Forcing Large Transitive Subtournaments

Eli Berger [1]
Haifa University
Haifa, Israel

Krzysztof Choromanski [2]
Columbia University
New York, NY, USA

Maria Chudnovsky [3]
Columbia University
New York, NY, USA

January 25, 2012; revised August 10, 2014

[1]Partially supported by BSF grant 2006099 and ISF grant 47391
[2]Partially supported by NSF grant IIS-1117631
[3]Partially supported by NSF grants DMS-1001091 and IIS-1117631 and BSF grant 2006099

Abstract

The Erdős-Hajnal Conjecture states that for every given H there exists a constant $c(H) > 0$ such that every graph G that does not contain H as an induced subgraph contains a clique or a stable set of size at least $|V(G)|^{c(H)}$. The conjecture is still open. However some time ago its directed version was proved to be equivalent to the original one. In the directed version graphs are replaced by tournaments, and cliques and stable sets by transitive subtournaments. Both the directed and the undirected versions of the conjecture are known to be true for small graphs (or tournaments), and there are operations (the so-called *substitution operations*) allowing to build bigger graphs (or tournaments) for which the conjecture holds. In this paper we prove the conjecture for an infinite class of tournaments that is not obtained by such operations. We also show that the conjecture is satisfied by every tournament on at most 5 vertices.

Keywords: the Erdős-Hajnal Conjecture, the regularity lemma, tournaments

1 Introduction

All graphs in this paper are finite and simple. Let G be an undirected graph. The vertex set of G is denoted by $V(G)$, and the edge set by $E(G)$. We write $|G|$ to mean $|V(G)|$. Given $X \subseteq V(G)$, we denote by $G|X$ the *subgraph of G induced by X*, that is the graph with vertex set X, in which $x, y \in X$ are adjacent if and only if they are adjacent in G. For an undirected graph H, we say that G is *H-free* if no induced subgraph of G is isomorphic to H. A *clique* in G is a subset of $V(G)$ all of whose elements are pairwise adjacent, and a *stable set* in G is a subset of $V(G)$ all of whose elements are pairwise non-adjacent. For a graph H and a vertex $v \in V(H)$ we denote by $H \backslash v$ a graph obtained from H by deleting v and all edges of H that are: adjacent to v in the undirected setting and: adjacent to or from v in the directed setting.

The Erdős-Hajnal Conjecture is the following:

1.1 *For every undirected graph H there exists a constant $c(H) > 0$ such that the following holds: every H-free graph G contains a clique or a stable set of size at least $|G|^{c(H)}$.*

This conjecture is still open. We say that an undirected graph H *satisfies the Erdős-Hajnal Conjecture* (or equivalently: *has the Erdős-Hajnal property*) if there exists $c(H) > 0$ such that every H-free graph G contains either a clique or a stable set of size at least $|G|^{c(H)}$.

A version of 1.1 for some class of directed graphs was formulated in [2]. To state it, we need some definitions. A *tournament* is a directed graph T, where for every two vertices u, v exactly one of (u, v), (v, u) is an edge of T (that is, a directed edge). If $(u, v) \in E(T)$, we say that u is *adjacent to v*, and that v is *adjacent from u*. A tournament is *transitive* if it contains no directed cycle (equivalently, no directed cycle of length three). Let T be a tournament. We denote its vertex set by $V(T)$ and its edge set by $E(T)$, and write $|T|$ for $|V(T)|$. We refer to $|T|$ as the *order* of T. Given $X \subseteq V(T)$, the *subtournament of T induced by X*, denoted by $T|X$, is the tournament with vertex set X, such that for $x, y \in X$, (x, y) is a directed edge of $T|X$ if and only if $(x, y) \in E(T)$. Given a tournament S, we say that T *contains* S if S is isomorphic to $T|X$ for some $X \subseteq V(T)$. If T does not contains S, we say that T is *S-free.* The conjecture from [2] is the following.

1.2 *For every tournament S there exists a constant $c(S) > 0$ such that the following holds: every S-free tournament T contains a transitive subtournament of order at least $|T|^{c(S)}$.*

It was also shown in [2] that 1.1 and 1.2 are equivalent to each other. We define what it means for a tournament to satisfy the Erdős-Hajnal Conjecture in a way similar to the that of undirected graphs. In other words, a tournament H *satisfies the Erdős-Hajnal Conjecture* if there exists $c(H) > 0$ such that every H-free tournament G contains a transitive set of size at least $|G|^{c(H)}$.

Both the directed and the undirected versions of the conjecture are known to be true for small graphs (or tournaments). There is also an operation, called the *substitution operation*, allowing to build bigger graphs (or tournaments) satisfying the Erdős-Hajnal Conjecture from smaller ones. We will define now the substitution operation. Let H_1, H_2 be two undirected graphs/tournaments with disjoint vertex sets. Assume that each H_i has at least two vertices. Let $w \in V(H_1)$. Then we say that H is obtained from H_1 by *substituting H_2 for w* if:

- $V(H) = (V(H_1) \cup V(H_2)) \backslash \{w\}$,
- $H|V(H_2) = H_2$,
- $H|(V(H_1) \backslash \{w\}) = H_1 \backslash w$ and,
- $v \in V(H_1)$ is adjacent in H to $u \in V(H_2)$ if and only if v is adjacent in H_1 to w.

It was proved in [2] (Theorem 2.1) that if H_1 and H_2 both satisfy the Erdős-Hajnal Conjecture then so does H. The proof was presented for the undirected setting however its version for the directed setting is completely analogous and therefore we leave it to the reader.

Let T be a tournament, and let $(v_1, \ldots, v_{|T|})$ be an ordering of its vertices; denote this ordering by θ. We say that an edge (v_j, v_i) of T is a *backward edge* under this ordering if $i < j$. The *graph of backward edges* under this ordering, denoted by $B(T, \theta)$, has vertex set $V(T)$, and $v_i v_j \in E(B(T, \theta))$ 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 $l_1, \ldots, l_t$. 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.

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. Finally, a *star* in $B(T, \theta)$, is a left star or a right star.

A tournament T is a *galaxy* if there exists an ordering θ of its vertices such that every connected component of $B(T, \theta)$ is either a star or a singleton, and

- no center of a star appears in the ordering between two leaves of another star.

We call such an ordering a *galaxy 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 theta*. If $V(T) = \bigcup_{i=1}^{l} V(\Sigma_l)$, we say that T is a *regular* galaxy.

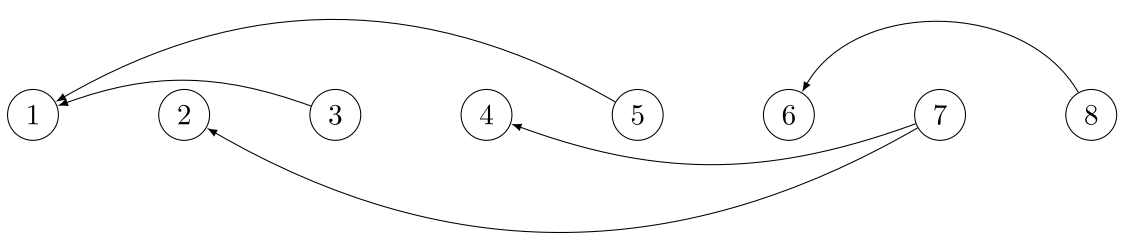

Fig.1 Galaxy consisting of one left and two right stars. All edges that are not drawn are forward.

Our first result in this paper is the following:

1.3 *Every galaxy H satisfies the Erdős-Hajnal Conjecture.*

We denote by P_k a tournament of order k whose vertices can be ordered so that the set of backward edges under this ordering is the set of edges joining consecutive vertices in the ordering. Thus formally, the vertices of P_k can be enumerated as: $v_1, ..., v_k$ such that the set of backward edges under this ordering is of the form: $\{(v_{i+1}, v_i) : i = 1, ..., k-1\}$.

As an interesting fact, we prove the following corollary:

1.4 *For every k, the tournament P_k satisfies the Erdős-Hajnal Conjecture.*

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 with vertex set $\{0, 1, 2, 3, 4\}$ and i is adjacent to $i + 1 \mod 5$ and $i + 2 \mod 5$.

Our second main result is:

1.5 *The tournament C_5 satisfies the Erdős-Hajnal Conjecture.*

As a corollary, using 1.5 and [3], we get

1.6 *Every tournament on at most 5 vertices satisfies the Erdős-Hajnal Conjecture.*

This paper is organized as follows:

- in Section 2 we present some definitions and technical lemmas used in the proofs of 1.3 and 1.5.
- in Section 3 we present most important ideas used in the proofs of main results of the paper without going deeply into technical details.
- in Section 4 we prove 1.3 and deduce 1.4.
- in Section 5 we prove 1.5.
- in Section 6 we prove 1.6.

2 Basic lemmas

In this section we prove a few lemmas used in the proofs of our main results. Let T be a tournament. We say that a vertex w is an *outneighbor* of a vertex v if $(v, w) \in E(T)$. Otherwise we say that it is an *inneighbor* of a vertex v. For disjoint subsets A, B of $V(T)$, we say that A is *complete to B* if every vertex of A is adjacent to every vertex of B. We say that A is *complete from B* if B is complete to A. Denote by $tr(T)$ the largest size of the transitive subtournament of T. For $X \subseteq V(T)$, write $tr(X)$ for $tr(T|X)$.

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: $|d(X, Y) - d(A, B)| \leq \epsilon$.

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 was proved in [1]:

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

The above lemma is a "tournament"-version of the celebrated *Regularity Lemma* proved by Endre Szeméredi and originally stated for undirected graphs ([6]). In the undirected setting we only need to change the definition of $e_{X,Y}$ which is now the number of edges between sets X and Y. The original version of the lemma is as follows:

2.2 *For every $\epsilon > 0$ and every $m \geq 1$ there exists an integer $DM = DM(m, \epsilon)$ such that every undirected graph of order 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:

2.3 *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 order 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 undirected version of the lemma above is another celebrated result, the so-called *Embedding Lemma*.

2.4 *For every natural number k and real number $0 < \lambda < 1$ there exists $0 < \eta = \eta(k, \lambda) < 1$ such that for every undirected graph H with vertex set $\{x_1, ..., x_k\}$ and undirected graph T with vertex set $V(T) = \bigcup_{i=1}^{k} V_i$, if the V_i's are disjoint sets, each of order 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 subgraph of T induced by $\{v_1, ..., v_k\}$.*

Its proof can be found in [4]. We will omit the proof of 2.3 since it is completely analogous to the proof of the Embedding Lemma.

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: $tr(S) \geq |S|^{\epsilon}$. Next we list some properties of ϵ-critical tournaments.

2.5 *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 order 2 so it suffices to take $\epsilon(N) = \log_N(2)$. ∎

2.6 *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 $|G| \geq f \cdot tr(T)$ we have: 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 $\epsilon < \log_c(1-f)$. ∎

2.7 *Let T be an ϵ-critical tournament with $|T| = n$ and $\epsilon, c > 0$ be constants such that $\epsilon < \log_{\frac{c}{2}}(\frac{1}{2})$. Then for every two disjoint subsets $X, Y \subseteq V(T)$ with $|X| \geq cn$, $|Y| \geq cn$ there exist an integer $k \geq \lfloor \frac{cn}{2} \rfloor$ and vertices $x_1, ..., x_k \in X$ and $y_1, ..., y_k \in Y$ such that y_i is adjacent to x_i for $i = 1, ..., k$.*

Proof. Assume otherwise. Write $m = \lfloor \frac{cn}{2} \rfloor$. Consider the bipartite graph G with bipartition (X, Y) where $\{x, y\} \in E(G)$ if $(y, x) \in V(T)$. Then we know that G has no matching of size m. By König's Theorem (see [5]) there exists $C \subseteq V(G)$ with $|C| < m$, such that every edge of G has an end in C. Write $C \cap X = C_X$ and $C \cap Y = C_Y$. We have $|C_X| \leq \frac{|X|}{2}$ and $|C_Y| \leq \frac{|Y|}{2}$. Therefore $|X \backslash C_X| \geq \frac{|X|}{2}$ and $|Y \backslash C_Y| \geq \frac{|Y|}{2}$, and by the definition of C and G, we know that $X \backslash C_X$ is complete to $Y \backslash C_Y$. Denote by T_1 a transitive subtournament of size $tr(T|(X \backslash C_X))$ in $T|(X \backslash C_X)$. Denote by T_2 a transitive subtournament of size $tr(T|(Y \backslash C_Y))$ in $T|(Y \backslash C_Y)$. From the ϵ-criticality of T and since $|X \backslash C_X| \geq \frac{cn}{2}$, $|Y \backslash C_Y| \geq \frac{cn}{2}$, we also have: $|T_1| \geq (\frac{cn}{2})^\epsilon$, $|T_2| \geq (\frac{cn}{2})^\epsilon$. We can merge T_1 and T_2 to obtain bigger transitive tournament T_3 with $|T_3| \geq 2(\frac{c}{2})^\epsilon n^\epsilon$. Therefore, since T is ϵ-critical, we have: $2(\frac{c}{2})^\epsilon < 1$. But this contradicts the condition $\epsilon < \log_{\frac{c}{2}}(\frac{1}{2})$.

∎

The next lemma is a starting point for all of our constructions. This is also the only step in the proof where we use 2.1. 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 2.1.

2.8 *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 tournament T not containing H 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 $R(t_1, t_2)$ denote the smallest integer such that every graph of order 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 [5]). Take $k = R(2^{P-1}, h)$. Take $\eta = \min(\frac{1}{2(k-1)}, \eta_0(h, \lambda))$ (where η_0 is as in the statement of 2.3). 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 2.1 there exists an integer $N > 0$ such that every tournament T with $|T| \geq N$ admits an η-regular partition with at least u parts. Denote by DM the upper bound (from 2.1) 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

$V(G)=\{W_1,\dots,W_r\}$ 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 (see [5]) 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 $W_1,...W_h$ 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 $W_1,...W_{2^{P-1}}$ 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 does not contain H and $\eta \leq \eta_0$, 2.3 implies that the former is impossible.

Now define $\hat{T}$ to be the 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 order at least 2^{P-1} contains a transitive subtournament of order at least P (see [7]), we conclude that $\hat{T}$ contains a transitive subtournament of order 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 order 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.

2.9 *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 2.8 further.

2.10 *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 $i,j \in \{1,2,...,P\}$, $i<j$ 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. Denote $\hat{I} = \{i_1, ..., i_{r-1}\}$. 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. We will now describe the procedure of extracting from A_{i_r} several transitive subtournaments of substantial sizes. 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 order $\lceil \frac{1}{2} n^\epsilon \rceil$. Denote this transitive subtournament by $T_1^{i_r}$. We have $|T_1^{i_r}| \geq \frac{1}{2} tr(T)$. The last inequality follows from the fact that, by ϵ-criticality, $tr(T) < n^\epsilon$ and from our previous observation regarding the lower bound on the size of $T_1^{i_r}$. We repeatedly remove transitive subtournaments: $T_1^{i_1}, T_2^{i_1}, ...$, each of size at least $\frac{1}{2} tr(T)$ as long as at least $\frac{|A_{i_r}|}{2}$ vertices remain. Denote the set of all extracted (and pairwise disjoint) transitive subtournaments as $\mathcal{W} = \{T_1^{i_r}, ..., T_w^{i_r}\}$, where w is some positive integer. We conclude that the following holds: $\bigcup_{j=1}^{w} |T_j^{i_r}| \geq \frac{|A_{i_r}|}{2}$ 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 2.9 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 2.9 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 $4P\hat{\Lambda} = (4P)^{|I|}\lambda = \Lambda$. ∎

Next 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 $|S_i| \geq c \cdot tr(T)$
- $d(S_i, S_j) \geq 1 - \lambda$ for all $1 \leq i < j \leq |w|$.

We now use 2.8 and 2.10 to prove the following:

2.11 *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.*

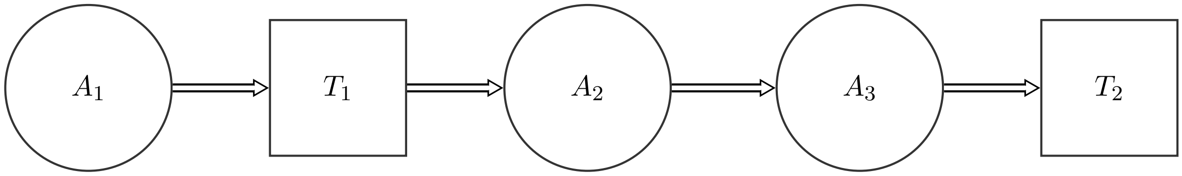

Fig.2 Schematical representation of the (c, λ, w)-structure. This structure consists of three linear sets: A_1, A_2, A_3 and two transitive sets: T_1 and T_2. The arrows indicate the orientation of most of the edges going between different elements of the (c, λ, w)-structure. Each T_i satisfies: $|T_i| \geq c \cdot tr(T)$ and each A_i satisfies: $|A_i| \geq c \cdot n$, where $n = |T|$. We have here: $w = (0, 1, 0, 0, 1)$.

Proof. Write $n = |T|$ and $w = (w_1, \ldots, w_P)$, where $P > 0$ is an integer. Define $C = |\{i : w_i = 1\}|$. Let $\Lambda = \frac{\lambda}{(4P)^C}$. By 2.5 we can choose ϵ_0 small enough such that $|T| > N$, where N is an integer from 2.8. Now it follows from 2.8 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 2.10 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 a 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 $j \in \{1, 2, ..., m\}$.

2.12 *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 will 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 2.6 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 $i \in \{1, \ldots, |w|\}$, 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.

2.13 *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 2.9 taking $X = B$, $Y = S_j$. ∎

2.14 *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 2.13 we know that $|B_j| \leq \frac{1}{M}|S_i|$. Therefore we have: $|B| \leq \frac{|w|}{M}|S_i|$. ∎

3 An overview

The goal of this section is to present the reader an overview of the key techniques that will be used to derive main results of the paper. Full proofs will be given in the subsequent sections.

The proofs use the directed version of Szemerédi's Regularity Lemma. Given a galaxy H, we start with a regular partition of a H-free tournament. Using the directed version of the embedding lemma along with a few standard techniques which we will not describe here, we can find subsets $V_{i_1}, \ldots, V_{i_t}$ (for an appropriately chosen constant t), such that $d(V_{i_p}, V_{i_q}) > .999$ for every $1 \leq p < q \leq t$. This means that for every $1 \leq p < q \leq t$, vertices of V_{i_p} tend to be adjacent to a substantial proportion of the vertices of V_{i_q}. On the other hand, if a substantial subset of V_{i_p} is complete to a substantial subset of V_{i_q}, then we can apply induction to get a large transitive subtournament in T, and so we may assume that no such subsets exist. We now construct a copy of H in T, choosing at most one vertex from each of $V_{i_1}, \ldots, V_{i_t}$, and using the fact that for $1 \leq p < q \leq t$ no substantial subset of V_{i_p} is complete to a substantial subset of V_{i_q} to obtain the backward edges in the galaxy ordering of G, thus obtaining the result of 1.3.

Obviously, every tournament obtained from a transitive tournament by adding a vertex is a galaxy. It is not difficult to check that there is only one tournament on five vertices that is not a galaxy. Here it is: its vertex set is $\{v_1, \ldots, v_5\}$, and v_iv_j is an edge if and only if $(j-i) \ mod \ 5 \in \{1, 2\}$. This is a tournament C_5. We remark that C_5 is an example of a tournament that is obtained from a transitive tournament by adding two vertices, and that is not a galaxy.

The proof that a tournament C_5 has the Erdös-Hajnal property. uses similar ideas to the ones in the proof of 1.3, but now instead of one specific ordering of vertices, two are used.

Theorem 1.3 and 1.5 together imply that every tournament on at most five vertices has the Erdös-Hajnal property. Another curious corollary of 1.3 is that P_k has the Erdös-Hajnal property. This follows from the fact that, somewhat surprisingly, P_k has a galaxy ordering.

4 Galaxies

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 $\delta^s : \{i : s_c = 1\} \rightarrow \mathbb{N}$ 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 galaxy, and let $(v_1, \dots, v_h)$ be a galaxy ordering of $V(H)$; denote this ordering by θ. 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 $w = s_c^{H,\theta}$.

Let $(S_1, S_2, ..., S_{|w|})$ be a (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 $m(i_r) = \lfloor \frac{|S_j|}{\delta^w(i_r)} \rfloor$. Denote $S_{i_r}^j = \{s_{i_r}^{(j-1)m(i_r)+1}, ..., s_{i_r}^{jm(i_r)}\}$ for $j \in \{1, 2, ..., \delta^w(i_r)\}$. For every $v \in S_{i_r}^j$ denote $\xi(v) = (|\{k < i_r : w(i) = 0\}| + \sum_{k<i_r:w(i)=1} \delta^w(k)) + j$. For every $v \in S_{i_r}$ such that $w(i_r) = 0$ denote $\xi(v) = (|\{k < i_r : w(i) = 0\}| + \sum_{k<i_r:w(i)=1} \delta^w(k)) + 1$. 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 $j \in \{1, \dots, h\}$.

Our main goal in this section is to prove 1.3. We then deduce 1.4. Let us start with one more technical lemma.

4.1 *Let H be a regular galaxy with $|H| = h$ and let θ be its galaxy-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 $(\frac{c}{(2(h+1))^{l-k}}, (2(h+1))^{2(l-k)}\lambda, w)$-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 2.5. 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 2.14 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 $i \in \{0, \dots, p\}$, define D_i to be the set of all vertices v of $\bigcup_{j=1}^{|w|} S_j$ with $\xi(v) = q_i$ that are M-good with respect to $(S_1, ..., S_{|w|})$. 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 2.12 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 of the same star, 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 $i \in \{1, ..., |w|\} \backslash \{x, y\}$. Denote $S_i^f = \bigcap_{j=0}^{p} S_{i,d_j}$. 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 t sets S_i^f, where $t = (|w| - 1)$ if S_y^f is defined and $t = (|w| - 2)$ if S_y^f is not defined. 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 2.9, 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 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 copy of H^k. This completes the proof. ∎

We need one more observation before proving 1.3.

4.2 *Every galaxy is a subtournament of a regular galaxy.*

Proof. Let H be a galaxy and let $\theta = (h_1, ..., h_{|H|})$ be its galaxy ordering. Let $\{h_{i_1}, ..., h_{i_s}\}$ for some $1 \leq i_1 < ... < i_s \leq |H|$ be the vertices-singletons. We can assume that this set is nonempty. Now let us consider tournament H^+ with vertices: $h_1, ..., h_{|H|}, h_{|H|+1}, ..., h_{|H|+s}$ such that under an ordering $\theta^+ = (h_1, ..., h_{|H|}, h_{|H|+1}, ..., h_{|H|+s})$ the backward edges are those of H under ordering θ and the edges of the form $(h_{|H|+j}, h_{i_j})$ for $j = 1, ..., s$. Ordering θ^+ is clearly the galaxy ordering of H^+. Under this ordering there are no longer singletons. Thus H^+ is a regular galaxy. Furthermore H^+ contains H as a subtournament. That completes the proof.

∎

Now we are ready to prove 1.3 that we restate

4.3 *Every galaxy H satisfies the Erdős-Hajnal Conjecture.*

Proof. Let H be a galaxy. By 4.2 we may assume that H is regular. Let θ be a galaxy-ordering of H and let $\Sigma_1, ..., \Sigma_l$ be the stars of H under θ. Let ϵ_l be as in 4.1. Suppose 4.3 is false. Then there exists an $\frac{\epsilon_l}{2}$-critical tournament T not containing H. By 2.11 T contains a $(c, \frac{1}{h^2(2h+2)^{2h+2}}, z)$-structure corresponding to H for an appropriate vector z and some constant $c > 0$. But now, by 4.1 with $k = l$ we deduce that T contains H, a contradiction. ∎

Now we prove an interesting corollary 1.4 which we restate below

4.4 *For every k the tournament P_k satisfies the Erdős-Hajnal Conjecture.*

Proof. Take a path P_k. We can assume without loss of generality that $k = 2l$ for some l. By 4.3, it is enough to prove that P_k is a galaxy. 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 galaxy and the result follows. ∎

5 The tournament C_5

In this section we prove 1.5. 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).

5.1 *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 2.7, 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 1.5 which we restate below:

5.2 *The tournament C_5 satisfies the Erdős-Hajnal Conjecture.*

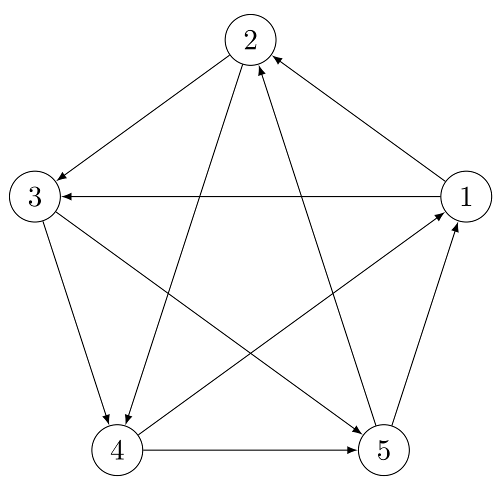

Fig.3 Tournament C_5 - the smallest tournament that is not a galaxy.

Proof. Assume otherwise. Taking $\epsilon > 0$ small enough, we may assume that there exists a C_5-free ϵ-critical tournament T. By 2.11 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 2.14 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 2.9 $(S_1^*, S_2, T_1^*, S_4, S_5^*)$ 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 2.7 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 inneighbor in T_3^*, and by Y the subset of $\{y_1, ..., y_k\}$ consisting of the vertices with an outneighbor in T_1^*.

We may assume that $\epsilon < \log_{\frac{5}{36}c}(1-\frac{c}{6})$, and thus 2.6 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 5.1 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. ∎

Fig.4 Two crucial orderings of vertices of C_5. The left one is the path ordering and the right one is the cyclic ordering. Notice that none of them is the galaxy ordering.

6 Small tournaments

Our goal in this section is to prove 1.6. First, we need some definitions. 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 [3].

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 [3].

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

We are ready to prove 1.6, which we restate.

6.2 *Every tournament on at most 5 vertices satisfies the Erdős-Hajnal Conjecture.*

Proof. Clearly every celebrity satisfies the Erdős-Hajnal Conjecture, so by 6.1 it is enough to prove the result for G_1, G_2, C_5. Since $(v_1, \ldots, v_5)$ is a galaxy-ordering of G_1, and $(w_1, \ldots, w_5)$ is a galaxy-ordering of G_2, 4.3 implies that both G_1 and G_2 satisfy the Erdős-Hajnal Conjecture, and by 1.5 so does C_5. This completes the proof. ∎

.

References

[1] Noga Alon and Asaf Shapira, Testing subgraphs in directed graphs, Proceedings of the thirty-fifth annual ACM symposium on Theory of computing (2003).

[2] Noga Alon, Janos Pach, Jozsef Solymosi, Ramsey-type theorems with forbidden subgraphs, Combinatorica, p. 155-170 (2001).

[3] Eli Berger, Krzysztof Choromanski, Maria Chudnovsky, Jacob Fox, Martin Loebl, Alex Scott, Paul Seymour, Stephan Thomassé, Tournaments and coloring, submitted for publication (2011).

[4] Béla Bollobás and Andrew Thomason, Hereditary and monotone properties of graphs, The mathematics of Paul Erdős II, Algorithms and Combinatorics, Vol. 14, Springer, Berlin, p. 7078 (1997).

[5] Reinhard Diestel, Graph Theory, Springer, p. 150, p. 190-211, p.30-31 (1999).

[6] Reinhard Diestel, Graph Theory, Springer, p. 153 (2000).

[7] R. Stearns, The voting problem, Amer. Math. Monthly 66, p. 761-763 (1959).