

Counting H -free orientations of graphs

BY MATIJA BUCIĆ

*School of Mathematics, Institute for Advanced Study, 1 Einstein Drive, Simonyi Hall 025,
Princeton, 08540, U.S.A.*

AND

*Department of Mathematics, Princeton University, Washington Road, Fine Hall 1203,
Princeton, 08540, U.S.A.*

e-mail: matija.bucic@ias.edu

OLIVER JANZER† AND BENNY SUDAKOV‡

Department of Mathematics, ETH Zurich, Ramistrasse 101, 8092 Zurich, Switzerland.

e-mails: oliver.janzer@math.ethz.ch;

benjamin.sudakov@math.ethz.ch

(Received 05 July 2021; revised 14 March 2021; accepted 02 March 2022)

Abstract

In 1974, Erdős posed the following problem. Given an oriented graph H , determine or estimate the maximum possible number of H -free orientations of an n -vertex graph. When H is a tournament, the answer was determined precisely for sufficiently large n by Alon and Yuster. In general, when the underlying undirected graph of H contains a cycle, one can obtain accurate bounds by combining an observation of Kozma and Moran with celebrated results on the number of F -free graphs. As the main contribution of the paper, we resolve all remaining cases in an asymptotic sense, thereby giving a rather complete answer to Erdős's question. Moreover, we determine the answer exactly when H is an odd cycle and n is sufficiently large, answering a question of Araújo, Botler and Mota.

2020 Mathematics Subject Classification: 05C20 (Primary); 05C35 (Secondary)

1. Introduction

Given a fixed graph H , over all n -vertex graphs G what is the maximum number of 2-edge colourings of G which contain no monochromatic copy of H ? This very natural question was first asked by Erdős and Rothschild [11] in 1974 for the special case of $H = K_3$. This case was resolved by Yuster [26] who in turn asked what happens for $H = K_k$. This problem, for large n , was solved by Alon, Balogh, Keevash and Sudakov [2] who in addition solved it for H being any edge-colour critical graph (defined as graphs in which the removal of some edge decreases the chromatic number). The question has attracted a lot of attention over the years and has been generalised in a number of ways; we point the interested reader to the numerous papers citing [2], e.g. [5, 20, 23, 24].

† Research supported by an ETH Zürich Postdoctoral Fellowship 20-1 FEL-35.

‡ Research supported in part by SNSF grant 200021_196965.

In the same paper from 1974, Erdős [11] also raised the following closely related problem. Given an oriented graph H , what is the maximum possible number of H -free orientations of an n -vertex graph? Let $D(n, H)$ denote the answer to this question. Erdős asked to determine or estimate $D(n, H)$. For an undirected graph F , let $\text{ex}(n, F)$ be the maximum number of edges in an n -vertex F -free graph. Writing F for the underlying undirected graph of H , we have a trivial lower bound $D(n, H) \geq 2^{\text{ex}(n, F)}$ since if G is an F -free graph, then any orientation of G is H -free. Alon and Yuster [3] showed that when H is a tournament, this simple lower bound gives the correct answer. That is, if T is a tournament on k vertices, then $D(n, T) = 2^{t_{k-1}(n)}$ holds for sufficiently large n , where $t_{k-1}(n)$ denotes the number of edges in the $(k-1)$ -partite Turán graph on n vertices. Their general argument, which follows the approach used in [2], relies on a regularity lemma and hence results in a requirement for n to be extremely large. For the special case of 3-vertex tournaments they give a different approach which solves the problem for the transitive tournament on three vertices for all n and only requires n to be larger than about 10000 for the directed 3-cycle C_3 (here and below, C_k denotes the directed k -cycle with the standard orientation). As an aside, we remark that the number of H -free orientations of a random graph $G = G(n, p)$ has also been studied for various choices of H , e.g. for $H = C_k$ (see [1, 10]). Recently, Araújo, Botler and Mota [4] determined $D(n, C_3)$ for all values of n and asked what happens if H is a directed cycle of arbitrary length, even if we are only interested in the case of large n . Our first result is an exact answer to their question for odd cycles.

THEOREM 1.1 *For any $k \geq 1$ there exists $n_0 = n_0(k)$ such that if $n \geq n_0$, then*

$$D(n, C_{2k+1}) = 2^{\lfloor n^2/4 \rfloor}.$$

In fact, our argument, which follows closely the ideas of both [2] and [3], applies for any H which is an orientation of an edge-colour critical graph, showing that $D(n, H) = 2^{\text{ex}(n, F)}$ for large enough n in such cases.

A natural next question is what happens for other graphs. As suggested by Erdős, obtaining an approximate understanding of the answer is already interesting. Using a short and beautiful argument involving a version of the classical Sauer–Shelah lemma on VC dimension of sets, Kozma and Moran [19] proved that the number of orientations of a fixed graph G without H is always at most the number of F -free subgraphs of G , where as usual F is the underlying graph of H . Hence, one can obtain upper bounds for $D(n, H)$ from known results on the number of n -vertex F -free graphs, which is an extensively studied subject on its own.

In the following result we trade precision for generality. It is obtained by combining the result of Kozma and Moran with that of Erdős, Frankl and Rödl [13].

PROPOSITION 1.2. *Given an oriented graph H with underlying graph F , then*

$$D(n, H) = 2^{\text{ex}(n, F) + o(n^2)}.$$

This result establishes the answer up to lower order terms for any oriented graph whose underlying graph is non-bipartite, since $\text{ex}(n, F) = \Theta(n^2)$ for any non-bipartite F . The case of bipartite underlying graphs turns out to be more difficult, mostly due to the fact that their Turán numbers are much less well understood. The next proposition relies on a result of Ferber, McKinley and Samotij [15].

PROPOSITION 1.3. *Let F be a graph containing a cycle, and assume that there exists a real number α such that $\text{ex}(n, F) = \Theta(n^\alpha)$. Then for any orientation H of F , we have*

$$D(n, H) = 2^{\Theta(n^\alpha)}.$$

Although it is generally believed that such an α exists for every bipartite graph F , this is still a wide open conjecture. For a survey of the vast literature about the extremal number of bipartite graphs, see [16].

The above results provide us with good understanding of $D(n, H)$ whenever the underlying graph of H contains a cycle. This leads to the natural question of what happens in the remaining case, namely when H is an orientation of a forest F . Since in this case $\text{ex}(n, F) = \Theta(n)$ (provided that F has at least two edges), we have $D(n, H) = 2^{\Omega(n)}$. On the other hand, the approach of bounding $D(n, H)$ with the number of n -vertex F -free graphs only gives $D(n, H) = 2^{O(n \log n)}$. Up to this point, all the results are consistent with $D(n, H) = 2^{\Theta(\text{ex}(n, F))}$, which might suggest that in the remaining cases, when H is an orientation of a forest F , the same should hold.

As a natural starting point one might ask what happens with perhaps the simplest example of an oriented forest, namely the directed path on k edges, which we denote by P_k . We show that in this case the trivial lower bound is indeed tight up to a multiplicative absolute constant in the exponent.

THEOREM 1.4. *For any $k \geq 2$ and any $n \in \mathbb{N}$,*

$$D(n, P_k) \leq 2^{3kn}.$$

This result also suggests that $D(n, H)$ should always be $2^{\Theta(\text{ex}(n, F))}$. However, perhaps surprisingly, it turns out that this is not the case. As we will see in a moment, there are even orientations of a path for which the answer is $2^{\Theta(n \log n)}$. We completely resolve the remaining cases by showing that for every oriented forest H with at least two edges, either $D(n, H) = 2^{\Theta(n)}$ or $D(n, H) = 2^{\Theta(n \log n)}$. We also provide a precise characterisation for when each case occurs.

An oriented graph H is said to be *antidirected* if there exists a bipartition $V(H) = A \cup B$ of the vertex set such that every $u \in A$ has 0 incoming edges and every $v \in B$ has 0 outgoing edges. It is not too hard to see that antidirected forests are exactly those oriented forests H with the property that any n -vertex directed graph with at least Cn edges, for sufficiently large C , contains H . So in particular, $D(n, H) = 2^{\Theta(n)}$ for any antidirected forest H . Despite this it turns out that there are many more oriented forests for which the answer is also $2^{\Theta(n)}$. The following definition precisely captures all such oriented forests.

Definition. We call an oriented graph H *1-almost antidirected* if there exists a bipartition $V(H) = A \cup B$ of the vertex set such that there are no edges inside $H[A]$ and $H[B]$, every $u \in A$ has at most one incoming edge and every $v \in B$ has at most one outgoing edge.

For example, the path with the usual orientation is 1-almost antidirected, but there are orientations which are not. See Figure 1 for an orientation of the path with 5 vertices which is not 1-almost antidirected (the second and the fourth vertex must be on the same side in a bipartition, but the former has two out-edges and the latter has two in-edges).

We are now ready to state the value of $D(n, H)$ for oriented forests.



Fig. 1. Q is an example of an oriented path which is not 1-almost antidirected.

THEOREM 1.5. *Let H be an oriented forest with at least two edges. Then*

$$D(n, H) = \begin{cases} 2^{\Theta(n)} & \text{if } H \text{ is 1-almost antidirected,} \\ 2^{\Theta(n \log n)} & \text{otherwise.} \end{cases}$$

Notation. In this paper, no loops, multiple edges or bidirected edges are allowed in our oriented graphs. The underlying (undirected) graph of an oriented graph is the graph obtained by removing the orientations from all edges. An in-neighbour u of a vertex v is a vertex for which uv is an edge. An in-leaf in an oriented tree is a leaf whose only edge is directed towards the leaf. We define out-neighbours and out-leaves analogously.

2. Counting H -free orientations of graphs

2.1. General oriented graphs

In this section, we show how to deduce Proposition 1.2 and Proposition 1.3 from known results.

Let H be an arbitrary oriented graph. As before, let us write F for the underlying undirected graph of H . For a graph G , write $D(G, H)$ for the number of H -free orientations of G and denote by $N(G, F)$ the number of F -free spanning subgraphs of G . Moreover, write $N(n, F)$ for the number of F -free graphs with vertex set $[n]$. Kozma and Moran [19] proved that $N(G, F)$ is an upper bound for $D(G, H)$.

THEOREM 2.1 (Kozma–Moran [19]). *Let F be an undirected graph and let H be an orientation of F . Then for any undirected graph G ,*

$$D(G, H) \leq N(G, F).$$

In particular, for any $n \in \mathbb{N}$,

$$D(n, H) \leq N(n, F).$$

Perhaps surprisingly, the proof uses an inequality about set systems. Given a set system \mathcal{A} on ground set X , we say that $S \subset X$ is *shattered* by \mathcal{A} if for every $T \subset S$, there exists some $A \in \mathcal{A}$ with $A \cap S = T$. Let us write $\text{str}(\mathcal{A})$ for the collection of subsets of X which are shattered by \mathcal{A} . A general version of the celebrated Sauer–Shelah lemma [22] states that $|\mathcal{A}| \leq |\text{str}(\mathcal{A})|$.

To see that this inequality implies Theorem 2.1, fix an orientation Q of G . Identify the power set of $E(G)$ with the set of orientations of G by identifying $A \subset E(G)$ with the orientation of G which differs from Q precisely on the edge set A . Let \mathcal{A} be the collection of subsets of $E(G)$ which are identified with H -free orientations of G . Clearly, $|\mathcal{A}| = D(G, H)$. On the other hand, assume that some $S \subset E(G)$ is shattered by \mathcal{A} . Then the graph formed by the edges in S is F -free. Indeed, if it did contain F as a subgraph, then there would exist an orientation of the edges in S which contains a copy of H . This means that S could not be shattered by \mathcal{A} . Thus, $|\text{str}(\mathcal{A})| \leq N(G, F)$ and Theorem 2.1 follows.

The function $N(n, F)$ has been extensively studied. Similarly to $D(n, H)$, we have the trivial lower bound $N(n, F) \geq 2^{\text{ex}(n, F)}$ since if G is an F -free graph, then any subgraph of G is also F -free. This has been shown to be almost tight first for complete graphs by Erdős, Kleitman and Rothschild [14] and then for general non-bipartite graphs by Erdős, Frankl and Rödl [13], who proved that $N(n, F) = 2^{\text{ex}(n, F) + o(n^2)}$ (see [6] for an improved error term). Combined with Theorem 2.1 and the trivial lower bound for $D(n, H)$, this implies Proposition 1.2.

However, this gives an unsatisfactory answer for bipartite graphs F as in that case $\text{ex}(n, F) = O(n^{2-\epsilon_F})$ for some $\epsilon_F > 0$. Since any F -free graph on vertex set $[n]$ has at most $\text{ex}(n, F)$ edges, we get a straightforward upper bound

$$N(n, F) \leq \sum_{k=0}^{\text{ex}(n, F)} \binom{\binom{n}{2}}{k},$$

which implies that $N(n, F) \leq 2^{C \text{ex}(n, F) \log n}$ for some constant C . The logarithmic factor is necessary when F is acyclic with maximum degree at least 2. Indeed, in this case $\text{ex}(n, F) = O(n)$, but it is easy to see that there are $2^{\Omega(n \log n)}$ graphs on vertex set $[n]$ with maximum degree 1. On the other hand, it is not known in general whether the logarithmic factor is necessary for graphs that contain a cycle. In this direction, settling a classical conjecture of Erdős it was shown by Morris and Saxton [21] that $N(n, C_{2k}) = 2^{O_k(n^{1+1/k})}$, generalising a previous result of Kleitman and Winston [17], and complementing the classical Bondy–Simonovits bound $\text{ex}(n, C_{2k}) = O_k(n^{1+1/k})$ [9]. Balogh and Samotij [7, 8] established a similar result for complete bipartite graphs in place of even cycles. These results were generalised greatly by Ferber, McKinley and Samotij [15]. They showed that if F is a graph containing a cycle and there are positive constants α and A such that $\text{ex}(n, F) \leq An^\alpha$, then there exists a constant C depending only on α, A and F such that for all n , $N(n, F) \leq 2^{Cn^\alpha}$. This result, combined with Theorem 2.1 and $D(n, H) \geq 2^{\text{ex}(n, F)}$, implies Proposition 1.3.

2.2. Directed path

In this subsection we will prove Theorem 1.4. Let G be an n -vertex graph. Our task is to show that there are at most 2^{3kn} orientations of G which do not contain P_k . Let us fix a canonical ordering of the vertices of G . We will count the number of orientations with the help of the following algorithm. It takes as input an orientation of G . In each step, it processes a vertex and updates the current “state” for every vertex that has not been processed yet. We will show that, provided the orientation is P_k -free, the potential states are severely restricted. We then use this to bound the number of possible orientations.

Algorithm: Initially, we assign to every vertex v a state $(a_v, b_v) = (0, 0)$. At step i we have a sequence of already processed vertices v_1, \dots, v_{i-1} and possibly the next vertex to be processed v_i .

- (i) If v_i is not specified, it is chosen as a vertex in $V(G) \setminus \{v_1, \dots, v_{i-1}\}$ with largest a_{v_i} , breaking ties by choosing such a v_i first in the canonical ordering.
- (ii) We now proceed to process v_i . We consider all edges between v_i and not already processed vertices.
 - (i) If all these edges are oriented towards v_i , we do not specify v_{i+1} and continue to the next step.

- (ii) Otherwise we choose v_{i+1} to be the out-neighbour v of v_i with largest value of b_v (breaking ties by choosing such a v first in the canonical ordering) and
- (a) we increase a_u by one for any (non-processed) out-neighbour u of v_i and
 - (b) we increase b_u by one for any (non-processed) in-neighbour u of v_i which had $b_u \leq b_{v_{i+1}}$.

Let us first look at what is going on here. The algorithm reveals the orientation of the edges of G bit by bit; specifically at step i it will reveal the orientation of all (not already revealed) edges incident to v_i , the vertex we are currently processing. In particular, by the end of step i the algorithm has revealed the orientation of all edges incident to v_1, \dots, v_i . Note that the orientation of these edges determines uniquely the first i steps of the algorithm, regardless of how the remaining edges are oriented (in other words, the algorithm is by this point completely independent of the orientation of the remaining edges). Roughly speaking, we think of the states a_v and b_v as the length of a directed path ending and starting at v , respectively, only using the already revealed edges, or in other words using only already processed vertices (apart from v itself). This is captured more precisely in the following lemma.

LEMMA 2.2. *If the algorithm assigned state (a_v, b_v) to v (at any time) then the orientation contains a directed path of length a_v ending in v and a path of length $b_v - 1$ starting with v .*

Proof. Let v_1, \dots, v_n be the order in which the vertices are processed. Observe that both a_v and b_v are non-decreasing throughout the process and are never updated once we process v . Let a_t, b_t denote the final value of a_{v_t}, b_{v_t} so in particular at the point when we process v_t .

We are first going to show by induction on t that there is a path of length at least a_t ending in v_t which only uses previously processed vertices. For the base case $t = 1$, we know that $a_1 = 0$ so the claim trivially holds. Now given v_t , let $i < t$ be the last step at which point a_{v_i} was updated (if a_{v_i} was never updated, then $a_t = 0$ and our claim is trivial). Take the smallest j such that v_j, v_{j+1}, \dots, v_i form a directed path oriented towards v_i . In particular, this means that v_j was not an out-neighbour of v_{j-1} and it was chosen as a vertex with maximum a_v among all yet unprocessed vertices. As v_t was not yet processed at this stage, it follows that we had $a_{v_i} \leq a_{v_j}$ at step j . Since after this point a_{v_i} could only have been incremented when processing v_j, \dots, v_i , we know that $a_t \leq a_j + i - j + 1$. In addition, we know by induction that there is a path of length a_j ending at v_j and using vertices only from $\{v_1, \dots, v_j\}$ to which we can append $v_j \dots v_i v_t$ to obtain the desired path of length $a_j + i - j + 1 \geq a_t$.

Turning now to the b_v 's, we are going to show by induction on t that there is a path of length at least $b_t - 1$ starting with v_t , which otherwise only uses vertices in $\{v_1, \dots, v_{t-2}\}$. The base case $t = 1$ is trivial. Also, observe that if b_{v_t} got incremented at most once, i.e. $b_t \leq 1$, the claim also holds trivially. Let us now consider the penultimate (second to last) step i at which point b_{v_i} was updated. Observe that by definition, our algorithm will only increment b_{v_i} when processing v_j if $v_i v_j$ is an edge and $v_j v_{j+1}$ is an edge. In particular, this implies that we did not update b_{v_i} at step $t - 1$, and hence $i \leq t - 3$, since we chose i to be the penultimate step which updated b_{v_i} . The fact that we incremented b_{v_i} when processing v_i means that at that time b_{v_i} was at most $b_{v_{i+1}}$. Since this was the penultimate time b_{v_i} was incremented, we know that $b_t \leq b_{i+1} + 2$. By induction we can find a path of length $b_{i+1} - 1$ starting with v_{i+1} which otherwise only uses vertices from the set $\{v_1, \dots, v_{i-1}\}$.

This means that we can prepend $v_t v_i v_{i+1}$ to this path to obtain a new one of length $b_{i+1} - 1 + 2 \geq b_t - 1$ which only uses vertices from the set $\{v_1, \dots, v_{i+1}\}$ (apart from v_t). But $i \leq t - 3$, so $\{v_1, \dots, v_{i+1}\}$ is a subset of $\{v_1, \dots, v_{t-2}\}$, as desired.

The above lemma tells us that if our orientation was P_k -free, then $a_v, b_v \leq k$ throughout the process. Keeping this in mind, the following lemma considers the part of the orientation revealed before step i of our algorithm and gives a bound on the number of ways in which one can complete this partial orientation into a P_k -free one.

LEMMA 2.3. *Assume that up to step i our algorithm processed vertices v_1, \dots, v_{i-1} and assigned the state (a_v, b_v) to any remaining vertex v . We fix the orientation of edges incident to vertices v_1, \dots, v_{i-1} which leads to this state. There are at most*

$$\prod_{v \in V(G) \setminus \{v_1, \dots, v_{i-1}\}} (k + 2) \cdot 2^{2k - a_v - b_v}$$

ways to orient the remaining edges to complete the orientation without creating a P_k .

Before turning to the proof, note that for $i = 1$ no part of the orientation was specified and $a_v, b_v = 0$ for every v , so the lemma tells us there are at most $(k + 2)^n 2^{2kn} \leq 2^{3kn}$ orientations of G without a P_k , establishing Theorem 1.4.

Proof. Observe first that Lemma 2.2 guarantees that $a_v, b_v \leq k$ as otherwise any orientation we produce would contain a P_k .

We will prove this by reverse induction on i . For the base case of $i = n$ there are no remaining edges to orient so the claim holds trivially, since $a_{v_n}, b_{v_n} \leq k$.

Let us assume that it holds if we start from step $i + 1$ and any state. Given the orientation of the edges incident to v_1, \dots, v_{i-1} , we can determine which vertex will be v_i . We will now consider all the ways in which we can extend our partial orientation to include the orientation of all edges incident to v_i . That is, we consider all the ways we can orient the edges between v_i and $W := V(G) \setminus \{v_1, \dots, v_i\}$. For each extended partial orientation, we will use the induction hypothesis to bound the number of ways it can be completed into a P_k -free orientation. Summing over all choices for the orientation of the edges between v_i and W , we will get the desired bound.

If v_i has all its unspecified edges (i.e. edges to W) oriented towards v_i then the state of every remaining vertex remains unchanged and the induction hypothesis tells us that there are $\prod_{v \in W} (k + 2) \cdot 2^{2k - a_v - b_v}$ ways to complete it into a P_k -free orientation.

Otherwise our algorithm will choose v_{i+1} to maximise b_v among out-neighbours $v \in W$ of v_i . Assume this maximum is equal to b . Let us denote by d the number of neighbours $v \in W$ of v_i (in the underlying undirected graph G) which have b_v at most b . We know that any other neighbour of v_i in W must be an in-neighbour of v_i (by maximality of b). In particular, there are at most 2^d orientations of the edges incident to v_i which result in this choice of b . For any such orientation, we claim that our algorithm increased the sum $\sum_{v \in W} (a_v + b_v)$ by d . Indeed, any out-neighbour of v_i had its a_v incremented and any in-neighbour among the d neighbours with $b_v \leq b$ had their b_v incremented. This means that regardless of how we orient, by the induction hypothesis the number of ways to complete any of these orientations is at most $1/2^d \prod_{v \in W} (k + 2) \cdot 2^{2k - a_v - b_v}$. Since there are 2^d possible orientations of the edges incident to v_i (given this choice of b) and $k + 1$ choices for $0 \leq b \leq k$, this gives at most $(k + 1) \prod_{v \in W} (k + 2) \cdot 2^{2k - a_v - b_v}$ different ways to complete our

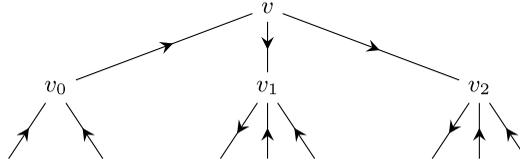


Fig. 2. $H_{1,2}$.

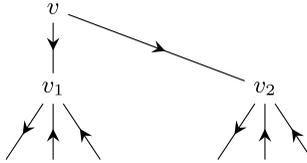


Fig. 3. $H_{1,2}^-$.

initial orientation. Adding this to the contribution of the case without any out-neighbours of v_i and using $a_{v_i}, b_{v_i} \leq k$ we obtain the claimed bound.

2.3. General trees

In this section we will complete the proof of Theorem 1.5. As already discussed in Section 2.1, we always have $D(n, H) \leq 2^{O(n \log n)}$ when H is an orientation of a forest. Let us now take $G = K_{\lfloor n/2 \rfloor, \lfloor n/2 \rfloor}$ (with an extra isolated vertex if n is odd), and orient all its edges from one part of the bipartition to the other (say from left to right), except for a single perfect matching which we orient in the other direction. Observe first that there are $\lfloor n/2 \rfloor! = 2^{\Omega(n \log n)}$ such orientations, since this is the number of choices for the matching that we have. As any such orientation of G is 1-almost antidirected, this shows that if H is not 1-almost antidirected then $D(n, H) = 2^{\Theta(n \log n)}$, as claimed in the second part of Theorem 1.5.

To establish the first part, we need to show that for any 1-almost antidirected forest H there are at most $2^{O(n)}$ H -free orientations of any n -vertex graph. We will actually show that this holds for a certain universal oriented tree H which contains all k -vertex 1-almost antidirected oriented forests. We define this universal oriented tree recursively, layer by layer.

The starting point is the tree $H_{1,t}$, which is defined as follows. It has a root v which has one in-neighbour v_0 and t out-neighbours v_1, \dots, v_t . Furthermore, v_0 has t in-neighbours and each of v_1, \dots, v_t has a single out-neighbour and t in-neighbours in addition to v . See Figure 2 for an illustration. In our recursive definition it will be convenient to have another building block which we call $H_{1,t}^-$ and which is obtained from $H_{1,t}$ by deleting the subtree rooted at v_0 . See Figure 3 for an illustration. $H_{s,t}$ is now defined by taking $H_{s-1,t}$ and appending a copy of $H_{1,t}$ to every out-leaf¹ and a copy of $H_{1,t}^-$ to every in-leaf. See Figure 4 for an illustration.

Observe first that $H_{k,k}$ contains any 1-almost antidirected tree on k vertices. This is due to the fact that in $H_{k,k}$ every non-leaf vertex at even depth has one in-neighbour and at least k out-neighbours and every vertex at odd depth has one out-neighbour and at least k

¹ Recall that an out-leaf is a leaf whose only edge is oriented away from the leaf.

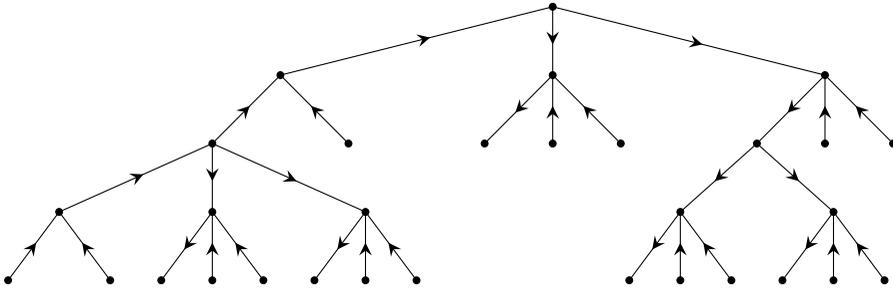
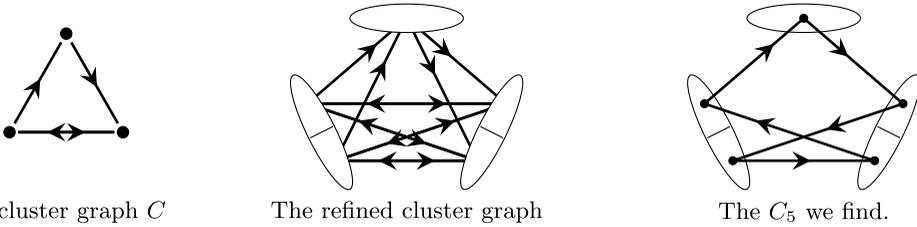


Fig. 4. Part of $H_{2,2}$.



In cluster graph C

The refined cluster graph

The C_5 we find.

Fig. 5. Illustration of how we find C_5 .

in-neighbours, which allows one to simply greedily embed any 1-almost antidirected tree (starting from the root of $H_{k,k}$ to ensure that a leaf of $H_{k,k}$ is not encountered). Observe also that $H_{k+1,k}$ contains any 1-almost antidirected forest on k vertices. This follows since if we remove the root and its neighbours, the remaining oriented graph contains many (at least k) pairwise vertex-disjoint subtrees which are isomorphic to $H_{k,k}$, and we can embed the 1-almost antidirected trees making up our 1-almost antidirected forest into separate ones. With this in mind, the following theorem is the key result we need to prove in order to establish the remaining case of Theorem 1.5.

THEOREM 2.4. *Let k, s, t be positive integers. Let $G = (X, Y)$ be a bipartite graph on n vertices. Then there are at most $2^{O_{k,s,t}(n)}$ orientations of G which contain no $H_{s,t}$ with the root in X and in which there are at most k out-edges from each $y \in Y$.*

Here and below, the notation $O_{k,s,t}$ means that the implied constant can depend on k, s and t . We will prove this theorem by induction on s . We will prove the case $s = 1$ separately, as it will both serve as the base case and be useful in the induction step.

LEMMA 2.5. *Let k, t be positive integers. Let $G = (X, Y)$ be a bipartite graph on n vertices. Then there are at most $2^{O_{k,t}(n)}$ orientations of G which contain no $H_{1,t}$ with the root in X and in which there are at most k out-edges from each $y \in Y$.*

Proof. Set $T := \max(|V(H_{1,t})|, kt) = O_{k,t}(1)$.

Let D be an orientation of G which contains no $H_{1,t}$ with the root in X and in which there are at most k out-edges from each $y \in Y$. Let $Y_1 = Y$ and let X_1 be the set of vertices in X which have at least one in-edge in D . Note that there are at most 2^n possibilities for X_1 . Let

$Y_2 \subseteq Y_1$ be the set of vertices in Y_1 which have at least $T + k$ neighbours in X_1 . There are at most $2^{(T+k)(|Y_1|-|Y_2|)}$ ways to orient the edges incident to $Y_1 \setminus Y_2$. Consider the subset $X_2 \subseteq X_1$ consisting of vertices which have an in-neighbour in Y_2 . Since every vertex $y \in Y_1 \setminus Y_2$ has at most $T + k$ neighbours in X_1 , there are at most $(T + k)(|Y_1| - |Y_2|)$ vertices in X_1 which have a neighbour in $Y_1 \setminus Y_2$, and we can specify the subset of them which got removed from X_1 to obtain X_2 in at most $2^{(T+k)(|Y_1|-|Y_2|)}$ many ways. We repeat as long as we can, i.e. until we obtain subsets $X_\ell \subset X_1$ and $Y_\ell \subset Y_1$ which have the following properties: every vertex of X_ℓ has an in-neighbour in Y_ℓ and every vertex in Y_ℓ has at least $T + k$ neighbours inside X_ℓ . By the above counting, the number of possibilities for X_ℓ, Y_ℓ and the orientations of the edges incident to $(X \setminus X_\ell) \cup (Y \setminus Y_\ell)$ is at most $2^n \cdot 2^{2(T+k)|Y_1|} \leq 2^{n+2(T+k)n}$.

If $Y_\ell = \emptyset$, then we have already revealed the entire orientation of G , so there are at most $2^{n+2(T+k)n}$ such suitable orientations. Assume that $Y_\ell \neq \emptyset$. We claim that this, together with the assumption that any vertex in $Y_\ell \subseteq Y$ has at most k out-neighbours in X , guarantees that we can find a copy of $H_{1,t}$ in $D[X_\ell \cup Y_\ell]$ with the root in X_ℓ , which is a contradiction. Indeed, fix one edge incoming from Y_ℓ at every vertex in X_ℓ . These edges span vertex disjoint out-directed stars of size at most k with centres in Y_ℓ . In particular, there are at least $|X_\ell|/k$ centres. Since each centre has at least $T + k$ neighbours in X_ℓ , at most k of which can be out-neighbours, there are at least T in-neighbours. Since $T \geq tk$, this means that some vertex $v \in X_\ell$ is an in-neighbour of at least t distinct centres v_1, \dots, v_t of our out-stars. Picking one out-edge per star gives us an out-directed tree consisting of root v and t vertex-disjoint paths of length 2. Note also that it is guaranteed that there is an in-neighbour $v_0 \in Y_\ell$ of v (which is distinct from v_1, \dots, v_t since they are its out-neighbours). What remains to be done is to find t in-neighbours for each of v_0, \dots, v_t which we can do greedily since each of them has at least $T \geq |V(H_{1,t})|$ in-neighbours in X_ℓ . So we found a copy of $H_{1,t}$ as claimed, and are done.

Let us now define the oriented tree $H_{s,t,t}'^*$ by modifying $H_{s,t}$ so that every vertex in the penultimate layer has t' instead of t out-leaves attached to it. Note that if $t' \geq t$, then $H_{s,t}'$ contains $H_{s,t,t}'^*$ as a subgraph.

In the induction step, we will use the following key lemma.

LEMMA 2.6. *For every k, s, t there exists $t' = t'(k, s, t)$ as follows. Let $D = (X, Y)$ be a bipartite oriented graph such that there are at most k out-edges from each $y \in Y$. Assume that D contains a copy of $H_{s,t,t}'^*$ with the root in X and assume that each leaf in this copy is the root of an $H_{1,t}'$. Then D contains $H_{s+1,t}$ with the root in X .*

Proof. Throughout the proof, we assume that t' is sufficiently large.

Take a subgraph K in D which is isomorphic to $H_{s,t,t}'^*$ with the root in X and in which every leaf is the root of an $H_{1,t}'$. Choose also a subgraph $L(w)$ isomorphic to $H_{1,t}'$ with root w for every leaf w in K . Observe that for every leaf w of K , w has a unique in-neighbour in $L(w)$, call this vertex $f(w)$ and note that $f(w) \in Y$.

Claim. K has a subgraph K' isomorphic to $H_{s,t}$ with the same root as K such that for every out-leaf w in K' , the vertex $f(w)$ is not in $V(K')$, and all these $f(w)$'s are distinct.

Proof of Claim. To get a subgraph of $H_{s,t,t}'^*$ isomorphic to $H_{s,t}$, we need to keep t of the t' out-leaves for every vertex of the penultimate layer. We can do this one by one in an arbitrary

order. We just need to pay attention that for each out-leaf w that we keep, the vertex $f(w)$ should be different from every vertex that is already in K' , and moreover all the $f(w)$'s for different w 's should be different. This can be done; indeed, every $f(w)$ is in Y and every $y \in Y$ has at most k out-edges in D , so at any point the number of forbidden choices for w is at most $2|V(H_{s,t})| \cdot k$, hence $t' > 2|V(H_{s,t})| \cdot k$ suffices.

It remains to extend K' to a copy of $H_{s+1,t}$. For this, we need to “attach” a copy of $H_{1,t}$ to each out-leaf in K' , and we need to attach a copy of $H_{1,t}^-$ to each in-leaf in K' in a way that all new vertices are distinct from each other and from the vertices of K' .

We begin by extending K' by joining $f(w)$ to w for each out-leaf w of K' . By the above claim, all of these $f(w)$'s are distinct and disjoint from K' . Next for every leaf w of K' we want to append t vertex disjoint (apart from sharing the start vertex w) paths of length 2 directed away from w . This can be done since $L(w)$ gives us $t' > |V(H_{s+1,t})|$ vertex disjoint paths of length 2 starting at w , so no matter how many vertices we already embedded we still have one of these paths available.

It remains to attach t in-neighbours to each vertex in the penultimate layer of our partial $H_{s+1,t}$. Since the vertices in the penultimate layer are in the middle layer of some $L(w)$, we know that each of these vertices has at least t' in-neighbours in D so once again we will always have enough of them available.

Proof of Theorem 2.4. We will use induction on s . The case $s = 1$ is Lemma 2.5. Assume now that we have verified the statement for $s - 1$. Let k, s, t be positive integers and let $t' = t'(k, s - 1, t)$ from Lemma 2.6.

Let D be an $H_{s,t}$ -free orientation of G with the property that there are at most k out-edges from every $y \in Y$. Let X_1 be the set of vertices in X which are roots of a copy of $H_{1,t'}$ in D and let $X_2 = X \setminus X_1$. Clearly there are at most 2^n possibilities for X_1 . The oriented graph $D[X_2 \cup Y]$ contains no $H_{1,t'}$ with the root in X_2 , so the number of possibilities for the orientation of $G[X_2 \cup Y]$ is at most $2^{O_{k,t'}(n)}$ by Lemma 2.5. Moreover, since D is $H_{s,t}$ -free, Lemma 2.6 implies that $D[X_1 \cup Y]$ contains no copy of $H_{s-1,t,t'}^*$, and hence also no copy of $H_{s-1,t'}$, with the root in X_1 . Then by induction there are at most $2^{O_{k,s-1,t'}(n)}$ possibilities for the orientation of $G[X_1 \cup Y]$. Combining all three upper bounds, the result follows.

Finally let us deduce the first part of Theorem 1.5.

COROLLARY 2.7. *Let $k \geq 1$, let G be an n -vertex graph and let H be a k -vertex 1-almost antidirected forest. There are at most $2^{O_k(n)}$ different H -free orientations of G .*

Proof. The proof is by induction on k . For the base case note that the $k \leq 2$ case is trivial. Let us now assume that the statement holds for forests with $k - 1$ vertices. We may w.l.o.g. assume that H has an in-leaf. Let H' denote the oriented forest obtained from H with this leaf removed. Given an orientation D of G , let X be the subset of $V(G)$ consisting of vertices with out-degree at least k . There are 2^n different options for X . Let $Y = V(G) \setminus X$. Observe first that if we could find H' inside X then we could extend it to a copy of H in D since every vertex of X has out-degree at least k . This means that by induction there are at most $2^{O_{k-1}(n)}$ many ways to orient the edges inside X . There are at most $|Y|k \leq kn$ edges inside Y , so the edges inside Y can be oriented in at most 2^{kn} many ways. Finally, the number of ways to orient the edges between X, Y in a way that any vertex in Y has at most k out-edges and without creating a copy of $H_{k+1,k}$ is at most $2^{O_k(n)}$ by Theorem 2.4. Since $H \subseteq H_{k+1,k}$, this completes the proof.

2.4. *Odd cycles*

We will assume some familiarity with the basic directed regularity lemma, the specific details needed are given in Section 2 of [3]. Since the use of regularity in our argument is essentially the same as in both [2, 3], we will not go into more technical details of these parts and will refer the reader to either of these papers, with the goal of making the key part of the argument easier to follow.

The following lemma says that if there are many orientations of G which are C_{2k+1} -free then G is not far from being bipartite. It is analogous to [3, Lemma 2.1] which replaces C_{2k+1} with an arbitrary tournament.

LEMMA 2.8. *Let $k \geq 1$ and $\delta > 0$ there exists $n_0 = n_0(\delta, k)$ such that if G is a graph of order $n \geq n_0$ which has at least $2^{\lfloor n^2/4 \rfloor}$ distinct C_{2k+1} -free orientations then there is a bipartition of $V(G)$ with at most δn^2 edges inside parts.*

Proof. Let us fix $0 < \varepsilon \ll \eta \ll \beta \ll \alpha \ll \delta$ as needed for various points of the upcoming argument.

Let \vec{G} be a C_{2k+1} -free orientation of G . We apply the directed regularity lemma to \vec{G} to obtain an ε -regular partition $V(\vec{G}) = V_1 \cup \dots \cup V_m$ (all V_i 's should have sizes as equal as possible, and all but εm^2 pairs (V_i, V_j) should satisfy that linear sized subsets have about the same density of edges in each direction as the density of edges between V_i, V_j in the same direction). We then consider a cluster (di)graph C of density η (its vertices are the parts of our partition and two parts are joined by a directed edge if they are ε -regular and the density of edges in the corresponding direction is at least η). Note that C can have bidirected edges.

We first want to show that there exists some orientation \vec{G} for which the resulting cluster graph has at least $m^2/4 - \beta m^2$ edges directed both ways. We claim that if this is not the case, then there would be too few (less than $2^{\lfloor n^2/4 \rfloor}$) orientations possible. Since the regularity lemma guarantees that $m \leq M = M(\varepsilon)$, there are at most M^n choices for the ε -regular partition \mathcal{P} , at most $2^{\binom{M}{2}}$ choices for which pairs are ε -regular and $4^{\binom{M}{2}}$ choices for C . In total there are at most $M^n 2^{3M^2/2}$ choices for \mathcal{P} , regular pairs and C . Let us now bound how many orientations could give rise to a fixed choice. There are few edges inside parts of our fixed \mathcal{P} and between non- ε -regular pairs (at most εn^2 in both cases) and each edge may be oriented in two ways, so the total contribution of these edges to the number of orientations is at most a factor of $2^{2\varepsilon n^2}$. For any ε -regular pair (V_i, V_j) which is not an edge of C in one of the directions, there are at most about $\eta n^2/m^2$ directed edges in that direction. An easy estimate tells us that the edges between V_i and V_j can be oriented like this in at most $2^{c_\eta n^2/m^2}$ many ways where $c_\eta \rightarrow 0$ as $\eta \rightarrow 0$ and c_η only depends on η . Since there are at most m^2 such pairs (V_i, V_j) , orienting edges between them contributes at most a factor of $2^{c_\eta n^2}$ to the total number of orientations. Finally, for any edge of C directed both ways, there are at most $2^{(n/m)^2}$ orientations of the edges between the corresponding pair of parts, but since we are assuming that C has at most $m^2/4 - \beta m^2$ such edges, they contribute at most a factor of $2^{n^2/4 - \beta n^2}$ to the total number of orientations. Putting it all together we get at most

$$M^n 2^{3M^2/2} \cdot 2^{2\varepsilon n^2} \cdot 2^{c_\eta n^2} \cdot 2^{n^2/4 - \beta n^2}$$

orientations. Choosing η to be small enough compared to β gives us a contradiction to having at least $2^{\lfloor n^2/4 \rfloor}$ orientations.

Let now \vec{G} be an orientation for which the resulting cluster graph C has at least $m^2/4 - \beta m^2$ edges directed both ways. We claim that C can not contain a bidirected triangle missing only a single directed edge² as otherwise \vec{G} would contain a C_{2k+1} . This is a consequence of a standard embedding lemma. One can easily deduce it from the classical (undirected) embedding lemma (see e.g. [18, Lemma 2.1]) by first refining the partition (splitting each part into k parts of size as equal as possible, which preserves the regularity while density drops to at worst $\eta - \varepsilon \geq \eta/2$) then only keeping edges in the desired direction, and applying the usual embedding lemma; see Figure 5 for an illustration.

In particular, this tells us that the graph consisting only of the bidirected edges of C is both triangle-free and has at least $m^2/4 - \beta m^2$ edges. The stability theorem of Simonovits [25] shows that there is a bipartition $V(C) = W_1 \cup W_2$ with at most αm^2 bidirected edges within a part (using that $\alpha \gg \beta$). Hence, the bipartite subgraph consisting of the bidirected edges of C between W_1 and W_2 has at least $m^2/4 - (\beta + \alpha)m^2$ edges. If C had in addition more than $8(\alpha + \beta)m^2$ directed edges inside parts, we would find a bidirected triangle with one directed edge removed in C . Indeed, more than $4(\alpha + \beta)m^2$ of these additional edges must be inside a single part, say W_1 , and we can pass to a bipartite subgraph of size more than $2(\alpha + \beta)m^2$ within W_1 . Taking into account these edges might also be bidirected there are more than $(\alpha + \beta)m^2$ distinct pairs spanning a directed edge. These edges together with the bidirected edges between W_1 and W_2 make a subgraph of C with more than $m^2/4$ edges, so by Mantel's theorem they give a triangle. This triangle has at most one edge inside W_1 (since the edges we used inside W_1 form a bipartite graph), so it has at least two bidirected edges, as desired.

It follows from the above that there are at most $\alpha m^2 + 8(\alpha + \beta)m^2$ edges of C inside W_1 and W_2 . Remove all edges of G which correspond to such edges of C . Moreover, remove all edges within V_i 's and between pairs (V_i, V_j) corresponding to non-edges in C . The remaining subgraph of G is bipartite (with the parts being the union of V_i 's corresponding to W_1 and to W_2). Since there are at most εn^2 edges within V_i 's, at most εn^2 edges between non- ε -regular pairs and at most $2\eta n^2$ edges between ε -regular pairs which are non-edges in C , we have in total removed at most $(\alpha + 8(\alpha + \beta))n^2 + \varepsilon n^2 + \varepsilon n^2 + 2\eta n^2 \leq \delta n^2$ edges, as desired.

The following lemma replaces the embedding [3, Lemma 3.1]. Let us introduce some notation for convenience. Given an oriented graph D and an integer k , we say that a pair of disjoint subsets $W_1, W_2 \subseteq V(D)$ with $|W_i| \geq 2k$ is k -rich if for any $X_i \subseteq W_i, |X_i| \geq |W_i|/20$, D has at least $\frac{1}{10}|X_1||X_2|$ edges from X_1 to X_2 , as well as at least $1/10|X_1||X_2|$ edges from X_2 to X_1 .

LEMMA 2.9. *Let D be an oriented graph and let $W_1, W_2 \subseteq V(D)$ be a k -rich pair. Then one can find a directed path of length $2k$ in the bipartite oriented graph $D[W_1, W_2]$ starting and ending in W_1 .*

Proof. We iteratively find our directed path. Assume that for some $i \leq k$, we have found a path $v_1 v_2 \dots v_{2i-1}$ and a subset $V_{2i-1} \subseteq W_2 \setminus \{v_2, v_4, \dots, v_{2i-2}\}$ of at least $|W_2|/20$ out-neighbours of v_{2i-1} . Then since D has at least $1/10|V_{2i-1}||W_1 \setminus \{v_1, v_3, \dots, v_{2i-1}\}|$ edges oriented from V_{2i-1} to $W_1 \setminus \{v_1, v_3, \dots, v_{2i-1}\}$, there must be a vertex v_{2i} in V_{2i-1} with a set V_{2i} of at least $(|W_1| - k)/10 \geq |W_1|/20$ out-neighbours in $W_1 \setminus \{v_1, v_3, \dots, v_{2i-1}\}$.

² In fact even having an oriented triangle with one bidirected edge would suffice, but this does not seem to be more useful.

Repeating from the other side completes the iteration. After k iterations, we find the desired path.

We now turn to the proof of the main result in this section.

Proof of Theorem 1.1. Let $n_0 = n_0(\delta^2, k)$ be given by Lemma 2.8 applied with δ^2 in place of δ , for some sufficiently small δ .

Let us take a graph G on $n > n_0^2 + n_0$ vertices which has at least $2^{\lfloor n^2/4 \rfloor + m} C_{2k+1}$ -free orientations for some $m \geq 0$. We will show that if G is not the Turán graph, then either we can find a vertex v such that $G \setminus v$ has at least $2^{\lfloor (n-1)^2/4 \rfloor + m + 1} C_{2k+1}$ -free orientations or we can find distinct vertices u and v such that $G \setminus \{u, v\}$ has at least $2^{\lfloor (n-2)^2/4 \rfloor + m + 2} C_{2k+1}$ -free orientations. We then iterate (note that no subgraph we consider can any longer be a Turán graph since it has too many orientations, so also edges) as long as our graph has at least n_0 vertices. When we stop, we obtain a graph with less than n_0 vertices which has at least $2^{n_0^2}$ orientations, which is impossible, since it has at most $n_0^2/2$ edges.

Let us assume that G is not the Turán graph on $n \geq n_0$ vertices and proceed to find a suitable vertex v .

Every vertex needs to have degree at least $\lfloor n/2 \rfloor$ as otherwise its edges contribute at most a factor of $2^{\lfloor n/2 \rfloor - 1}$ to the number of orientations so it would immediately work as our vertex v above.

Let V_1, V_2 form a bipartition of $V(G)$ which minimises the number of edges within parts. Since $n \geq n_0$, by Lemma 2.8 we have at most $\delta^2 n^2$ edges within parts. This implies $|V_1|, |V_2| \leq (1/2 + \delta)n$ as otherwise G would have less than $n^2/4$ edges, so too few orientations. Similarly, there can be at most $\delta^2 n^2$ edges missing between parts.

We first claim that there can be only few orientations for which there exists a pair of disjoint subsets $X_1 \subseteq V_1, X_2 \subseteq V_2$, both of size at least $2\delta n$, which have at most $|X_1||X_2|/10$ edges directed from X_1 to X_2 . The number of such orientations of edges between X_1, X_2 is at most

$$\sum_{i=0}^{|X_1||X_2|/10} \binom{e(X_1, X_2)}{i} \leq \sum_{i=0}^{|X_1||X_2|/10} \binom{|X_1||X_2|}{i} \leq 2^{|X_1||X_2|/2},$$

where $e(X_1, X_2)$ stands for the number of edges between X_1 and X_2 . Since the total number of edges is at most $n^2/4 + \delta^2 n^2$, there are at most $2^{n^2/4 + \delta^2 n^2 - |X_1||X_2|/2} \leq 2^{n^2/4 - \delta^2 n^2}$ such orientations of the whole graph. Since we can choose X_1 and X_2 in at most 2^{2n} many ways, there can be at most $2^{2n} \cdot 2^{n^2/4 - \delta^2 n^2} \leq 2^{\lfloor n^2/4 \rfloor - 1}$ orientations for which such a pair X_1, X_2 exists.

Let us now consider only C_{2k+1} -free orientations such that for any pair of disjoint subsets X_1, X_2 of size at least $2\delta n$, there are at least $|X_1||X_2|/10$ edges oriented from X_1 to X_2 and also from X_2 to X_1 . Then any pair of subsets, both of size at least $40\delta n$, is k -rich. We call such an orientation *relevant* and by the above counting and our assumption on the number of C_{2k+1} -free orientations of G , there are at least $2^{\lfloor n^2/4 \rfloor + m} - 2^{\lfloor n^2/4 \rfloor - 1} \geq 2^{\lfloor n^2/4 \rfloor + m - 1}$ relevant orientations.

Case 1. Some vertex v has at least $800\delta n$ neighbours in its own part, say V_1 .

Note that v must have at least $800\delta n$ neighbours in V_2 as well, by maximality of the number of edges between V_1 and V_2 . If in a relevant orientation v has at least $40\delta n$ out-neighbours and at least $40\delta n$ in-neighbours belonging to different parts, then since these sets make a

k -rich pair we can use Lemma 2.9 to find a path of length $2k - 1$ and join it with v to give a C_{2k+1} , a contradiction. This means that at least 2 out of the 4 sets: out-neighbours of v in V_1 , in-neighbours of v in V_1 , out-neighbours of v in V_2 and in-neighbours of v in V_2 need to have size at most $40\delta n$. These two parts can not belong to the same V_i or be of different types in different parts. The only remaining option is for v to have at most $80\delta n$ in-neighbours or at most $80\delta n$ out-neighbours. In particular, its edges may be oriented in such a way in at most

$$2 \sum_{i=0}^{80\delta n} \binom{d(v)}{i} \leq 2 \sum_{i=0}^{d(v)/10} \binom{d(v)}{i} \leq 2^{0.49d(v)} \leq 2^{0.49n}$$

many ways, where $d(v) \leq n$ denotes the degree of v in G . In other words $G \setminus \{v\}$ must have at least

$$2^{\lfloor n^2/4 \rfloor + m - 1 - 0.49n} \geq 2^{\lfloor (n-1)^2/4 \rfloor + m + 1}$$

C_{2k+1} -free orientations, as desired.

Case 2. Every vertex of G has at most $800\delta n$ neighbours in its own part.

Since G is not the Turán graph there must exist an edge uv inside a part, say in V_2 . Both u and v have at least $\lfloor n/2 \rfloor - 800\delta n \geq n/3$ neighbours in V_1 , in particular they have $d(u, v) \geq n/8$ common neighbours in V_1 since parts have size at most $n/2 + \delta n$. If in a relevant orientation uv is an edge, then the set W_1 of out-neighbours of v in V_1 which are also in-neighbours of u has size at most $40\delta n$. This is due to Lemma 2.9 (applied with W_1 and $W_2 = V_2 \setminus \{u, v\}$) which allows us to find a path of length $2k - 2$ starting and ending in W_1 , which in turn can be completed into a C_{2k+1} using uv . This will severely reduce the number of possible orientations of edges incident to u or v . More precisely, the edges from u and v to their common neighbours can be oriented in at most

$$\sum_{i=0}^{40\delta n} \binom{d(u, v)}{i} \cdot 4^i \cdot 3^{d(u, v) - i} \leq 41\delta n \cdot \binom{d(u, v)}{40\delta n} \cdot 4^{40\delta n} \cdot 3^{d(u, v) - 40\delta n} \leq 4^{0.99d(u, v)}.$$

The same bound analogously holds if vu is the edge instead. In particular, there are at most

$$2^{d(u)+d(v)-0.02d(u, v)} \leq 2^{n-n/1000}$$

possible orientations of edges incident to u or v (we are using that both u and v have degree at most $n/2 + \delta n + 800\delta n$ and that δ is small). In particular, the total number of orientations of $G \setminus \{u, v\}$ is at least

$$2^{\lfloor n^2/4 \rfloor + m - 1} / 2^{n-n/1000} \geq 2^{\lfloor (n-2)^2/4 \rfloor + m + 2}$$

and we are done.

3. Concluding remarks and open problems

A classical result of Erdős and Gallai [12] states that $ex(n, P_k) \leq (k - 1)n/2$ and this is tight when k divides n . (With a slight abuse of notation, here P_k refers to the unoriented path of length k .) The tightness of this bound implies that $D(n, P_k) \geq 2^{\frac{(k-1)n}{2}}$ when k divides n , showing that our Theorem 1.4 is tight up to an absolute constant factor in the exponent. Our bound coming from Lemma 2.3 is actually slightly better than 2^{3kn} , namely $2^{(1+o_k(1))2kn}$,

where $o_k(1) \rightarrow 0$ as $k \rightarrow \infty$. It seems not too hard to improve the constant further. We pose the following question.

Question 3.1. Let k be a positive integer. Does there exist $n_0 = n_0(k)$ such that for every $n \geq n_0$, we have $D(n, P_k) \leq 2^{\frac{(k-1)n}{2}}$?

We were able to answer the question in the affirmative for $k \leq 3$. For general k , proving an upper bound of the form $2^{(1+o_k(1))kn/2}$ would already be interesting.

Turning to general oriented forests H , we have seen that $D(n, H) = 2^{\Theta(\text{ex}(n, F) \log n)}$ when H is not 1-almost antidirected (as before, F is the underlying undirected graph). We also proved that the logarithmic factor is not present when H is 1-almost antidirected. However, even in that case, $D(n, H)$ need not be very close to $2^{\text{ex}(n, F)}$ as the following example shows. Let H be the oriented star on $k + 1$ vertices whose edges are all oriented out from the centre. The underlying undirected graph has Turán number about $(k - 1)n/2$. On the other hand, we claim that for every $\varepsilon > 0$ there exists k such that for all sufficiently large n , we have $D(n, H) \geq 2^{(1-\varepsilon)kn}$. Indeed, given $\varepsilon > 0$, let k be large and take a $(1 - \varepsilon/2)2k$ -regular graph G_0 on $2k$ vertices. If k is large enough, using standard Chernoff estimates we can show that in a random orientation of G_0 the out-degree of every vertex is at most $(1 - \varepsilon/4)k$ and in particular the orientation is H -free, with probability at least $1/2$. Hence, there are at least $(1/2)2^{(1-\varepsilon/2)2k^2} \geq 2^{(1-\varepsilon)k|V(G_0)|}$ H -free orientations of G_0 . So for large n we can take G to be the union of $n/2k$ vertex-disjoint copies of G_0 , and then $D(G, H) \geq 2^{(1-\varepsilon)kn}$.

However, it remains possible that there exists an absolute constant C such that for every 1-almost antidirected oriented tree H there exists $n_0 = n_0(H)$ such that for all $n \geq n_0$, we have $D(n, H) \leq 2^{C \text{ex}(n, F)}$.

Turning to non-bipartite underlying graphs, Proposition 1.2 shows that $\log_2 D(n, H)$ is asymptotically $\text{ex}(n, F)$ in this case. On the other hand, one can construct non-bipartite examples for which $D(n, H) > 2^{\text{ex}(n, F)}$ even for arbitrarily large n . For example, let F be the complete bipartite graph $K_{5,5}$ with a path of length 4 placed in one of the parts. Moreover, let H be an orientation of F in which every vertex has at least two out-neighbours and at least two in-neighbours in the other part and in which the path of length 4 is oriented in a not 1-almost antidirected way (see Figure 1 for such an orientation of the path). Then $\text{ex}(n, F) = n^2/4 + O(n)$, but if G is the complete tripartite graph with parts of size $n/2, n/4, n/4$, then $D(G, H) \geq 2^{n^2/4} \cdot (n/4)!$, which shows that $D(n, H) > 2^{\text{ex}(n, F)}$.

One might also try to determine $D(n, H)$ more precisely when H is a forest which is not 1-almost antidirected. Here it was observed by Alon that if the maximum degree of the forest H is d , then $\log D(n, H) = O(nd \log n)$ (now the implied constant does not depend on H). On the other hand, by an obvious modification of our lower bound construction, one can show that there are oriented trees with maximum degree d such that $\log D(n, H) = \Omega(nd \log n)$. This determines $\log D(n, H)$ up to an absolute constant factor for many, but not for all, non 1-almost antidirected forests H .

Acknowledgement. We are grateful to Noga Alon and Rob Morris for helpful comments on a draft of this paper. We also thank the anonymous referee for many useful suggestions which improved the presentation.

REFERENCES

- [1] P. ALLEN, Y. KOHAYAKAWA, G. O. MOTA and R. F. PARENTE. On the number of orientations of random graphs with no directed cycles of a given length, *Electron. J. Combin.* (2014), P1–52.
- [2] N. ALON, J. BALOGH, P. KEEVASH and B. SUDAKOV. The number of edge colorings with no monochromatic cliques, *J. London Math. Soc.* (2) **70** (2004), no. 2, 273–288.
- [3] N. ALON and R. YUSTER. The number of orientations having no fixed tournament, *Combinatorica* **26** (2006), no. 1, 1–16.
- [4] P. ARAÚJO, F. BOTLER and G. O. MOTA. Counting graph orientations with no directed triangles, Preprint arXiv:2005.13091 (2020).
- [5] J. BALOGH. A remark on the number of edge colorings of graphs, *European J. Combin.* **27** (2006), no. 4, 565–573.
- [6] J. BALOGH, B. BOLLOBÁS and M. SIMONOVITS. The number of graphs without forbidden subgraphs, *J. Combin. Theory Ser. B* **91** (2004), no. 1, 1–24.
- [7] J. BALOGH and W. SAMOTIJ. The number of $K_{m,m}$ -free graphs, *Combinatorica* **31** (2011), no. 2, 131–150.
- [8] J. BALOGH and W. SAMOTIJ. The number of $K_{s,t}$ -free graphs, *J. Lond. Math. Soc.* (2) **83** (2011), no. 2, 368–388.
- [9] J. BONDY and M. SIMONOVITS. Cycles of even length in graphs, *J. Combin. Theory, Series B* **16** (1974), no. 2, 97–105.
- [10] M. COLLARES, Y. KOHAYAKAWA, R. MORRIS and G. O. MOTA. Counting restricted orientations of random graphs, *Random Structures Algorithms* **56** (2020), no. 4, 1016–1030.
- [11] P. ERDŐS. *Some new applications of probability methods to combinatorial analysis and graph theory* Proceedings of the Fifth Southeastern Conference on Combinatorics, Graph Theory and Computing (Florida Atlantic Univ., Boca Raton, Fla., 1974), pp. 39–51. Congressus Numerantium, No. X.
- [12] P. ERDŐS and T. GALLAI. On maximal paths and circuits of graphs, *Acta Math. Acad. Sci. Hungar.* **10** (1959), 337–356.
- [13] P. ERDŐS, P. FRANKL and V. RÖDL. The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent, *Graphs Combin.* **2** (1986), no. 1, 113–121.
- [14] P. ERDŐS, D. J. KLEITMAN and B. L. ROTHSCHILD. Asymptotic enumeration of K_n -free graphs, International Colloquium on Combinatorial Theory, 1976.
- [15] A. FERBER, G. MCKINLEY and W. SAMOTIJ. Supersaturated sparse graphs and hypergraphs, *Internat. Math. Res. Notices* (2020), no. 2, 378–402.
- [16] Z. FÜREDI and M. SIMONOVITS. The history of degenerate (bipartite) extremal graph problems, (Springer, 2013), pp. 169–264.
- [17] D. J. KLEITMAN and K. J. WINSTON. On the number of graphs without 4-cycles, *Discrete Math.* **41** (1982), no. 2, 167–172.
- [18] J. KOMLÓS and M. SIMONOVITS. Szemerédi’s regularity lemma and its applications in graph theory, *Combinatorics*, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993), *Bolyai Soc. Math. Stud.*, vol. 2 (János Bolyai Math. Soc., Budapest, 1996), pp. 295–352.
- [19] L. KOZMA and S. MORAN. Shattering, graph orientations, and connectivity, *The Electronic Journal of Combinatorics* **20** (2013), no. 3, P.44.
- [20] H. LEFMANN, Y. PERSON, V. RÖDL and M. SCHACHT. On colourings of hypergraphs without monochromatic Fano planes, *Combin. Probab. Comput.* **18** (2009), no. 5, 803–818.
- [21] R. MORRIS and D. SAXTON. The number of $C_{2\ell}$ -free graphs, *Advances in Mathematics* **298** (2016), 534–580.
- [22] A. PAJOR. *Sous-Espaces ℓ_1^n des Espaces de Banach*. (Editions Hermann, 1985).
- [23] O. PIKHURKO, K. STADEN and Z. B. YILMA. The Erdős–Rothschild problem on edge-colourings with forbidden monochromatic cliques, *Math. Proc. Camb. Phil. Soc.* **163** (2017), no. 2, 341–356.
- [24] O. PIKHURKO and Z. B. YILMA. The maximum number of K_3 -free and K_4 -free edge 4-colorings, *J. London Math. Soc.* **85** (2012), no. 3, 593–615.
- [25] M. SIMONOVITS. A method for solving extremal problems in graph theory, stability problems, *Theory of Graphs* (Proc. Colloq., Tihany, 1966)(Academic Press, New York, 1968), pp. 279–319.
- [26] R. YUSTER. The number of edge colorings with no monochromatic triangle, *J. Graph Theory* **21** (1996), no. 4, 441–452.