 3.2.5 to prove the following:

3.2.6 *Let S be a tournament, let w be a $\{0,1\}$-vector, and let $0 < \lambda < \frac{1}{2}$ be a constant. Then there exist $\epsilon_0, c_1 > 0$ such that for every $0 < \epsilon < \epsilon_0$, every S-free ϵ-critical tournament contains a (c_1, λ, w)-structure.*

Proof. Write $n = |T|$ and $w = (w_1, \dots, w_P)$, where $P > 0$ is an integer. Define $C = |\{i : w_i = 1\}|$. Let $\Lambda = \frac{\lambda}{(4P)^C}$. By 3.2.1 we can choose ϵ_0 small enough such that $|T| > N$, where N is an integer from 3.2.3. Now it follows from 3.2.3 that there exist a constant $c > 0$ and sets $A_1, ..., A_P$ such that $|A_i| \geq cn$ for $i \in \{1, 2, ..., P\}$ and $d(A_i, A_j) \geq 1 - \Lambda$ for $i, j \in \{1, 2, ..., n\}, i < j$. We may assume that $\epsilon_0 < \log_{\frac{c}{2}}(\frac{1}{2})$. We now use 3.2.5 to complete the proof. ∎

Let U be a transitive tournament with $V(U) = \{u_1, u_2, ..., u_{|U|}\}$, where $(u_1, u_2, ..., u_{|U|})$ is a transitive ordering. An (m, c)-*subdivision* of U is defined as the sequence $\mathcal{U}_m^c = (U_1, U_2, ..., U_m)$, where $U_j = \{u_{i_j}, u_{i_j+1}, ..., u_{k_j}\}$ for $i_1, i_2, ..., i_m, k_1, k_2, ..., k_m$ satisfying $1 \leq i_1 \leq k_1 < i_2 \leq k_2 < ... < i_m \leq k_m \leq |U|$ and $|U_j| \geq c|U|$ for $j \in \{1, 2, ..., m\}$.

3.2.7 *Let $m, c_1, c_2, c_3, \epsilon > 0$, be constants, where $m > 0$ is an integer, $0 < c_1, c_2, c_3 < 1$, and $0 < \epsilon < \log_{\frac{c_1}{m}}(1 - c_2c_3)$. Let T be an ϵ-critical tournament with $|T| = n$, and let $A \subseteq V(T)$ with $|A| \geq c_1 n$. Let U be a transitive subtournament of T with $|U| \geq c_2 tr(T)$ and $V(U) \subseteq V(T) \backslash A$, and let $\mathcal{U}_m^{c_3} = (U_1, ..., U_m)$ be an (m, c_3)-subdivision of U. Then there exist vertices $u_1, u_2, ..., u_m, x$ such that*

$x \in A$, $u_i \in U_i$ and u_i is adjacent to x for $i \in \{1, 2, ..., m\}$. Similarly, there exist vertices $w_1, w_2, ..., w_m, d$ such that $d \in A$, $w_i \in U_i$ and d is adjacent to w_i for $i \in \{1, 2, ..., m\}$.

Proof. We prove only the first statement because the latter can be proved analogously. Suppose no such $u_1, u_2, ..., u_m, x$ exist. That means that every $a \in A$ is complete to U_i for at least one $i \in \{1, 2, ..., m\}$. Therefore there exists $i^* \in \{1, 2, ..., m\}$ such that at least $\frac{|A|}{m}$ vertices of A are complete to U_{i^*}. But this contradicts 3.2.2 since T is ϵ-critical and $\epsilon < \log_{\frac{c_1}{m}}(1 - c_2c_3)$. ∎

We continue with more definitions related to (c, λ, w)-structures. Let $(S_1, S_2, ..., S_{|w|})$ be a (c, λ, w)-structure, let $i \in \{1, \dots, |w|\}$, and let $v \in S_i$. We say that v is *M-good with respect to the set S_j* if either $j > i$ and $d(S_j, \{v\}) \leq M\lambda$ or $j < i$ and $d(\{v\}, S_j) \leq M\lambda$; and that v is *M-good with respect to* $(S_1, S_2, ..., S_{|w|})$ if it is M-good with respect to every S_j for $j \in \{1, 2, ..., |w|\} \backslash \{i\}$. Denote by $S_{j,v}$ the set of the vertices of S_j adjacent from v for $j > i$ and adjacent to v for $j < i$. Now, if $v \in S_i$ is M-good with respect to $(S_1, S_2, ..., S_{|w|})$, then $|S_{j,v}| \geq (1 - M\lambda)|S_j|$ for all $j \neq i$. Next we list some easy facts about (c, λ, w)-structures.

3.2.8 *Let $(S_1, S_2, ..., S_{|w|})$ be a (c, λ, w)-structure. Then for every $i, j \in \{1, 2, ..., |w|\}$, $i \neq j$ all but at most $\frac{1}{M}|S_i|$ of the vertices of S_i are M-good with respect to S_j.*

Proof. We may assume without loss of generality that $i < j$ (for $i \geq j$ the proof is analogous). Denote by $B \subseteq S_i$ the set of the vertices of S_i that are not M-good with respect to S_j. From the definition of M-goodness we have $d(B, S_j) < (1 - M\lambda)$. Therefore $|B| \leq \frac{1}{M}|S_i|$ because otherwise we get a contradiction to 3.2.4 taking $X = B$, $Y = S_j$. ∎

3.2.9 *Let $(S_1, S_2, ..., S_{|w|})$ be a (c, λ, w)-structure. Then for every $i \in \{1, 2, ..., |w|\}$ all but at most $\frac{|w|}{M}|S_i|$ of the vertices of S_i are M-good with respect to $(S_1, S_2, ..., S_{|w|})$.*

Proof. Denote by B_j the subset of vertices of S_i that are not M-good with respect to S_j for $j \in \{1, 2, ..., |w|\} \backslash \{i\}$. Denote by B the subset of vertices of S_i that are not M-good with respect to $(S_1, S_2, ..., S_{|w|})$. We have: $B = \bigcup_{j \neq i} B_j$. From 3.2.8 we know that $|B_j| \leq \frac{1}{M}|S_i|$. Therefore we have: $|B| \leq \frac{|w|}{M}|S_i|$. ∎

4

Constellations

4.1 Basic definitions

For a long time the only prime tournaments for which the Erdös-Hajnal Conjecture was known to be true were all prime tournaments on at most 4 vertices and some prime tournaments on 5 vertices. In this chapter we will prove the conjecture for an infinite family of tournaments called *constellations* that contains infinitely many prime tournaments ([10]). All the methods that allow us to prove the conjecture for several classes of tournaments can be used to obtain lower bounds on the coefficient $\epsilon(H)$ as a function of the order of a tournament.In fact even though in the proofs that constellations and C_5 satisfy the conjecture we use regularity lemma, it is possible to avoid it. Thus our approach may be easily applied to construct algorithms finding large transitive subsets in certain families of tournaments with forbidden substructures. We will not discuss algorithmic results in the thesis. An algorithmic approach that does not use regularity lemma is presented in [9]. We mentioned in the introduction that the conjecture is still open for undirected paths of at least 5 vertices. In the directed scenario there is an analogous family of so-called *directed paths* (it will be defined later). Every directed path on at least 5 vertices is prime. It turns out that every directed path is also a constellation. Thus as a corollary of our main result we will prove that every directed path satisfies the Erdös-Hajnal Conjecture. Another corollary is that all tournaments on at most five vertices satisfy the conjecture. However to prove the latter we will also need to prove the conjecture for C_5 which is not a constellation. Both corollaries were first proven in [6] and were implied by the fact that the family of so-called *galaxies* satisfies the conjecture and so does C_5. However now we will prove the conjecture for a larger family of tournaments than galaxies. We now introduce definitions that will be used later in this chapter.

Let T be a tournament with vertex set $V(T)$. Fix some ordering of its vertices. The *graph of backward edges* under this ordering, denoted by $B(T, \theta)$, has vertex set $V(T)$, and $v_iv_j \in E(B(T, \theta))$ if and only if (v_i, v_j) or (v_j, v_i) is a backward edge of T under the ordering θ. For an integer t, we call the graph $K_{1,t}$ a *star*. Let S be a star with vertex set $\{c, l_1, \ldots, l_t\}$, where c is adjacent to vertices $l_1, \ldots, l_t$. We call c the *center of the star*, and $l_1, \ldots, l_t$ *the leaves of the star*. Note that in the case $t = 1$ we may choose arbitrarily any one of the two vertices to be the center of the star, and the other vertex is then considered to be the leaf. Let $\theta = (v_1, v_2, ..., v_n)$ be an ordering of the vertex set $V(T)$ of a n-vertex tournament T. For a subset $S \subseteq V(T)$ we say that $v_i \in S$ is a *left point of S under θ* if $i = \min\{j : v_j \in S\}$. We say that $v_i \in S$ is a *right point of S under θ* if $i = \max\{j : v_j \in S\}$. If from the context it is clear which ordering is taken we simply say: *left point of S* or *right point of S*. For an ordering θ and two vertices v_i, v_j with $i \neq j$ we say that v_i is *before* v_j if $i < j$ and *after* v_j otherwise. We say that a vertex v_j is *between* two vertices v_i, v_k under an ordering $\theta = (v_1, ..., v_n)$ if $i < j < k$ or $k < j < i$. We denote by P_k for $k = 1, 2, ...$ a tournament for which there exists an ordering of vertices under which the graph of backward edges is a path of k vertices. We call it the *directed path* with k vertices.

A *right star* in $B(T, \theta)$ is an induced subgraph with vertex set $\{v_{i_0}, \ldots, v_{i_t}\}$, such that $B(T, \theta)|\{v_{i_0}, \ldots, v_{i_t}\}$ is a star with center v_{i_t}, and $i_t > i_0, \ldots, i_{t-1}$. In this case we also say that $\{v_{i_0}, \ldots, v_{i_t}\}$ is a right star in T. A *left star* in $B(T, \theta)$ is an induced subgraph with vertex set $\{v_{i_0}, \ldots, v_{i_t}\}$, such that $B(T, \theta)|\{v_{i_0}, \ldots, v_{i_t}\}$ is a star with center v_{i_0}, and $i_0 < i_1, \ldots, i_t$. In this case we also say that $\{v_{i_0}, \ldots, v_{i_t}\}$ is a left star in T. A *star* in $B(T, \theta)$ is a left star or a right star.

Let H be a tournament and assume there is an ordering θ of its vertices such that every connected component of $B(H, \theta)$ is either a star or a singleton under this ordering. We call this ordering a *star ordering.* A star ordering of the vertices of the tournament under which no center of a star is between leaves of another star is called a *galaxy ordering.* A tournament is a *galaxy* if its set of vertices has a galaxy ordering. In [6] in joined work with Eli Berger and Maria Chudnovsky we proved that:

4.1.1 *Every galaxy has the Erdös-Hajnal property.*

The *interstellar graph* of H under a star ordering θ is an undirected graph, whose vertices are the sets of leaves of the stars of H under θ and two vertices L_1 and L_2 are adjacent if:

- the left point of L_1 preceeds the right point of L_2 in θ and
- the left point of L_2 preceeds the right point of L_1 in θ

For each connected component C of the interstellar graph of H denote by $Z(C)$ the union of subsets of $V(H)$ corresponding to its vertices (this is the union of some subsets of $V(H)$ of the vertices of H). Next let us define $\mathcal{C}(Z(C))$ as follows. We say that a vertex $v \in \mathcal{C}(Z(C))$ if $v \in Z(C)$ or v is between some two vertices of $Z(C)$ under the ordering θ. Let $C_1, ..., C_k$ be the connected components of the interstellar graph. Note that for any given $1 \leq i < j \leq k$ either every vertex of $\mathcal{C}(Z(C_i))$ is before every vertex of $\mathcal{C}(Z(C_j))$, or every vertex of $\mathcal{C}(Z(C_j))$ is before every vertex of $\mathcal{C}(Z(C_i))$. Thus there is a natural ordering of the sets $\mathcal{C}(Z(C_i))$ for $i = 1, 2, ..., k$ induced by the ordering of the vertices. Denote the ordered sequence of the sets $\mathcal{C}(Z(C_i))$ for $i = 1, 2, ..., k$ as $(\mathcal{W}_1, ..., , \mathcal{W}_k)$, where a set $\mathcal{W}_i$ is before a set $\mathcal{W}_j$ for $1 \leq i < j \leq k$. Denote $\mathcal{W}_0 = \mathcal{W}_{k+1} = \emptyset$. For $i = 1, 2, ..., k+1$ denote by $\mathcal{R}_i$ the set of the vertices of H that are after all the vertices of $\mathcal{W}_{i-1}$ and before all the vertices of $\mathcal{W}_i$ under the ordering θ. Note that if $\mathcal{R}_i$ is nonempty then all its elements are centers of the stars of H. Denote the set of nonempty sets $\mathcal{R}_i$ as $\{\mathcal{M}_1, ..., \mathcal{M}_r\}$ for some $r \geq 0$. Note that $\{\mathcal{W}_1, ..., \mathcal{W}_k, \mathcal{M}_1, ..., \mathcal{M}_r\}$ is a partition of the vertices of H. Denote this partition by $P_\theta(H)$. We are ready to define constellations.

A tournament T is a *constellation* if there exists a star ordering θ of its vertices such that if a center of a star is in some $P \in P_\theta(H)$ then no leaf of this star is in P.

We call such an ordering a *constellation ordering* of T. Let $\Sigma_1, \ldots, \Sigma_l$ be the non-singleton components of $B(T, \theta)$. We say that $\Sigma_1, \ldots, \Sigma_l$ are the *stars of* T *under* θ. If $V(T) = \bigcup_{i=1}^{l} V(\Sigma_l)$, we say that T is a *regular constellation*.

The goal of this chapter is to prove the following result that first appeared in [**10**]:

4.1.2 *Every constellation has the Erdös-Hajnal property.*

Theorem 4.1.2 extends Theorem 4.1.1 since every galaxy is a constellation.

4.2 Erdös-Hajnal property of constellations

In this section we prove Theorem 4.1.2.

Let s be a $\{0,1\}$-vector. Denote by s_c the vector obtained from s by replacing every subsequence of consecutive $1's$ by single 1. Let $\delta^s : \{i : s_c = 1\} \rightarrow N$ be a function that assigns to every nonzero entry of s_c the number of consecutive $1's$ of s replaced by that entry of s_c.

Let H be a regular constellation, and let $\theta = (v_1, \dots, v_h)$ be its constellation ordering. Let $\Sigma_1, \dots, \Sigma_l$ be the stars of H. For $i \in \{0, \dots, l\}$ define $H^i = H|\bigcup_{j=1}^{i} V(\Sigma_j)$, where $H^l = H$, and H^0 is the empty tournament. Let $s^{H,\theta}$ be a $\{0,1\}$-vector such that $s^{H,\theta}(i) = 1$ if and only if v_i is a leaf of one of the stars of H. We say that a (c, λ, w)-structure *corresponds* to H under the ordering θ if $w = s_c^{H,\theta}$. We say that a (c, λ, w)-structure is *constellation-correlated* with H under the ordering θ if the length of w is the number of parts of the partition $P_\theta(H)$ (see the definition of a constellation) and $w_i = 1$ for $i = 1, 2, ..., |P_\theta(H)|$.

Let $(S_1, S_2, ..., S_{|w|})$ be a strong (c, λ, w)-structure that corresponds to H under θ, and let i_r be such that $w(i_r) = 1$. Assume that $S_{i_r} = \{s_{i_r}^1, ..., s_{i_r}^{|S_{i_r}|}\}$ and $(s_{i_r}^1, ..., s_{i_r}^{|S_{i_r}|})$ is a transitive ordering. Write $m(i_r) = \lfloor \frac{|S_j|}{\delta^w(i_r)} \rfloor$.
Write $S_{i_r}^j = \{s_{i_r}^{(j-1)m(i_r)+1}, ..., s_{i_r}^{jm(i_r)}\}$ for $j \in \{1, 2, ..., \delta^w(i_r)\}$. For every $v \in S_{i_r}^j$ write $\xi(v) = (|\{k < i_r : w(i) = 0\}| + \sum_{k<i_r:w(i)=1} \delta^w(k)) + j$. For every $v \in S_{i_r}$ such

that $w(i_r)=0$ write $\xi(v)=(|\{k<i_r : w(i)=0\}|+\sum_{k<i_r:w(i)=1}\delta^w(k))+1$. We say that H is *well-contained* in $(S_1,S_2,...,S_{|w|})$ that corresponds to H if there is a homomorphism f of H into $T|\bigcup_{i=1}^{|w|}S_i$ such that $\xi(f(v_j))=j$ for every $j\in\{1,\dots,h\}$.

Write $P_\theta(H)=\{P_1,P_2,...,P_z\}$ and assume that first $|P_1|$ vertices of H under θ are in P_1, next $|P_2|$ are in P_2, etc. Let $(S_1,S_2,...,S_{|w|})$ be a (c,λ,w)-structure constellation-correlated with H under θ. Assume that $S_i=\{s_i^1,...,s_i^{|S_i|}\}$ and that $(s_1^i,...,s_{|S_i|}^i)$ is a transitive ordering. Write $n(i)=\lfloor\frac{|S_i|}{|P_i|}\rfloor$. Write $S_i^j=\{s_i^{(j-1)n(i)+1},...,s_i^{jn(i)}\}$ for $j\in\{1,2,...,|P_i|\}$. For every $v\in S_i^j$ write $\xi(v)=\sum_{k<i}|P_i|+j$. We say that H is *constellation-contained* in $(S_1,S_2,...,S_{|w|})$ that is constellation-correlated with H if there is a homomorphism f of H into $T|\bigcup_{i=1}^{|w|}S_i$ such that $\xi(f(v_j))=j$ for every $j\in\{1,\dots,h\}$.

At the very beginning we need the following technical lemma:

4.2.1 *Let H be a regular constellation with $|H|=h$ and let θ be its constellation-ordering. Let $\Sigma_1,\Sigma_2,...,\Sigma_l$ be the stars of H under θ. Let $c>0$, $0<\lambda\leq\frac{1}{h^2(2(h+1))^{2h+2}}$ be constants, and w be a vector. Fix $k\in\{0,...,l\}$. Let T be a tournament and let $(S_1,...,S_{|w|})$ be a strong $(\frac{c}{(2(h+1))^{l-k}},(2(h+1))^{2(l-k)}\lambda,w)$-structure in T corresponding to H^k. Then there exists $\epsilon_k>0$ such that if $0<\epsilon<\epsilon_k$ and T is ϵ-critical, then H^k is well-contained in $(S_1,...,S_{|w|})$.*

Proof. Let $h_1,...,h_{|H|}$ be the vertices of H in order θ. Let $\Sigma_1,...,\Sigma_l$ be the stars of H under θ. Write $|T| = n$. Taking $\epsilon_k > 0$ small enough we may assume that $tr(T) \geq \frac{h(h+1)}{c}$ by 3.2.1. The proof is by induction on k. For $k = 0$ the statement is obvious since H^0 is the empty tournament. Write $M = 2h(h+1), \hat{c} = \frac{c}{(2(h+1))^{l-k}}, \hat{\lambda} = (2(h+1))^{2(l-k)}\lambda$. By 3.2.9 we know that for every $i \in \{1, ..., |w|\}$ every S_i contains at least $(1 - \frac{1}{2(h+1)})|S_i|$ M-good vertices with respect to $(S_1, ..., S_{|w|})$. We call this property the *purity property* of $(S_1, ..., S_{|w|})$. Assume that h_{q_0} is the center of Σ_k and $h_{q_1}, ...h_{q_p}$ are its leaves for some integer $p > 0$. For $i \in \{0, \ldots, p\}$, define D_i to be the set of all vertices v of $\bigcup_{i=1}^{|w|} S_i$ with $\xi(v) = q_i$ that are M-good with respect to $(S_1, ..., S_{|w|})$. Note that for $1 \leq i < j \leq p$ subset D_i is complete to the subset D_j. Besides each D_i for $i = 1, 2, ..., p$ induces a transitive subtournament. From the purity property and the fact that $tr(T) \geq \frac{h(h+1)}{c}$ it follows that $|D_i| \geq \frac{\hat{c}}{2(h+1)} tr(T)$ for $i = \{1, ..., p\}$, and $|D_0| \geq \frac{\hat{c}}{2} n$. We may assume that $\epsilon_k < \log_{\frac{\hat{c}}{2h}}(1 - \frac{\hat{c}}{2(h+1)})$. Now we use 3.2.7 to conclude that there exist vertices: $d_0, ..., d_p$ such that $d_i \in D_i$ for $i = 0, ..., p$ and

- $d_1, ..., d_p$ are all adjacent to d_0 if Q is a left-star, and
- $d_1, ..., d_p$ are all adjacent from d_0 if Q is a right-star.

Therefore $\{d_0, ..., d_p\}$ induces a copy of Σ_k. Let $x \in \{1, ..., |w|\}$ be such that $d_0 \in S_x$. Now since $(S_1, ..., S_{|w|})$ corresponds to H^k and $h_{q_1}, ..., h_{q_p}$ are leaves , we also know that there exist $y_1, ..., y_p \in \{1, ..., |w|\} \backslash \{x\}$ so that $d_i \in S_{y_i}$ for all $i \in \{1, ..., p\}$, and $T|(S_{y_1} \bigcup ... \bigcup S_{y_p})$ is a transitive tournament. Let $i \in \{1, ..., |w|\} \backslash \{x, y_1, ..., y_p\}$.

Write $S_i^f = \bigcap_{j=0}^{p} S_{i,d_j}$. Since each d_j is M-good with respect to $(S_1, ..., S_{|w|})$ we have $|S_{i,d_j}| \geq (1 - M\hat{\lambda})|S_i|$. Therefore $|S_i^f| \geq (1 - Mh\hat{\lambda})|S_i|$. By the definition of $\hat{\lambda}$ we conclude that $|S_i^f| \geq (1 - \frac{1}{2(h+1)})|S_i|$. Write $\mathcal{H} = \{1, ..., h\} \backslash \{q_0, ...q_p\}$. If $\{v \in S_{y_i} : \xi(v) \in \mathcal{H}\} \neq \emptyset$, then we define $S^*_{y_i} = S_{y_i, d_0}$. By a similar argument as above we conclude that if $S^*_{y_i}$ is defined then $|S^*_{y_i}| \geq (1 - \frac{1}{2(h+1)})|S_{y_i}|$. If $S^*_{y_i}$ is defined then define $\hat{S_{y_i}} = \{v \in S^*_{y_i} : \xi(v) \in \mathcal{H}\}$. Let $I_{y_i} = \{j : \exists_{v \in \hat{S_{y_i}}} \xi(v) = j\}$. Note that if $\hat{S_{y_i}}$ is defined then $I_{y_i} \neq \emptyset$. Assume now that $\hat{S_{y_i}}$ is defined. For every $j \in I_{y_i}$ select arbitrarily $\lceil \frac{\hat{c}}{2(h+1)} \rceil$ vertices v in $\hat{S_{y_i}}$ with $\xi(v) = j$ and denote the union of these $|I_{y_i}|$ sets by $S^f_{y_i}$. We can always do this selection since for every $j \in I_{y_i}$ we have $|v : \xi(v) = j| \geq \frac{|S_{y_i}|}{h+1}$ and also $|S^*_{y_i}| \geq (1 - \frac{1}{2(h+1)})|S_{y_i}|$. Thus we have defined some number of sets S_i^f. Denote this number by t. We have: $|S_i^f| \geq \frac{\hat{c}}{2(h+1)} tr(T)$ for every (defined) S_i^f with $w(i) = 1$ and $|S_i^f| \geq \frac{\hat{c}}{2(h+1)} n$ for every (defined) S_i^f with $w(i) = 0$. Now 3.2.4, implies that the sets $S_1^f, ..., S_t^f$ form a strong $(\frac{\hat{c}}{2(h+1)}, 4(h+1)^2\hat{\lambda}, z)$-structure that corresponds to H^{k-1} for an appropriate vector z. Inductively H^{k-1} is well-contained in this structure for $\epsilon_k > 0$ small enough. But now we can merge the well-contained copy of H^{k-1} and a copy of Σ_k that we have already found to get a well-contained copy of H^k. This completes the proof. ∎

From the previous lemma we get the following lemma:

4.2.2 *Let H be a regular constellation with $|H| = h$. Let c be a positive constant. Then there exists $\epsilon > 0$ such that every ϵ-critical tournament containing a strong*

$(c, \frac{1}{h^2(2h+2)^{2h+2}}, z)$*-structure corresponding to* H *for an appropriate vector* z *is not* H*-free.*

Proof. Let θ be a constellation-ordering of H and let $\Sigma_1, ..., \Sigma_l$ be the stars of H under θ. Then Lemma 4.2.2 follows from Lemma 4.2.1 if we take $k = l$. ∎

We also need the following lemma:

4.2.3 *Let* H *be a regular constellation with* $|H| = h$ *and let* θ *be its constellation-ordering. Let* $\Sigma_1, \Sigma_2, ..., \Sigma_l$ *be the stars of* H *under* θ. *Let* $c > 0$, $0 < \lambda \leq \frac{1}{h^2(2(h+1))^{2h+2}}$ *be constants, and* w *be a vector. Fix* $k \in \{0, ..., l\}$. *Let* T *be a tournament and let* $(S_1, ..., S_{|w|})$ *be* $(\frac{c}{(2(h+1))^{l-k}}, (2(h+1))^{2(l-k)}\lambda, w)$*-structure in* T *which is constellation-correlated with* H^k. *Then there exists* $\epsilon_k > 0$ *and* $c_1 > 0$ *such that if* $0 < \epsilon < \epsilon_k$ *and* T *is* ϵ*-critical, then*

- H^k *is constellation-contained in* $(S_1, ..., S_{|w|})$ *or*
- *there exist* $1 \leq i < j \leq |w|$ *and* $\mathcal{S}_i \subseteq S_i$, $\mathcal{S}_j \subseteq S_j$ *such that* $\mathcal{S}_i$ *is complete to* $\mathcal{S}_j$ *and* $|\mathcal{S}_i| \geq \frac{c}{(2(h+1))^l h} tr(T)$, $|\mathcal{S}_j| \geq \frac{c}{(2(h+1))^l h} tr(T)$.

Proof. The proof is similar to the proof of Lemma 4.2.1. Let $h_1, ..., h_{|H|}$ be the vertices of H in order θ. Let $\Sigma_1, ..., \Sigma_l$ be the stars of H under θ. Write $|T| = n$. Taking $\epsilon_k > 0$ small enough we may assume that $tr(T) \geq \frac{h(h+1)}{c}$ by 3.2.1. The proof is by induction on k. For $k = 0$ the statement is obvious since H^0 is an empty tournament. So it suffices to prove the statement for $k \geq 1$. Write $M = 2h(h+1), \hat{c} = \frac{c}{(2(h+1))^{l-k}}, \hat{\lambda} = (2(h+1))^{2(l-k)}\lambda$. By 3.2.9 we know that for every $i \in \{1, ..., |w|\}$ every S_i contains at least $(1 - \frac{1}{2(h+1)})|S_i|$ M-good vertices with respect to $(S_1, ..., S_{|w|})$. We call this property the *purity property* of $(S_1, ..., S_{|w|})$.

Assume that h_{q_0} is the center of Σ_k and $h_{q_1}, ...h_{q_p}$ are its leaves for some integer $p > 0$. For $i \in \{0, \ldots, p\}$, define D_i to be the set of all vertices v of $\bigcup_{i=1}^{|w|} S_i$ with $\xi(v) = q_i$ that are M-good with respect to $(S_1, ..., S_{|w|})$. Note that each D_i for $i = 0, 1, 2, ..., p$ induces a transitive subtournament. From the purity property and the fact that $tr(T) \geq \frac{h(h+1)}{c}$ it follows that $|D_i| \geq \frac{\hat{c}}{2(h+1)}tr(T)$ for $i = \{0, 1, ..., p\}$. Assume first that there are no vertices $d_0, d_1, ..., d_p$ such that $d_i \in D_i$ for $i = 0, ..., p$ and

- $d_1, ..., d_p$ are all adjacent to d_0 if Q is a left-star, and
- $d_1, ..., d_p$ are all adjacent from d_0 if Q is a right-star.

Without loss of generality assume that Σ_k is a left star. Then for every $d \in D_0$ there exists $i_d \in \{1, 2, ..., p\}$ such that $\{d\}$ is complete to D_{i_d}. Therefore, by pigeonhole principle, there exists $i^* \in \{1, 2, ..., p\}$ and a set $D^* \subseteq D_0$ such that $|D^*| \geq \frac{|D_0|}{p}$ and D^* is complete to D_{i^*}. But then we can take $\mathcal{S}_i = D^*$ and $\mathcal{S}_j = D_{i^*}$ and we are done. Thus we conclude that there exist vertices: $d_0, ..., d_p$ such that $d_i \in D_i$ for $i = 0, ..., p$ and

- $d_1, ..., d_p$ are all adjacent to d_0 if Q is a left-star, and

- $d_1, ..., d_p$ are all adjacent from d_0 if Q is a right-star.

Therefore $\{d_0, ..., d_p\}$ induces a copy of Σ_k. Let $x \in \{1, ..., |w|\}$ be such that $d_0 \in S_x$. Now, since $(S_1, ..., S_{|w|})$ is constellation-corellated with H^k and $h_{q_1}, ..., h_{q_p}$ are leaves , we also know that there exists $y \in \{1, ..., |w|\} \backslash \{x\}$ so that $d_i \in S_y$ for all $i \in \{1, ..., p\}$, and $T|S_y$ is a transitive tournament. Let $i \in \{1, ..., |w|\} \backslash \{x, y\}$. Write $S_i^f = \bigcap_{j=0}^{p} S_{i,d_j}$. Since each d_j is M-good with respect to $(S_1, ..., S_{|w|})$ we have $|S_{i,d_j}| \geq (1 - M\hat{\lambda})|S_i|$. Therefore $|S_i^f| \geq (1 - Mh\hat{\lambda})|S_i|$. By the definition of $\hat{\lambda}$ we conclude that $|S_i^f| \geq (1 - \frac{1}{2(h+1)})|S_i|$. Write $\mathcal{H} = \{1, ..., h\} \backslash \{q_0, ...q_p\}$. If $\{v \in S_y : \xi(v) \in \mathcal{H}\} \neq \emptyset$, then we define $S_y^* = S_{y,d_0}$. By a similar argument as above we conclude that if S_y^* is defined then $|S_y^*| \geq (1 - \frac{1}{2(h+1)})|S_y|$. If S_y^* is defined then define $\hat{S}_y = \{v \in S_y^* : \xi(v) \in \mathcal{H}\}$. Let $I_y = \{j : \exists_{v \in \hat{S}_y} \xi(v) = j\}$. Note that if $\hat{S}_y$ is defined then $I_y \neq \emptyset$. Assume now that $\hat{S}_y$ is defined. For every $j \in I_y$ select arbitrarily $\lceil \frac{\hat{c}}{2(h+1)} \rceil$ vertices v in $\hat{S}_y$ with $\xi(v) = j$ and denote the union of these $|I_y|$ sets by S_y^f. We can always do this selection since for every $j \in I_y$ we have $|v : \xi(v) = j| \geq \frac{|S_y|}{h+1}$ and also $|S_y^*| \geq (1 - \frac{1}{2(h+1)})|S_y|$. Thus we have defined some number of sets S_i^f. Denote this number by t. We have: $|S_i^f| \geq \frac{\hat{c}}{2(h+1)} tr(T)$ for every (defined) S_i^f. Now 3.2.4, implies that the sets $S_1^f, ..., S_t^f$ form a $(\frac{\hat{c}}{2(h+1)}, 4(h+1)^2\hat{\lambda}, z)$-structure that is constellation-correlated with H^{k-1} for an appropriate vector z. Inductively either H^{k-1} is constellation-contained in this structure for $\epsilon_k > 0$ small enough or there exists $1 \leq i < j \leq |w|$ and $\mathcal{S}_i \subseteq S_i$, $\mathcal{S}_j \subseteq S_j$ such that $\mathcal{S}_i$ is complete to $\mathcal{S}_j$ and $|\mathcal{S}_i| \geq \frac{c}{(2(h+1))^l h} tr(T)$, $|\mathcal{S}_j| \geq \frac{c}{(2(h+1))^l h} tr(T)$. If the latter follows then we are obviously done. However if the former follows then we are also done since if this is the case we can merge the constellation-contained copy of H^{k-1} and a copy of Σ_k that we have already found to get a constellation-contained copy of H^k. This completes the proof. ∎

From the previous lemma we get the following lemma:

4.2.4 *Let H be a regular constellation with $|H| = h$. Let c be a positive constant. Then there exists $\epsilon_1 > 0$ such that for every $0 < \epsilon < \epsilon_1$, for every ϵ-critical tournament T containing a $(c, \frac{1}{h^2(2h+2)^{2h+2}}, z)$-structure $\Pi = (S_1, ..., S_{|z|})$, constellation-correlated with H for an appropriate vector z, the following holds:*

- *T is not H-free or*
- *there exists $1 \leq i < j \leq |z|$ and $\mathcal{S}_i \subseteq S_i$, $\mathcal{S}_j \subseteq S_j$ such that $\mathcal{S}_i$ is complete to $\mathcal{S}_j$ and $|\mathcal{S}_i| \geq \frac{c}{(2(h+1))^l h} tr(T)$, $|\mathcal{S}_j| \geq \frac{c}{(2(h+1))^l h} tr(T)$.*

Proof. Let θ be a constellation-ordering of H and let $\Sigma_1, ..., \Sigma_l$ be the stars of H under θ. Then Lemma 4.2.4 follows from Lemma 4.2.3 if we take $k = l$. ∎

We are ready to prove Theorem 4.1.2.

Proof. Fix some constellation H. We can assume that it is regular since every constellation is a subtournament of some regular constellation. Denote its constellation ordering by θ and related partition by $P_\theta(H)$. Let $\Sigma_1, ..., \Sigma_l$ be the stars of H under θ. Assume that the maximal number of consecutive 0's in a vector w of a strong (c, λ, w)-structure that corresponds to H is s and the number of 1's in w is u (note that vector w depends only on H and the ordering θ). Let T be an ϵ-critical tournament. It is enough to show that for $\epsilon > 0$ small enough T is not H-free. Denote by ϵ_1 the ϵ from Lemma 4.2.4, taken for $c = c_0((2(h+1))^l h)^{-\binom{r}{2}}$. We may assume that for the ϵ-critical tournament T that we consider we have $\epsilon < \epsilon_1$. Write $\Lambda_0 = h^{-2-2\binom{r}{2}}(2h+2)^{-2h-2-2l\binom{r}{2}}$ and $r = 2^{\max\{u, |P_\theta(H)|\}+1}$. We may assume, according to Lemma 3.2.6, that T contains a (c_0, Λ_0, z)-structure for $z = (z_0, ..., z_{(r+1)s+r-1})$, where $z_i = 0$ for $i \bmod (s+1) \leq (s-1)$ and $z_i = 1$ for $i \bmod (s+1) = s$ for some constant $c_0 > 0$. Now consider the sequence of undirected graphs $(G_0, G_1, ...)$ and the sequence of (c, λ, w)-structures $(\Pi_0, \Pi_1, ...)$ defined recursively as follows. G_0 is a graph on r vertices with no edges such that $V(G_0) = \{1, 2, ..., r\}$ and $\Pi_0 = (S^0_1, ..., S^0_{(r+1)s+r})$ is some fixed (c_0, Λ, z)-structure in T. Assume that we have defined G_i and $\Pi_i = (S^i_1, ..., S^i_{(r+1)s+r})$. Assume furthermore that Π_i is a (c_i, Λ_i, z)-structure for some $c_i > 0$ and $\Lambda_i > 0$. For every $1 \leq m \leq r$ define $\rho^i(m) = S^i_{m^*}$, where m^* is the index that corresponds to the m^{th} nonzero entry in z. If G_i does not contain stable sets of size $|P_\theta(H)|$ then G_{i+1} and Π_{i+1} are not defined. If this is the case we say that we reached *state* 0.

If G_i contains a stable set of size $|P_\theta(H)|$ then take any of them and denote it by $\mathcal{A} = \{g_1, ..., g_{|P_\theta(H)|}\}$. Without loss of generality assume that $g_1 < g_2 < ... < g_{|P_\theta(H)|}$. Consider the following sequence $(S^i_{\rho^i(g_1)}, ..., S^i_{\rho^i(g_{|P_\theta(H)|})})$. It is a (c_i, Λ_i, q)-structure for some appropriate vector q. Note that it is constellation-correlated with H. If $\Lambda_i > \frac{1}{h^2(2h+2)^{2h+2}}$ then we do not define G_{i+1} and Π_{i+1} and we say that the *state* 1 was reached. So assume that $\Lambda_i \leq \frac{1}{h^2(2h+2)^{2h+2}}$. Assume also that we have $c_i \geq c_0((2(h+1))^l h)^{-\binom{r}{2}}$. But now we can use Lemma 4.2.4 and conclude (taking ϵ to be small enough) that either T is not H-free or there exists $g_{i_1} < g_{j_1}$ and $\mathcal{S}_{\rho^i(g_{i_1})} \subseteq S^i_{\rho^i(g_{i_1})}$, $\mathcal{S}_{\rho^i(g_{j_1})} \subseteq S^i_{\rho^i(g_{j_1})}$ such that $\mathcal{S}_{\rho^i(g_{i_1})}$ is complete to $\mathcal{S}_{\rho^i(g_{j_1})}$ and $|\mathcal{S}_{\rho^i(g_{i_1})}| \geq \frac{c_i}{(2(h+1))^l h}$, $|\mathcal{S}_{\rho^i(g_{j_1})}| \geq \frac{c_i}{(2(h+1))^l h}$. If the former holds then we say that we reached *state* 2. So assume that the former does not hold, but the latter holds. Then G_{i+1} is obtained from G_i by adding an edge $\{i_1, j_1\}$. Structure Π_{i+1} is obtained from Π_i by replacing $S^i_{\rho^i(g_{i_1})}$ by $\mathcal{S}_{\rho^i(g_{i_1})}$ and $S^i_{\rho^i(g_{j_1})}$ by $\mathcal{S}_{\rho^i(g_{j_1})}$.

If we have $c_i < c_0((2(h+1))^l h)^{-\binom{r}{2}}$, then we do not define G_{i+1} and Π_{i+1} and we say that *state* 3 was reached. Note that the sequences $(G_0, G_1, ...)$ and $(\Pi_0, \Pi_1, ...)$ are finite. Indeed, graph G_{i+1} has one more edge than G_i and if $V(G_i)$ induces a clique then we reach state 0. Therefore we can write both sequences as: $(G_0, G_1, ...G_l)$, $(\Pi_0, \Pi_1, ...\Pi_l)$, for some $l \leq \binom{r}{2}$. Using induction and Lemma 3.2.4 we easily get that $c_i \geq \frac{c_0}{((2(h+1))^l h)^i}$ and $\Lambda_i \geq \Lambda_0((2(h+1))^l h)^{2i}$. Therefore we can conclude that state 3 can never be reached. Assume that we constructed two sequences: $(G_0, G_1, ...G_l)$ and $(\Pi_0, \Pi_1, ...\Pi_l)$, according to the rules above. Since those sequences are finite, we reached one of the states: $0, 1, 2$. State 1 in fact cannot be reached because of the lower bound on Λ_i we derived above and the formula on Λ_0. Reaching state 2 implies that a copy of H was found in T so we are done. Therefore we can assume that state 0 was reached. Since G_l has $2^{\max(u,|P_\theta(H)|)+1}$ vertices, it has either a clique or stable set of size at least $\max\{u, |P_\theta(H)|\}$. Since state 0 was reached, G_l does not have stable sets of size $|P_\theta(H)|$. Therefore it has a clique of size u. Denote this clique by $\{c_1, c_2, ..., c_u\}$. Let $S^l_{\rho^l(c_1)}, S^l_{\rho^l(c_2)}, ..., S^l_{\rho^l(c_u)}$ be the corresponding sets. Note that from the definition of a vector z we know that $\rho^l(c_j) - \rho^l(c_i) \geq (s+1)$ for $1 \leq i < j \leq u$. But then, again from the definition of u and z, we can easily complete the sequence $(S^l_{\rho^l(c_1)}, S^l_{\rho^l(c_2)}, ..., S^l_{\rho^l(c_u)})$ by the sets of the form S^l_i (for which $z_i = 0$), to get a (c_l, Λ_l, q)-structure (for an appropriate vector q) that corresponds to H. Note that this is a strong (c_l, Λ_l, q)-structure since for every $1 \leq i < j \leq u$ we have: $S^l_{\rho^l(c_i)}$ is complete to $S^l_{\rho^l(c_j)}$. Now, using the derived lower bound on Λ_l and the formula on Λ_0, we can conclude for $\epsilon > 0$ small enough that, according to Lemma 4.2.2, a tournament T contains a copy of H as a subtournament. That completes the proof of Theorem 4.1.2. ∎

Now we prove an interesting corollary of Theorem 4.1.2 ([**6**]) which we restate below:

4.2.5 *For every k the tournament P_k satisfies the Erdös-Hajnal Conjecture.*

Proof. Take a path P_k. We can assume without loss of generality that $k = 2l$ for some l. By Theorem 4.1.2, it is enough to prove that P_k is a constellation. Assume that $V(P_k) = \{1, ..., 2l\}$ and that under the ordering given by this labeling the only backward edges are of the form $(i+1, i)$ for $i = 1, ..., 2l-1$. Now take the following ordering of the vertices of P_k: $\theta = (2, 1, 4, 3, 6, 5, ..., 2l, 2l-1)$. Under this ordering the set of backward edges is the collection of edges of the form $(2s+1, 2s)$ for $s = 1, ..., l-1$. Therefore P_k is a constellation and the result follows. ∎

5

The Erdös-Hajnal Conjecture for small tournaments

5.1 Introduction

In this chapter we prove that every tournament on at most 5 vertices has the Erdös-Hajnal property. Thus, we prove the following result of [6]:

5.1.1 *The Erdös-Hajnal Conjecture is true for all tournaments on at most* 5 *vertices.*

In the undirected scenario there are three graphs on at most 5 vertices for which the conjecture is still open (path of 5 vertices, its complement and cycle of length

5). Unfortunately the result for tournaments does not seem to provide clues to prove the undirected version of the conjecture for those three remaining undirected graphs.

We need some definitions. Denote by C_5 the (unique) tournament on 5 vertices in which every vertex is adjacent to exactly two other vertices. One way to construct this tournament is to take a vertex set $\{0, 1, 2, 3, 4\}$ and make vertex i adjacent to $i + 1 \mod 5$ and $i + 2 \mod 5$ for $i = 0, 1, 2, 3, 4$. The tournament C_5 is a prime tournament that is not a constellation. It turns out that Theorem 5.1.1 is implied by Theorem 4.1.2, the fact that C_5 has the Erdös-Hajnal property and some results of [7].

Therefore in this chapter we also prove the following result:

5.1.2 *The tournament C_5 has the Erdös-Hajnal property.*

This result first appeared in [6].

5.2 Small tournaments

Our goal in this section is to prove Theorem 5.1.1. We asssume here that C_5 has the Erdös-Hajnal property. The proof of this fact is given in the next section. First, we need some definitions. Let us remind that a tournament S is a *celebrity* if there exists a constant $c(S)$, with $0 < c(S) \leq 1$, such that every S-free tournament T satisfies $tr(T) \geq c(S)|T|$. Celebrities were fully characterized in [7].

Let G_1 be the tournament with 5 vertices $v_1, \ldots, v_5$, such that under the ordering $(v_1, \ldots, v_5)$ the backward edges are: $(v_4, v_1), (v_5, v_2)$. Let G_2 be the tournament with 5 vertices $w_1, \ldots, w_5$, such that under the ordering $(w_1, \ldots, w_5)$ the backward edges are: $(w_5, w_1), (w_5, w_3)$.

We need the following result from [7].

5.2.1 *Every tournament on at most* 5 *vertices, except* C_5, G_1, G_2, *is a celebrity.*

We are ready to prove 5.1.1:

Proof. Clearly every celebrity satisfies the Erdös-Hajnal Conjecture, so by 5.2.1 it is enough to prove the result form G_1, G_2, C_5. Since $(v_1, \ldots, v_5)$ is a constellation ordering of G_1, and $(w_1, \ldots, w_5)$ is a constellation ordering of G_2, Theorem 4.1.2 implies that both G_1 and G_2 satisfy the Erdös-Hajnal Conjecture, and by 5.1.2 so does C_5. This completes the proof. ∎

.

5.3 The tournament C_5

We prove in this section Theorem 5.1.2. We start with some preliminary observations. Let $v_1, ..., v_5$ be the vertices of C_5. Then there exists an ordering $(v_{\theta(1)}, v_{\theta(2)}, v_{\theta(3)}, v_{\theta(4)}, v_{\theta(5)})$ of $v_1, ..., v_5$ where the set of backward edges is the following $\{(v_{\theta(5)}, v_{\theta(1)}), (v_{\theta(4)}, v_{\theta(1)}), (v_{\theta(5)}, v_{\theta(2)})\}$. We call this ordering the *path ordering of* C_5 since under this ordering the set of backward edges forms a path (and one isolated vertex). There also exists an ordering $(v_{\rho(1)}, v_{\rho(2)}, v_{\rho(3)}, v_{\rho(4)}, v_{\rho(5)})$ of the vertices of C_5 where the set of backward edges is

$\{(v_{\rho(5)}, v_{\rho(3)}), (v_{\rho(3)}, v_{\rho(1)}), (v_{\rho(5)}, v_{\rho(1)}), (v_{\rho(4)}, v_{\rho(2)})\}$. We call this ordering the *cyclic ordering of* C_5, since under this ordering the set of backward edges forms a graph containing a cycle (a triangle plus an edge).

5.3.1 *Let $c, d > 0$, $0 < \lambda < 1$, $\epsilon < \log_{\frac{dc}{2}}(\frac{1}{2})$ and $w = (0, 0, 1, 0, 0)$. Let $(S_1, ..., S_5)$ be a (c, λ, w)-structure of an ϵ-critical tournament T. Let $s_1 \in S_1, s_3 \in S_3, s_5 \in S_5$. Assume that s_5 is adjacent to both s_1 and s_3 and s_3 is adjacent to s_1. Let $\hat{S}_2$ be the subset of the vertices of S_2 adjacent to s_3, s_5 and from s_1. Let $\hat{S}_4$ be the subset of the vertices of S_4 adjacent to s_5 and from s_1, s_3. Assume that $|\hat{S}_i| \geq d|S_i|$ for $i \in \{2, 4\}$. Then T contains a copy of C_5.*

Proof. By 4.2.2, and since T is ϵ-critical and $\epsilon < \log_{\frac{dc}{2}}(\frac{1}{2})$, there exist $s_2 \in \hat{S}_2$ and $s_4 \in \hat{S}_4$ such that s_4 is adjacent to s_2. But now $\{s_1, ..., s_5\}$ induces a copy of

C_5 in T and the ordering $(s_1, ..., s_5)$ is a cyclic ordering. ∎

We will now prove 5.1.2 which we restate below:

5.3.2 *The tournament C_5 satisfies the Erdös-Hajnal Conjecture.*

Proof. Assume otherwise. Taking $\epsilon > 0$ small enough, we may assume that there exists a C_5-free ϵ-critical tournament T. By 3.2.6 T contains a (c, λ, w)-structure $(S_1, ..., S_5)$ for some $c > 0$, $\lambda = \frac{1}{720}$ and $w = (0, 0, 1, 0, 0)$. We may assume without loss of generality that $|S_3| \mod 3 = 0$. Let (T_1, T_2, T_3) be a $(3, \frac{1}{3})$-subdivision of S_3. Let $M = 30$. Let S_i^* be the subset of S_i of M-good vertices with respect to $(S_1, ..., S_5)$. By 3.2.9 we have $|S_i^*| \geq (1 - \frac{5}{M})|S_i|$. Denote $T_i^* = S_3^* \cap T_i$ for $i \in \{1, 2, 3\}$. We have: $|T_i^*| \geq \frac{1}{2}|T_i|$. So by 3.2.4 $(S_1^*, S_2, T_1^*, S_4, S_5^*)$ is a $(\frac{c}{6}, 36\lambda, w)$-structure. Similarly, $(S_1^*, S_2, T_3^*, S_4, S_5^*)$ is a $(\frac{c}{6}, 36\lambda, w)$-structure. Write $\delta = \frac{1}{2}(1 - \frac{5}{M})$. We may assume that $\epsilon < \log_\delta(\frac{1}{2})$, and so by 4.2.2 there exists an integer $k \geq \frac{5}{12}c$ and vertices $x_1, ..., x_k, y_1, ..., y_k$ such that $x_i \in S_1^*$, $y_i \in S_5^*$ and y_i is adjacent to x_i for $i \in \{1, ..., k\}$. Denote by X the subset of $\{x_1, ..., x_k\}$ consisting of the vertices with an inneighbour in T_3^*, and by Y the subset of $\{y_1, ..., y_k\}$ consisting of the vertices with an outneighbour in T_1^*. We may assume that $\epsilon < \log_{\frac{5}{36}c}(1 - \frac{c}{6})$, and thus 3.2.2 implies that $|X| > \frac{k}{2}$ and $|Y| > \frac{k}{2}$. Consequently, there exists an index $j \in \{1, ..., k\}$ and vertices x_j, y_j, t_1, t_3 such that $t_1 \in T_1^*$, $t_3 \in T_3^*$, t_3 is adjacent to x_j, and y_j is adjacent to t_1. If x_j is adjacent to t_1 and t_3 is adjacent to y_j then write $E_* = S_{3,x_j} \cap S_{3,y_j} \cap T_2^*$. From the fact that x_j, y_j are M-good with respect to $(S_1, ..., S_5)$ and since $|T_2| = \frac{|S_3|}{3}$, it follows that $|E_*| \geq \frac{1}{2}|T_2|$, in particular $|E_*| > 0$. Let $q \in E_*$. Then x_j, t_1, q, t_3, y_j induce a copy of C_5 in T, where the ordering (x_j, t_1, q, t_3, y_j) is the tree ordering, a contradiction. Therefore we may assume that either t_1 is adjacent to x_j, or y_j is adjacent to t_3. Write $E_i = S_{i,x_j} \cap S_{i,t_1} \cap S_{i,t_3} \cap S_{i,y_j}$ for $i \in \{2, 4\}$. From the fact that x_j, y_j, t_1, t_3 are M-good with respect to $(S_1, ..., S_5)$ it follows that $|E_i| \geq (1 - 4M\lambda)|S_i| \geq \frac{1}{2}|S_i|$ for $i \in \{2, 4\}$.

We may assume that $\epsilon < \log_{\frac{c}{12}}(\frac{1}{2})$. Observe that $(S_1^*, S_2, T_1^*, S_4, S_5^*)$ and $(S_1^*, S_2, T_3^*, S_4, S_5^*)$ are both $(\frac{c}{6}, 36\lambda, w)$-structures. But now, applying 5.3.1 to $(S_1^*, S_2, T_1^*, S_4, S_5^*)$ if t_1 is adjacent to x_j, and to $(S_1^*, S_2, T_3^*, S_4, S_5^*)$ if y_j is adjacent to t_3, we deduce that T contains a copy of C_5 (with the path ordering), a contradiction. ∎

6

Upper bounds for EH-suprema

6.1 Introduction

In this chapter we give upper bounds for EH-suprema of several classes of tournaments. As a byproduct of our methods we will also obtain some results on EH-suprema for undirected graphs. Our main results of the first part of this chapter are the following:

6.1.1 *The Erdös-Hajnal supremum of every undirected graph H is at most $\frac{4}{|H|}$.*

6.1.2 *There exists $\eta > 0$ such that the Erdös-Hajnal supremum of almost every tournament T on k vertices is at most $\frac{4}{k}(1+\eta\frac{\sqrt{\log(k)}}{\sqrt{k}})$, i.e. the proportion of tour-*

naments on k vertices with the supremum exceeding $\frac{4}{k}(1+\eta\frac{\sqrt{\log(k)}}{\sqrt{k}})$ goes to 0 as k goes to infinity.

Besides we show how the parameter $t(H)$ (the number of directed triangles) of a given tournament H can be used to obtain upper bounds on its EH-supremum $\xi(H)$. We also show that tournaments with big EH-suprema have very nonrandom properties by establishing an inequality combining two important parameters of the tournament: EH-supremum $\xi(H)$ and chromatic number $\chi(H)$:

6.1.3 *Every tournament H satisfies:* $\chi(H) < \lfloor \frac{2}{\xi(H)} \rfloor + 1$.

We show these results in the first part of this chapter. This part of the chapter is based on [**11**].

The upper bound $\xi(H) \leq \frac{4}{h}(1+o(1))$ is valid for tournaments with quadratic number of backward edges under every ordering such as random tournaments. However these results do not say anything about $\xi(H)$ for an arbitrarily chosen tournament with no nontrivial homogeneous sets, i.e. an arbitrary prime tournament. We address that problem in the subsequent part, proving the following:

6.1.4 *There exists $C > 0$ such that every prime h-vertex tournament H satisfies* $\xi(H) \leq C\frac{\log(h)}{h}$.

We also introduce a parameter called the *partition number* of a tournament that measures how well it can be decomposed into homogeneous sets. We show its close relation to the EH-supremum of the tournament. We also show that tournaments with small nontrivial homogeneous sets have small EH-suprema. Using quotient graphs of tournaments we also give some structural characterization and upper bounds on sizes of families of tournaments with given lower bound on their EH-suprema. This part of the chapter is based on [**10**].

6.2 EH-suprema of almost all graphs

6.2.1 Probabilistic tools

The most renowned previous work in this area is the determination of the assymptotics of the Ramsey number $R(k, t)$ for fixed k, as t goes to infinity. $R(k, t)$ is the smallest n such that every graph with n vertices contains a clique of size k or an independent set of size t. Thus, the Erdös-Hajnal supremum of a clique of size k is the inverse of the infimum of those c such that $R(k, n)$ exceeds n^c for all sufficiently large n. The earliest results in this area are due to Spencer[1975]. His approach was to choose a parameter p such that the sum of the number of cliques of size k plus the stable sets of size t in a graph with n vertices, where each edge is present independently with probability p, was, on average, less than one. This implied that there was a choice, where the sum was zero, i.e. an n-vertex graph with no clique of size k and no stable set of size at least t, and hence $R(k,t) > n$. One can refine this approach, if the expectation of the sum is at most $\frac{n}{2}$. It can be done by deleting one vertex from each large clique and each large stable set to get a graph which shows that $R(k,t) > \frac{n}{2}$. This gives a slightly better result. One can refine the result even further, using the Lovasz Local Lemma (for details see [**26**]).

To prove Theorem 6.1.1, we proceed in a similar manner. We can assume that H has more edges than non-edges (since we can consider H or its complement). We choose p such that the sum of the number of copies of H plus the number of stable sets of size $(\frac{n}{2})^f$ in a graph with n vertices, where n is sufficiently large and each edge is present independently with probability p is, on average, less than $\frac{n}{2}$. That proves that $\xi(H) \leq f$.

To prove Theorem 6.1.2 we proceed as follows. We consider a random tournament on vertices $v_1, ..., v_n$ with sufficiently large n, where for each pair i, j with $1 \leq i < j \leq n$ independently, $v_i v_j$ is an edge with probability p and $v_j v_i$ is an edge with probability $1-p$. We choose a p such that the sum of the number of vertex sets of copies of T plus the number of transitive subtournaments of size $(\frac{n}{2})^f$ in this random tournament is, on average, less than $\frac{n}{2}$. This proves that the Erdös-Hajnal supremum is at most f.

More precisely, in the proof of Theorem 6.1.2, we choose p so that the expected number of transitive subtournaments of size $(\frac{n}{2})^f$ in the random tournament is, on average, less than one. We then examine what properties of a tournament T ensure that the expected number of vertex sets of copies of T in this random tournament is less than $\frac{n}{2}$.

Using the probabilistic method we show the following:

6.2.1 *The vertices of every tournament H can be ordered in such a way that the set of backward edges of every subtournament S of H has size at most $\frac{|S|-1}{\xi(H)}$.*

Theorem 6.1.2 is a direct consequence of this result since almost every tournament on k vertices has at least $(1+o(1))\frac{k(k-1)}{4}$ backward edges under every ordering of its vertices.

Bounding the size of the set of backward edges also bounds the number of directed triangles of the tournament. Thus, as we will show, Theorem 6.2.1 also implies:

6.2.2 *Let H be an n-vertex tournament H. Then if $t(H) > 0$, the following holds*

$$\xi(H) \leq \min\left\{\frac{2(n-1)(n-2)}{t(H)}, \frac{(n-1)(\sqrt{(2n-5)^2+8t(H)}+(2n-5))}{4t(H)}\right\}$$

Theorem 6.2.1 and 6.1.2 clarify that not all tournaments have Erdös-Hajnal supremum 1. Within the class of tournaments which do, the family of celebrities is a special subclass.

The following has been proven in [7]:

6.2.3 *The vertices of a celebrity H can be ordered in such a way that backward edges form a forest.*

Even though Proposition 6.2.3 does not explicitly follow from Theorem 6.2.1, the techniques used to prove Theorem 6.2.1 can be also adapted to prove Proposition 6.2.3. We remark that Theorem 6.2.1 easily implies Corollary 6.1.3.

6.2.2 Proof of Theorem 6.1.1 and Theorem 6.2.1

We start by proving Theorem 6.1.1. In fact we will prove a slightly stronger result from which Theorem 6.1.1 easily follows. Denote by N the set of nonedges of a given undirected graph H.

6.2.1 *The Erdös-Hajnal supremum of any undirected graph H with edge set E and the set of nonedges N on h vertices satisfies:* $\xi(H) \leq \frac{(h-1)}{\max\{|N|,|E|\}}$.

Proof. We can assume that $|E| > |N|$ since the result holds for H if and only if it holds for the complement of H. Denote $m = |E|$. We will the use probabilistic method. All we need to do is to show for arbitrary $f^* > \frac{h-1}{m}$ an infinite family of graphs G without H as an induced subgraph and without cliques and independent sets of size at least $|G|^{f^*}$. Fix some $f^* > \frac{h-1}{m}$. Denote $\epsilon = \frac{(f^* - \frac{h-1}{m})}{2}$ and let $f = \frac{h-1}{m} + \epsilon$. Consider a random graph G^n_f on n vertices, where for any two given vertices i, j the probability p that i is adjacent to j is equal to $\frac{4\log(n)}{n^f}$.

We introduce three random variables: X, Y and Z such that:

- X counts the number of of independent sets of size n^f in G_f^n
- Y counts the number of induced subgraphs of G_f^n isomorphic to H
- Z counts the number of cliques of size n^f in G_f^n

Using the inequality $(1-x) \le e^{-x}$ we get:

$$EX = \binom{n}{n^f}(1-p)^{\binom{n^f}{2}} \le n^{n^f} e^{-p\frac{n^f(n^f-1)}{2}} \tag{6.1}$$

Therefore

$$EX \le e^{n^f \log(n) - p\frac{n^f(n^f-1)}{2}} = o(1) \tag{6.2}$$

because of the choice of p.

We also have:

$$EZ = \binom{n}{n^f} p^{\binom{n^f}{2}} \tag{6.3}$$

and again we get: $EZ = o(1)$.

Finally, we have:

$$EY = O(\binom{n}{h} p^m) \leq C(\log(n))^m n^{h-fm} \tag{6.4}$$

for some constant $C > 0$.

From the assumption about f we get:

$$EY \leq C(\log(n))^m n^{1-\delta} \tag{6.5}$$

for some $\delta > 0$.

Now, using Markov's inequality we can conclude that for n large enough the probability that $X > 0$ and the probability that $Z > 0$ are both less than $\frac{1}{3}$ (we use here the fact that $EX = o(1)$ and $EZ = o(1)$). Again, using Markov's inequality we conclude that $Pr(Y > 3EY) < \frac{1}{3}$. So, from the union bound, we get that the event $E = \{X = Z = 0, Y \leq 3EY\}$ has positive probability. That is, there is some G_f^n for which $X = Z = 0$ and $Y \leq 3EY$. From each copy of H in that G_f^n we can delete one vertex. Therefore altogether we delete at most $3EY$ vertices, i.e. at most $3C(\log(n))^m n^{1-\delta}$ vertices. Denote by D_f^n the graph obtained from G_f^n by these deletions. Note that this graph has no clique of size at least n^f, no independent set of size at least n^f and no induced subgraphs isomorphic to H.

Besides it has at least $n - 3C(\log(n))^m n^{1-\delta}$ vertices. We can take n big enough such that $|D_f^n|^{f+\epsilon} \geq n^f$. Therefore D_f^n has no cliques or stable sets of size at least $|D_f^n|^{f+\epsilon}$ and no induced subgraphs isomorphic to H. That completes the proof of Theorem 6.1.1, since $f + \epsilon = f^*$. ∎

Now we prove Theorem 6.2.1, which we restate:

6.2.4 *The vertices of every tournament H can be ordered in such a way that the set of backward edges of every subtournament S of H has size at most $\frac{|S|-1}{\xi(H)}$.*

Proof. We use the probabilistic argument again. Fix some $0 < f^* < 1$. Assume that for every ordering o of the vertices of a tournament H there exists a subtournament S of H such that $|b_{o,S}| > \frac{|S|-1}{f^*}$, where $b_{o,S}$ is the set of backward edges of the subtournament S in the ordering o. All we need to prove then is that for all n large enough there exists a n-vertex tournament without transitive subtournaments of size n^{f^*} and without copies of H as subtournaments. Our n-vertex tournament will be constructed randomly. From what we said so far we know that exists $0 < f < 1$, $\epsilon > 0$ such that $f^* = f + \epsilon$ and for every ordering o of the vertices of a tournament H there exists a subtournament S of H such that $|b_{o,S}| > \frac{|S|-1}{f}$, where $b_{o,S}$ is the set of backward edges of the subtournament S in the ordering o. Take a set of n vertices $\{1, 2, ..., n\}$. For each pair $\{i, j\}$, where $i, j \in \{1, 2, ..., n\}$ and $i < j$ we choose independently at random the edge (i, j) with probability p

and (j, i) with probability $(1-p)$, where $p = \frac{4\log(n)}{n^f}$. As a result we obtain some random tournament T_f^n.

We introduce two random variables: X and Y such that:

- X counts the number of transitive subtournaments of size n^f in T_f^n
- Y counts the number of subtournaments H_S of at most $|H|$ vertices of T_f^n with the following property: if the vertices of H_S are ordered according to decreasing index, namely: $v_1 > v_2 > ... > v_{|H_S|}$, then this ordering induces more than $\frac{|H_S|-1}{f}$ backward edges.

Our first goal is to bound EX. We need the following lemma.

6.2.5 *For the probabilistic model introduced above the following is true for n large enough:*

$$EX \leq n^{n^f} e^{-p\binom{n^f}{2}}$$

Proof. We have

$$EX = \binom{n}{n^f} P, \tag{6.6}$$

where P is defined as a probability that the set of vertices $\{1, 2, ..., m\}$, where $m = n^f$ induces transitive subtournament. It is enough to show that

$$P \leq (n^f)! e^{-p\binom{m}{2}}, \tag{6.7}$$

since $\binom{n}{n^f}(n^f)! \leq n^{n^f}$.

Note that from the definition of the introduced probabilistic model we immediately have:

$$P \leq (n^f)! P_b, \tag{6.8}$$

where P_b is the probability that the set $M = \{1, 2, ..., m\}$ induces a transitive subtournament such that m is an inneighbour of all other vertices in M, $m-1$ is an inneighbour of all other vertices in M but m, etc. This configuration is clearly the most likely configuration giving the transitive subtournament on the set M. The total probability P differs from the probability of that particular configuration at most by the factor $(n^f)!$. In the Appendix to this chapter exact closed expression on P is given. This result is of its own interest, we will not prove it here since a very rough estimation is all that we need to finish the proof of Lemma 6.2.5.

Since we have $P_b = (1-p)^{\binom{m}{2}}$ and $(1-x) \leq e^{-x}$, we get inequality 6.7 and that completes the proof of Lemma 6.2.5. ∎

Note that Lemma 6.2.5 immediately implies that $EX \to 0$ as $n \to \infty$, since

$$EX \leq e^{n^f \log(n)} e^{-p\frac{n^f(n^f-1)}{2}} e^{\log(\frac{1-p}{1-2p})n^f} = o(1). \tag{6.9}$$

Now we need to calculate EY.

Clearly, we have for some $\epsilon > 0$:

$$EY = O(\sum_{h=1}^{|H|} \binom{n}{h} p^{\frac{h-1}{f}+\epsilon}), \tag{6.10}$$

because every edge of the form: (v_j, v_i) for $j > i$ is chosen with probability p.

Therefore we have

$$EY \leq C \sum_{h=1}^{|H|} n^h \frac{\log^{\frac{h-1}{f}+\epsilon}(n)}{n^{h-1+\epsilon f}} \tag{6.11}$$

for some constant $C > 0$.

But this means that

$$EY = O(n^{1-\eta}) \tag{6.12}$$

for some $\eta > 0$.

Now note that the condition $Y = 0$ implies that there are no copies of H in the random n-vertex tournament. To see why this is true assume by contradiciton that H appears somewhere in a random tournament. Assume that tournament H is induced by the following vertices of a random tournament: $v_1, v_2, ..., v_{|H|}$, where $v_1 > v_2 > ... > v_{|H|}$. Then we can denote by o^* the ordering of vertices $\{1, 2, ..., |H|\}$

of H that is induced by the ordering $\{v_1, ..., v_{|H|}\}$. Parameter f was chosen in such a way that for every ordering o of vertices of H there exists a subtournament S of H such that $|b_{o,S}| > \frac{|S|-1}{f}$. So now we can take such a subtournament S^* for o^* and from the fact that $|b_{o^*,S^*}| > \frac{|S^*|-1}{f}$ and the definition of Y, we obtain: $Y > 0$.

Using Markov's inequality we see that $Pr(Y \geq 3EY) \leq \frac{1}{3}$. Besides $Pr(X > 0) \leq E(X) \leq \frac{1}{3}$ for n large enough, since $EX \rightarrow 0$ as $n \rightarrow \infty$. So using the union bound, we see that with probability at least $\frac{1}{3}$ random n-vertex tournament T_f^n, constructed according to our probabilistic model, has at most $s = Cn^{1-\eta}$ subtournaments isomorphic to H (for some constant C) and no transitive subtournaments of size n^f. We can denote by D_f^n the tournament obtained from T_f^n by deleting one vertex from every subtournament isomorphic to H. So with probability at least $\frac{1}{3}$ tournament D_f^n is created by deleting at most s vertices, has no subtournaments isomorphic to H and has no transitive subtournaments of size n^f. So for n large enough there exists (n-s)-vertex tournament D that has no subtournaments isomorphic to H and has no transitive subtournaments of size n^f. We can take n sufficiently large such that $(n-s)^{f+\epsilon} > n^f$. So we know that for every $\epsilon > 0$ there exists tournament D that has no subtournaments isomorphic to H and has no transitive subtournaments of size $|D|^{f+\epsilon}$. That completes the proof of Theorem 6.2.1, since $f + \epsilon = f^*$. ∎

6.2.3 Properties of tournaments with large EH suprema

Below we give some applications of Theorem 6.2.1. In particular we prove Theorem 6.1.2. But first, we prove Corollary 6.1.3.

Proof. By Theorem 6.2.1 we know that there exists an ordering o of the vertices of a tournament H such that: $\forall_{S \subseteq V(H)} b_{o,S} < \frac{|S|-1}{c}$. Take an undirected graph G with the set of vertices V(H) induced by backward edges of H under an ordering o. It suffices to prove that $\chi(G) < \lfloor \frac{2}{c} \rfloor + 1$. Take any induced subgraph of G with the set of vertices S. Such a graph has fewer than $\frac{|S|-1}{c}$ edges, so it has a vertex of degree less than $\frac{2}{c}\frac{|S|-1}{|S|}$. So *the coloring number of* G ([**28**]) is less than $\lfloor \frac{2}{c} \rfloor + 1$. Thus $\chi(G) < \lfloor \frac{2}{c} \rfloor + 1$. ∎

Now we will prove Theorem 6.1.2.

Proof. Take a random tournament T on k vertices where for every two vertices u, v an edge uv is chosen independently at random with probability $\frac{1}{2}$. Denote by P the probability that such a random tournament has an ordering σ under which the set of backward edges is of size less than $\frac{1}{2}\binom{k}{2} - a$, where $a = C\binom{k}{2}\frac{\sqrt{\log(k)}}{\sqrt{k}}$ for some sufficiently big constant $C > 0$. Using the classic Chernoff's bound and the union bound we get: $P \leq k!e^{-\frac{a^2}{2\binom{k}{2}}}$. Now, using the formula for a we get that $P \rightarrow 0$ as $k \rightarrow \infty$. So almost all tournaments on k vertices contain at least $\frac{1}{2}\binom{k}{2} - a$ backward edges under each ordering of vertices. But that, according to

Theorem 6.2.1, completes the proof. ∎

We conclude with a few explicit constructions of the families of tournaments with the Erdös-Hajnal supremum of order $O(\frac{1}{n})$, where n is the order of the tournament.

One of the explicit constructions of families of tournaments with the Erdös-Hajnal supremum around $\frac{4}{n}$, where n is the order of the tournament, is the family of so-called *quadratic residue tournaments* ([4], p.106-109). A quadratic residue tournament H_p of order p, for p being a prime number of the form $4k+3$, is a tournament where (i, j) is an edge if $(i - j)$ is a quadratic residue modulo p. It can be shown that a quadratic residue tournament H_p has at least $\frac{1}{2}\binom{p}{2} - O(p^{\frac{3}{2}} \log(p))$ backedges under every ordering of its vertices. So according to what we have proven so far, we obtain: $\xi(H_p) \leq \frac{4}{p}(1 + \frac{c \log(p)}{\sqrt{p}})$ for some constant c.

In fact there exist much easier constructions of tournaments H with $\xi(H) = O(\frac{1}{|H|})$ than quadratic residue tournaments. We will present one of them below. First we need to note that

6.2.6 *If $\Delta(H)$ denotes the maximal number of edge disjoint directed triangles of the tournament H then $\xi(H) \leq \frac{|H|-1}{\Delta(H)}$.*

The remark is an immediate consequence of Theorem 6.2.1 and the fact that under every ordering each directed triangle induces at least one backward edge.

Define the following family of tournaments

For a fixed $k \in N$ a tournament H belongs to the family $B(k)$ if:

- H consists of three pairwise disjoint sets of vertices: A,B,C such that: $|A| = |B| = |C| = k$
- every vertex in B is an outneighbour of every vertex in A
- every vertex in C is an outneighbour of every vertex in B
- every vertex in A is an outneighbour of every vertex in C.

6.2.7 *For $H \in B(k)$ we have $\xi(H) \leq \frac{3k-1}{k^2}$*

Proof. We will use remark 6.2.6. It suffices to prove that if $H \in B(k)$ then $\Delta(H) \geq k^2$. So we only need to find a family S of at least k^2 edge-disjoint directed triangles in H. But this is easy. From the definition of H we know that there are k edge-disjoint perfect matchings: $M_1, M_2, ..., M_k$ that match vertices in B and C. We construct now the family S. Denote $A = \{a_1, a_2, ..., a_k\}$. For every $a_i \in A$, where $i = 1, 2, ..., k$ we take matching M_i, obtaining exactly k edge-disjoint triangles of the form: $((a_i, b_j), (b_j, c_j), (c_j, a_i))$ for $(b_j, c_j) \in M_i$ and $j = 1, 2, ..., k$. Altogether we obtain k^2 triangles which are edge-disjoint. Those triangles establish the family S. ∎

6.2.4 Directed triangles in tournaments with large EH suprema

In this section we connect the Erdös-Hajnal supremum of the tournament H with the number of its directed triangles. Below we prove Theorem 6.2.2.

Proof. Take the ordering from Theorem 6.2.1. Denote by b the number of backward edges in such an ordering. We know that b satisfies an inequality

$$\xi(H) \leq \frac{n-1}{b} \tag{6.13}$$

Therefore it suffices to prove that

$$b \geq \max\left\{\frac{t(H)}{2(n-2)}, \frac{4t(H)}{\sqrt{(2n-5)^2+8t(H)}+(2n-5)}\right\} \tag{6.14}$$

On the one hand we have

$$t(H) \leq \binom{b}{2} + b(n-2), \tag{6.15}$$

where the last inequality follows directly from counting directed triangles of the tournament using given ordering and the set of backward edges (every directed triangle uses either two backward edges and one forward edge or two forward edges and one backward edge from this ordering).

The expression $\binom{b}{2}$ came from counting number of directed triangles using exactly two backward edges. The expression: $b(n-2)$ is an upper bound on the number of triangles using exactly one backward and two forward edges. But on the other hand for every single backward edge we can choose at most $(n-2)$ other backward edges to form a directed triangle of two backward edges. Therefore we also have

$$t(H) \leq 2b(n-2), \tag{6.16}$$

Using both inequality 6.15 and 6.16 and choosing bigger lower bound on b we get our desired lower bound 6.14. ∎

Using the bound: $\xi(H) \leq \frac{2(n-1)(n-2)}{t(H)}$ we obtain the following corollary

6.2.8 *Every δ-dense tournament H of order n for $\delta > 0$ satisfies: $\xi(H) \leq \frac{2}{\delta n}$.*

So the family of δ-dense tournaments is the family for which we obtain the bound on the Erdös-Hajnal supremum of the same order as for a random h-vertex tournament, namely: $O(\frac{1}{h})$.

6.3 EH-suprema of tournaments without large nontrivial homogeneous sets

6.3.1 Partition number and k-modular partitions

All results of this section can be found in [10].

We will give now explicit upper bounds for EH-suprema of tournaments without large nontrivial homogeneous sets, in particular for prime tournaments. Let us first describe the main results of this section. We prove here Theorem 6.1.4, which we restate:

6.3.1 *There exists $C > 0$ such that every prime h-vertex tournament H satisfies $\xi(H) \leq C\frac{\log(h)}{h}$.*

Note that this bound is worse from the bound obtained in previous sections only by a logarithmic factor, however it can be applied to much wider families of tournaments, some with very nonrandom properties.

For a tournament T with $|V(T)| > 1$ a *k-modular partition* is a partition of $V(T)$ into k nonempty pairwise disjoint parts $\{V_1, ..., V_k\}$ such that for every $1 \leq i < j \leq k$ V_i is complete to V_j or from V_j. Note that $\{V_1, V_2, ..., V_k\}$ is in fact a partition of $V(T)$ into k homogeneous sets. A *partition number* $p(T)$ of a tournament T with $|V(T)| > 1$ is the smallest $k > 1$ such that there exists a k-modular partition of

$V(T)$. For T with $|V(T)| = 1$ we define $p(T) = 1$.

Note that clearly $p(T) \leq |V(T)|$. A tournament T is prime if and only if $p(T) = |V(T)|$.

Our next result of this section combines partition numbers and EH-suprema.

6.3.2 *There exists a constant $C > e$ such that if $\phi(x) : [e^e, \infty] \to (0, \infty)$ is a function defined as $\phi(x) = \frac{\log(\log(x))}{\log(x)}$, then for any tournament H with $\xi(H) > 0$ the following holds:*

$$p(H) \leq \lfloor \phi^{-1}(\frac{\xi(H)}{C}) \rfloor,$$

where ϕ^{-1} is the inverse of ϕ.

(Note that ϕ decreases on $[e^e, \infty]$ and $\phi(e^e) > \frac{1}{C}$. Thus $\phi^{-1}(\frac{\xi(H)}{C})$ is well defined.)

Therefore the bigger EH-supremum of a tournament, the more structured a tournament is (since there exists a k-modular partition of its vertices consisting of fewer parts). On the other hand it can be easily proven that a random h-vertex tournament H with high probability satisfies $p(H) = h$.

We have just proven that almost all labeled h-vertex tournaments have very small EH-suprema, namely of order $O(\frac{1}{h})$. In this section we give upper bounds on the sizes of families of tournaments characterized by big EH-suprema. These bounds cannot be derived using methods introduced by us earlier. Denote by $C_{\epsilon}(h)$ the number of all labeled h-vertex tournaments H with $\epsilon(H) \geq \epsilon$. We prove that:

6.3.3 *There exists a constant $C > e$ such that if $\phi(x) : [e^e, \infty) \rightarrow (0, \infty)$ is a function defined as $\phi(x) = \frac{\log(\log(x))}{\log(x)}$, then for any $1 \geq \epsilon > 0$ we have:*

$$h! \leq C_\epsilon(h) \leq h!(2\pi)^{\frac{k}{2}} 2^{(h-1)\binom{k}{2}} e^{-k(h+1+\log(2))}(h+1)^{k(h+\frac{3}{2})}(1+o(1)),$$

where $k = \lfloor \phi^{-1}(\frac{\epsilon}{C}) \rfloor$.

We already know that $C_\epsilon(h) = o(2^{\binom{h}{2}})$. However using the upper bound given above and Stirling's formula we see that in fact $C_\epsilon(h) = o(2^{h\log(h)r(h)})$, where r is any function satisfying $r(h) \rightarrow \infty$ as $h \rightarrow \infty$. From the inequality $C_\epsilon(h) \geq h!$, by Stirling's formula, we also have: $C_\epsilon(h) = \omega(2^{h\log(h)})$.

In the next subsection we prove Theorem 6.1.4, Theorem 6.3.2 and Theorem 6.3.3.

6.3.2 Proof of Theorem 6.1.4

For $X \subseteq V(T)$, write $tr(X)$ for $tr(T|X)$. Let H be a tournament. Assume that $V(H)$ admits a k-modular partition $P = \{V_1, ..., V_k\}$. We associate with the partition P a k-vertex tournament H_P with $V(H_P) = \{v_1, v_2, ..., v_k\}$ such that for $1 \leq i < j \leq k$ vertex v_i is adjacent to a vertex v_j in H_P if V_i is complete to V_j. We call H_P the *quotient tournament* of P. We say that a tournament T is H-far if T is H_P-free for every $k > 1$ and every k-modular partition P of H.

First we prove the following result:

6.3.4 *Let H be a tournament with at least two vertices. Assume that T is H-far. Then*

$$\xi(H) \leq \frac{\log(tr(T))}{\log(|T|)}.$$

Proof. Denote $V(T) = \{1, 2, ..., |T|\}$. Consider a family of tournaments $\{F_0, F_1, ...\}$ defined in the following recursive way. A tournament F_0 is just a single vertex. For $i > 0$ a tournament F_i is defined as follows. $V(F_i) = P_1^i \bigcup P_2^i \bigcup ... P_{|T|}^i$, where each P_j^i for $j = 1, 2, ..., |T|$ induces a tournament isomorphic to F_{i-1} and besides for any two $1 \leq j_1 < j_2 \leq |T|$ the set $P_{j_1}^i$ is complete to the set $P_{j_2}^i$ if j_1 is adjacent to j_2 in T and complete from the set $P_{j_2}^i$ if j_1 is adjacent from j_2. Note first that every F_i is H-free. To see this we use induction on i. For $i = 0$ this is trivial. Now take tournament F_{i+1}. If F_{i+1} is not H-free, then since T is H-far and F_i is H-free, we can conclude that $V(H)$ has k-modular partition for some $1 < k < p(H)$. That contradicts definition of $p(H)$. Knowing that every F_i is H-free we calculate the size of the biggest transitive subtournament of F_i. For $i = 0$ we have $tr(F_i) = 1$. Assume that $i > 0$. Let Tr_i be the biggest transitive subtournament of F_i. Write $S_j = V(Tr_i) \bigcap P_j^i$ for $j = 1, 2, ..., |T|$. Assume that $\{S_{j_1}, ..., S_{j_k}\}$ is the set of nonempty sets S_j. Note that the subtournament of T induced by the set $\{j_1, ..., j_k\}$ must be transitive. Otherwise, according to the definition of the family $\{F_j\}_{j=0,1,2,...}$, we conclude that Tr_i contains vertices inducing directed triangle (that contradicts the fact that Tr_i is transitive). Therefore we must have $k \leq tr(T)$. Since $S_j \subseteq P_j^i$ we must have $|S_j| \leq tr(F_{i-1})$. Therefore we have $|V(Tr_i)| = tr(F_i) \leq tr(T)tr(F_{i-1})$. So by induction, $tr(F_i) \leq tr(T)^i$. In fact from our analysis we easily see that we have $tr(F_i) = tr(T)^i$. We also have $|V(F_i)| = |T|^i$. Therefore we have $tr(F_i) = |F_i|^{\log_{|T|}(tr(T))}$. So we have $tr(F_i) = |F_i|^{\frac{\log(tr(T))}{\log(|T|)}}$. We conclude that each F_i is H-free and does not contain transitive subtournaments of size at least $|F_i|^{\epsilon}$, where $\epsilon = \frac{\log(tr(T))}{\log(|T|)}$. This implies that $\xi(H) \leq \epsilon$. ∎

We are now ready to prove Theorem 6.1.4 and Theorem 6.3.2. We encapsulate them both in the following statement:

6.3.5 *There exists $C > 0$ such that every h-vertex tournament H satisfies $\xi(H) \leq C\frac{\log(\log(p(H)))}{\log(p(H))}$. Furthermore, if $p(H) = h$ then $\xi(H) \leq C\frac{\log(h)}{h}$ for some universal constant $C > 0$.*

Theorem 6.1.4 follows from Theorem 6.3.5 since every prime tournament H satisfies $p(H) = |H|$.

Proof of Theorem 6.3.5. We may assume that $p(H)$ is large enough since for every tournament H we trivially have: $\xi(H) \leq 1$. Let G be a n-vertex tournament, where for any two vertices $1 \leq i < j \leq n$ an edge (i, j) is chosen with probability $\frac{1}{2}$. Let c be some large constant. Denote by X the number of transitive subtournaments of G of size at least $c\log(n)$ and by Y the number of copies in G of subtournaments isomorphic to some H_P, where P is some k-modular partition of H. Write $r = c\log(n)$. Note that we have $EX \leq r!\binom{n}{r}(\frac{1}{2})^{\binom{r}{2}}$. Therefore $EX \leq e^{r\log(n) - \frac{r(r-1)}{2}\log(2)}$. Taking c large enough we have $EX < \frac{1}{3}$. Assume first that $p(H) = h$. Note that in this case there is a unique H_P and it is isomorphic to H. Write $n = e^{dh}$, where $d > 0$ is a small enough constant. We have: $EY \leq \binom{n}{h}h!2^{-\binom{h}{2}} \leq n^h 2^{-\binom{h}{2}} < \frac{1}{3}$ for d small enough. Therefore for c large enough and d small enough we have: $EX < \frac{1}{3}$ and $EY < \frac{1}{3}$. Thus, using Markov's inequality, we conclude that with probability less than $\frac{1}{3}$ we have $Y \geq 1$ and with probability less than $\frac{1}{3}$ we have $X \geq 1$. So from the union bound we know that with probability bigger than $\frac{1}{3}$ we have $X < 1$ and $Y < 1$. So there exists a tournament G that is H-far and does not contain transitvie subtournaments of size $c\log(n)$. Since we have $n = e^{p(H)}$, using Theorem 6.3.4, we immediately obtain Theorem 6.1.4. In the general case when the condition $p(H) = h$ is not necessarily satisfied, we use the same analysis. The only difference is the choice of n. Let $n = p(H) - 1$. In this scenario Y is trivially 0 since every H_P has at least $p(H)$ vertices so cannot be contained in the tournament on $p(H) - 1$ vertices. The rest of the proof is exactly the same as in the case when $p(H) = h$. That completes the proof of Theorem 6.3.2. ∎

Let us make the following remark. Celebrities satisfy Conjecture 1.2.2 in the strongest, linear sense. In [7] it has been proven that every celebrity H has an ordering of vertices such that the set of backward edges forms a forest. The converse is not true. There do exist tournaments that are not celebrities but have an ordering of vertices under which the set of backward edges forms a forest. However all the examples of h-vertex tournaments with EH-suprema of order $O(\frac{1}{h})$ given in the first part of this chapter involved tournaments with qudratic number of backward edges under every ordering. Therefore the following question is natural: does the existence of an ordering with the set of backward edges forming a forest imply big EH-suprema ? Theorem 6.1.4 shows that this is not the case since there are many examples of prime tournaments having ordering under which the set of backward edges forms a forest (one of them is a long enough directed path).

Now we prove Theorem 6.3.3.

Proof. In the context of this proof the partition of a given integer h into k parts is a set of integers $\{h_1, ..., h_k\}$ such that $h_1, ..., h_k \geq 0$ and $h_1 + ... + h_k = h$. We call such a partition a *valid partition* if in addition there exist $1 \leq i < j \leq k$ with $h_i, h_j > 0$. We use the following notation for a valid partition: $< h_1, ..., h_k >$. Fix some $0 < \epsilon \leq 1$. Write $g(h) = h!e^{t(h)}$, where $t(h) = (h-1)\binom{k}{2}\log(2) + k\log((h+1)!) - k\log(2)$ for $h \geq 1$ and $t(0) = 0$. Note first that every transitive tournament H on h vertices satisfies $\xi(H) = 1$. Therefore, since the number of all labeled h-vertex transitive tournaments is equal to the number of orderings of the set $\{1, 2, ..., h\}$, we have: $C_\epsilon(h) \geq h!$. We will prove now that we also have: $C_\epsilon(h) \leq g(h)$. This will be done by induction on h. For $h = 0, 1$ the inequality is trivial. Thus we can assume that

$h > 1$. The following is true:

$$C_\epsilon(h) \leq \sum_{<h_1,...,h_k>} 2^{\binom{k}{2}} \binom{h}{h_1} \binom{h-h_1}{h_2} ... \binom{h-h_1-h_2-...-h_{k-1}}{h_k} C_\epsilon(h_1)...C_\epsilon(h_k),$$

where the sum goes over all possible valid partitions of h into k parts. To see why the inequality above is true note that $V(H)$ has k_1-modular partition for some $1 < k_1 \leq k$. This comes from Theorem 6.3.5. Each h_i corresponds to the size of one of the parts of the partition of $V(H)$, where we allow empty parts. The part of the fixed size h_1 can be chosen in $\binom{h}{h_1}$ ways, then next one given the first one in $\binom{h-h_1}{h_2}$ ways, etc. Finally, given all the parts of the modular partition, the type of the connection between any two of them P_i, P_j may be chosen in 2 ways (either P_i is complete to or from P_j). Note also that at least two parts must be nonempty because $k_1 > 1$. That is why we assume that there exist $1 \leq i < j \leq h$ such that $h_i, h_j > 0$. In the following calculations we omit index ϵ since we use only one ϵ in the whole proof. We have:

$$C(h) \leq \sum_{<h_1,...,h_k>} 2^{\binom{k}{2}} \frac{h!}{h_1!...h_k!} C(h_1)...C(h_k).$$

Therefore we have:

$$\frac{C(h)}{h!} \leq \sum_{<h_1,...,h_k>} 2^{\binom{k}{2}} \frac{C(h_1)}{h_1!} ... \frac{C(h_k)}{h_k!}.$$

Note that $h_1, ..., h_k < h$. From the induction hypothesis we have:

$$\sum_{<h_1,...,h_k>} 2^{\binom{k}{2}} \frac{C(h_1)}{h_1!} ... \frac{C(h_k)}{h_k!} \leq \sum_{<h_1,...,h_k>} 2^{\binom{k}{2}} e^{t(h_1)+...+t(h_k)}.$$

Thus it suffices to prove that:

$$e^{t(h)} \geq \sum_{<h_1,...,h_k>} 2^{\binom{k}{2}} e^{t(h_1)+...+t(h_k)}.$$

The number of valid partitions $< h_1, ..., h_k >$ is bounded by $\binom{h+1}{k}$. To see that, take a sequence of h elements and cut it in k places. The place for every cut can be chosen in at most $h+1$ ways. Clearly the number of ways we can do such a k-cut is an upper bound for the number of partitions $< h_1, ..., h_k >$. Denote by M the maximum over all valid partitions $< h_1, ..., h_k >$ of the expression $e^{t(h_1)+...+t(h_k)}$. We have

$$\sum_{<h_1,...,h_k>} 2^{\binom{k}{2}} e^{t(h_1)+...+t(h_k)} \leq \binom{h+1}{k} M 2^{\binom{k}{2}},$$

so

$$\sum_{<h_1,...,h_k>} 2^{\binom{k}{2}} e^{t(h_1)+...+t(h_k)} \leq (h+1)^k M 2^{\binom{k}{2}}.$$

Therefore it suffices to prove that $e^{t(h)} \geq e^{\log(M)+k\log(h+1)+\binom{k}{2}\log(2)}$, i.e. that:

$$t(h) \geq \log(M) + k\log(h+1) + \binom{k}{2}\log(2) \tag{6.17}$$

Note first that:

6.3.6 $M \leq e^{t(h-1)}$.

Proof. We need one more definition. We say that the sequence $(t(0), ..., t(h))$ is *strictly convex* if

$$t(1) - t(0) < t(2) - t(1) < ... < t(h) - t(h-1)$$

. We need to prove that the maximum over all valid partitions $< h_1, ..., h_k >$ of the expression $t(h_1) + ... + t(h_k)$ is $t(h-1) + t(1)$ (since $t(1) = 0$). Denote this maximum by H. Note that $t(i) - t(i-1) = \binom{k}{2} \log(2) + k \log(i+1)$ for $i = 2, ..., h$. Therefore $t(1) - t(0) < t(2) - t(1) < ... < t(h) - t(h-1)$. Thus $(t(0), t(1), ..., t(h))$ is strictly convex. Denote by $< h_1^{opt}, ..., h_k^{opt} >$ the valid partition for which $t(h_1^{opt}) + ... + t(h_k^{opt}) = H$. We will prove that in every such partition exactly two h_i^{opt}, h_j^{opt} are nonzero and besides one of them is 1 and one is $h-1$. Assume this is not the case. Let assume first that there are three nonzero elements and without loss of generality assume that these are $0 < h_1^{opt} \leq h_2^{opt} \leq h_3^{opt}$. But then we may replace h_1^{opt} by $h_1^{opt} - 1$ and h_3^{opt} by $h_3^{opt} + 1$ to obtain another valid partition. Denote this partition by $< h_1^{'}, ..., h_k^{'} >$. From the strict convexity property of the sequence $(t(1), ..., t(k))$ we have $t(h_1^{'}) + ... + t(h_k^{'}) > t(h_1^{opt}) + ... + t(h_k^{opt})$ which contradicts definition of $(h_1^{opt}, ..., h_k^{opt})$. Thus we may assume that there are exactly two nonzero elements in $< h_1^{opt}, ..., h_k^{opt} >$. Without loss of generality we may assume that $0 < h_1^{opt} \leq h_2^{opt}$. Assume by contradiction that $h_1^{opt} > 1$. But then we may replace h_1^{opt} by $h_1^{opt} - 1$ and h_2^{opt} by $h_2^{opt} + 1$ getting a contradiction as in the previous case. That completes the proof of Lemma 6.3.6. ∎

Note that we have: $t(h) - t(h-1) = \binom{k}{2}\log(2) + k\log(h+1)$. This and Lemma 6.3.6 imply 6.17. Thus we proved that

$$C_\epsilon(h) \leq h! e^{(h-1)\binom{k}{2}\log(2) + k\log((h+1)!) - k\log(2)}. \tag{6.18}$$

To finish the proof of Theorem 6.3.3 it is enough to use inequality 6.18 and standard Stirling's formula therefore we leave it to the Reader. ∎

We prove one more structural result about tournaments H with $\xi(H) \geq \epsilon$ that may be of interest on its own. We already know that $p(H) \leq \lfloor \phi^{-1}(\frac{\epsilon}{C}) \rfloor$, where ϕ^{-1} is the inverse of $\phi(x) : [e^e, \infty] \rightarrow \infty$ defined as $\phi(x) = \frac{\log(\log(x))}{\log(x)}$ and C is some constant that does not depend on ϵ. Let $\mathcal{P}$ be a $p(H)$-modular partition $(V_1, ..., V_{p(H)})$ of the set $V(H)$ and let $H_\mathcal{P}$ be the corresponding quotient tournament. For every vertex $v_i \in V(H_\mathcal{P})$ denote by w_i the number of vertices of H that correspond to v_i. We say that $H_\mathcal{P}$ is *x-transitive* if there exists $S \subseteq V(H_\mathcal{P})$ such that $|S| > 1$, $\sum_{v_i \in S} w_i \geq x$ and S induces a transitive subtournament. Note that $H_\mathcal{P}$ is transitive if and only if it is $|V(H)|$-transitive.

Our next result is the following:

6.3.7 *Let H be a tournament with $\xi(H) > 0$. Then $H_{\mathcal{P}}$ is $(V(H) - \frac{3}{\xi(H)}(p(H)-2))$-transitive.*

Proof. Let $(V_1, ..., V_{p(H)})$ be a $p(H)$-modular partition of $V(H)$. Let $A = \{V_1, ..., V_{p(H)}\}$. As long as there exist in A three sets V_i,V_j,V_k with $1 \leq i < j < k \leq p(H)$ such that in $H_{\mathcal{P}}$ set $\{v_i, v_j, v_k\}$ induces a directed triangle, we remove from A the smallest one. Note that when no three sets with these properties can be found in A then vertices of $H_{\mathcal{P}}$ that correspond to elements left in A induce a transitive tournament. Since we start with $|A| \geq 2$ then when we stop removing elements from A we also have $|A| \geq 2$. Thus, since we can remove from A at most $(p(H) - 2)$ elements, it suffices to prove that whenever element V_i is removed from A we have $w_i \leq \frac{3}{\xi(H)}$. Assume that this is not true. Write $r = \frac{3}{\xi(H)}$. Therefore at some stage we removed from A three sets V_a, V_b, V_c such that $|V_a|, |V_b|, |V_c| > r$ and besides we have either:

- V_a is complete to V_b, V_b is complete to V_c, V_c is complete to V_a or
- V_a is complete to V_c, V_c is complete to V_b, V_b is complete to V_a.

In both scenarios the tournament $H^{'}$ induced by $V_a \bigcup V_b \bigcup V_c$ satisfies $\xi(H^{'}) \leq \frac{3r-1}{r^2}$. That has been already proven by us (see family $B(k)$). Thus we have $\xi(H^{'}) < \xi(H)$. But on the other hand, since $H^{'}$ is a subtournament of H, we have: $\xi(H^{'}) \geq \xi(H)$, contradiction. ∎

6.4 Appendix: More precise evaluation of EX from Lemma 6.2.5

Note that in the proof of Lemma 6.2.5 we derived the following equality: $EX = \binom{n}{n^f}P$. We will now give exact closed-form expression on P.

6.4.1 *For the probabilistic model considered in Lemma 5.2.7 the following is true:*

$$P = (1-p)^{\binom{m}{2}} \frac{\prod_{i=1}^{m}(1-q^i)}{(1-q)^m},$$

where m is the size of the transitive tournament and $q = \frac{p}{1-p}$.

Proof. Denote by Σ_m the set of all permutations of the set $\{1, 2, ..., m\}$. Each possible transitive tournament induced by the set $\{1, 2, ..., m\}$ correspondence to exactly one permutation $\sigma \in \Sigma_m$. This permutation is obtained by putting at the i^{th} place in the permutation, where $i = 1, 2, ..., m$, this element v from $\{1, 2, ..., m\}$ that has (m-i) inneighbours in $\{1, 2, ..., m\}$. Denote by $\zeta : Tr_m \rightarrow \Sigma_m$ the bijection that maps every m-vertex transitive tournament to the corresponding permutation. Denote by A_t for $t \in Tr_m$ the event that a set $\{1, 2, ..., m\}$ induces transitive tournament t. Then we have

$$P = \sum_{t \in Tr_m} Pr(A_t) \tag{6.19}$$

An easy observation leads to the conclusion that

$$Pr(A_t) = (1-p)^{\binom{m}{2} - I(\zeta(t))} p^{I(\zeta(t))}, \tag{6.20}$$

where by $I(\sigma)$ we denote the number of inversions of a permutation $\sigma = \{\sigma(1), ..., \sigma(m)\}$, i.e. the number of pairs $(\sigma(i), \sigma(j))$, where $i < j$ and $\sigma(i) > \sigma(j)$. Therefore we have

$$P = \sum_{t \in \ Tr_m} (1-p)^{\binom{m}{2} - I(\zeta(t))} p^{I(\zeta(t))} \tag{6.21}$$

So we have

$$P = (1-p)^{\binom{m}{2}} \sum_{\sigma \in \Sigma_m} q^{I(\sigma)} \tag{6.22}$$

If we now introduce new denotation: $K_t = \sum_{\sigma \in \Sigma_t} q^{I(\sigma)}$ for $t = 1, 2, ...$, then we have the following recursive formula that can be easily checked

- $K_1 = 1$,
- $K_t = \frac{1-q^t}{1-q} K_{t-1}$, for $t > 1$

From this recursion we obtain

$$K_t = \frac{\prod_{i=1}^{t} (1-q^i)}{(1-q)^t}, \tag{6.23}$$

for $t = 1, 2, ...$

That allows us us to express P as

$$P = (1-p)^{\binom{m}{2}} \frac{\prod_{i=1}^{m}(1-q^i)}{(1-q)^m} \tag{6.24}$$

and completes the proof of Lemma 6.4.1. ∎

7

Conclusions

In this thesis we showed several results concerning tournaments characterized by forbidden substructures. All of them were motivated by the celebrated Erdös-Hajnal Conjecture. We proved the conjecture for new families of tournaments for which it was open before. In particular, we showed that the conjecture is satisfied by all tournaments on at most 5 vertices and proved the conjecture for infinitely many prime tournaments. We described all tournaments satisfying the conjecture in almost linear sense and obtained several results on upper bounds for EH-suprema.

There are still many open questions that are worth to work on.

An obvious one is to determine whether the conjecture is true or not in the most general setting.

We have already mentioned one open problem at the end of Chapter 1.

It would be also interesting to know whether we can obtain upper bounds on the EH-suprema of order $o(\frac{1}{h})$ since the best upper bounds we obtained so far (for random tournaments) were of order $O(\frac{1}{h})$.

One may also try to improve the lower bounds on EH-suprema obtained here. Even though all the proofs presented in the thesis are constructive, the methods that were used give very weak lower bounds. In a recent paper ([9]) we propose a new algorithmic proof of the fact that constellations satisfy the conjecture. This proof gives much better lower bounds since it does not use regularity lemma. It would be interesting to reduce the gap between the best upper and lower bounds we obtained for prime tournaments for which we proved the conjecture.

Another nontrivial problem is to better understand possible values of EH-suprema. In this thesis we showed that there are no EH-suprema in the range $(\frac{5}{6}, 1)$, but the following question seems to be interesting: are there tournaments H that are not pseudo-celebrities but satisfy $\xi(H) \geq \frac{1}{2}$? A related task would be to improve the bound $\frac{5}{6}$ since it is almost certainly not the best possible.

All results of this thesis concerning lower bounds on EH-suprema were obtained only in the directed scenario. One may consider answering the question whether techniques used in that scenario can be somehow translated to the undirected case.

Other nontrivial problems involve excluding families of tournaments rather than just a single tournament. Let V_1 and V_2 be two nonempty and disjoint sets of vertices. Fix directed edges going between V_1 and V_2. Denote this set of edges by E_{V_1,V_2}. Consider the set $\mathcal{T}(E_{V_1,V_2})$ of all tournaments T with $V(T) = V_1 \bigcup V_2$ and such that $E_{V_1,V_2}(T) = E_{V_1,V_2}$, where: $E_{V_1,V_2}(T)$ is the set of edges going between V_1 and V_2 in T. The following question is still open: is it true that for every set of directed edges E_{V_1,V_2} the family $\mathcal{T}(E_{V_1,V_2})$ has the Erdös-Hajnal property ?

Bibliography

[1] N. Alon, J.Pach, R. Pinchasi, R. Radoicic, and M. Sharir. Crossing patterns of semi-algebraic sets. *JCTA*, 111, 2005.

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

[3] Noga Alon and Asaf Shapira. Testing subgraphs in directed graphs. *Proceedings of the 35th annual ACM symposium on theory of computing*, 2003.

[4] Noga Alon and Joel Spencer. The probabilistic method in combinatorics. 1990.

[5] L. W. Beineke and A.J. Schwenk. On a bipartite form of the Ramsey problem. *Proc. 5th British Combin. Conf.*, 15:17–22, 1975.

[6] E. Berger, K. Choromanski, and M. Chudnovsky. Forcing Large Transitive Subtournaments. *submitted for publication*, 2012.

[7] E. Berger, K. Choromanski, M. Chudnovsky, J. Fox, M . Loebl, A. Scott, P. Seymour, and Stephan Thomassé. Tournaments and coloring. *JCTB*, 2012.

[8] B. Bollobas and A. Thomason. Hereditary and monotone properties of graphs. *The mathematics of Paul Erdös II, Algorithms and Combinatorics*, 14:70–78, 1997.

[9] K. Choromanski. Coloring tournaments with forbidden substructures. *under preparation*.

[10] K. Choromanski. EH-suprema of tournaments with no nontrivial homogeneous sets. *submitted for publication*, 2012.

[11] K. Choromanski. Upper bounds for Erdös-Hajnal coefficients of tournaments. *Journal of Graph Theory*, 2012.

[12] K. Choromanski, M. Chudnovsky, and P. Seymour. Tournaments with near-linear transitive subsets. *submitted for publication*, 2012.

[13] M. Chudnovsky, N. Robertson, P. Seymour, and R. Thomas. The strong perfect graph theorem. *Annals of Math*, 164:51–229, 2006.

[14] M. Chudnovsky and S. Safra. The Erdös-Hajnal conjecture for bull-free graphs. *JCTB*, 98:1301–1310, 2008.

[15] M. Chudnovsky and P. Seymour. Excluding Paths and Antipaths. *submitted for publication*, 2012.

[16] M. Chudnovsky and Y. Zwols. Large cliques or stable sets in graphs with no four-edge path and no five-edge path in the complement . *Journal of Graph Theory*, 70:449–472, 2012.

[17] R. Diestel. Graph Theory. *Springer*, pages 30–31, 150, 190–211, 1999.

[18] P. Erdös and A. Hajnal. Ramsey-type theorems. *Discrete Applied Mathematics*, 25:37–52, 1989.

[19] J. Fox and J. Pach. Erdös-Hajnal-type results on intersection patterns of geometric objects. *Horizon of Combinatorics (G.O.H. Katona et al., eds.), Bolyai Society Studies in Mathematics*, 2008.

[20] J. Fox, J. Pach, and Cs.D.Toth. Turan-type results for partial orders and intersection graphs of convex sets. *Israel J. Mathematics*, 178:29–50, 2010.

[21] J. Fox and B. Sudakov. Density theorems for bipartite graphs and related Ramsey-type results. *Combinatorica*, 29:153–196, 2009.

[22] Jacob Fox. *private communication*.

[23] J. H. Kim. The Ramsey number $R(3,5)$ has order of magnitude $\frac{t^2}{\log(t)}$. *Random Structures and Algorithms*, 7, 1995.

[24] D. Larman, J. Matousek, J. Pach, and J. Torocsik. A Ramsey-type result for convex sets. *Bull. London Math. Soc*, 26, 1994.

[25] M. Loebl, B. Reed, A. Scott, S. Thomassé, and A. Thomason. Almost all H-free graphs have the Erdös-Hajnal property. *An Irregular Mind (Szemerédi is 70), in Bolyai Society Mathematical Studies, Springer, Berlin, 21*, pages 405–414, 2010.

[26] Ivan Matosevic. The Probabilistic Method and the Bounds of Classical Ramsey Numbers. 2005.

[27] G. L. Miller, S.-H. Teng, W. Thurston, and S. A. Vavasis. Separators for sphere-packings and nearest neighbor graphs. *J. ACM*, 44:1–29, 1997.

[28] Michael Molloy and Bruce Reed. Graph coloring and the probabilistic method. *Algorithms and Combinatorics*, 23, 2000.

[29] D. Seinsche. On a property of the class of n-colorable graphs. *JCTB*, 16:191–193, 1974.

[30] R. Stearns. The voting problem. *Amer. Math. Monthly*, pages 761–763, 1959.