

Oriented Ramsey numbers of graded digraphs

Patryk Morawski*

Yuval Wigderson†

Abstract

We show that any graded digraph D on n vertices with maximum degree Δ has an oriented Ramsey number of at most $C^\Delta n$ for some absolute constant $C > 1$, improving upon a recent result of Fox, He, and Wigderson. In particular, this implies that oriented grids in any fixed dimension have linear oriented Ramsey numbers, and gives a polynomial bound on the oriented Ramsey number of the hypercube.

We also show that this result is essentially best possible, in that there exist graded digraphs on n vertices with maximum degree Δ such that their oriented Ramsey number is at least $c^\Delta n$ for some absolute constant $c > 1$.

1 Introduction

1.1 Background

The Ramsey number $r(G)$ of a graph G is the minimum number N such that every 2-coloring of the edges of the complete graph K_N on N vertices contains a monochromatic copy of G . Studying how $r(G)$ depends on G is the main question in graph Ramsey theory, and is one of the most studied topics in combinatorics. In particular, one would like to obtain comparable upper and lower bounds on $r(G)$ in terms of basic parameters of G . Such bounds are known in certain regimes; for example, it is known that if G has n vertices, then $r(G)$ grows exponentially in n if and only if G has $\Omega(n^2)$ edges [11, 44].

In particular, sparse graphs—those with $o(n^2)$ edges—have subexponential Ramsey numbers. However, it was observed very early in the history of graph Ramsey theory (see e.g. [18]) that certain sparse graphs, such as trees and cycles, have *much* smaller Ramsey numbers, namely *linear* in their order. In 1975, Burr and Erdős [6] conjectured that this is in fact true for all sparse graphs. Their notion of sparsity is the degeneracy¹ of the graph, where a graph G is called d -degenerate if every subgraph of G has a vertex of degree at most d , and the degeneracy of G is the least d for which it is d -degenerate. The Burr–Erdős conjecture then states that for every d -degenerate graph G we have that $r(G) \leq C_d |V(G)|$, where C_d is a constant that depends only on d . The first major step towards proving this conjecture was made by Chvátal, Rödl, Szemerédi and Trotter [8], who proved that that for every graph G with maximum degree Δ , we have $r(G) \leq C_\Delta |V(G)|$, that is, that the conjecture holds under the stronger assumption of bounded maximum degree. After a sequence of further partial results (e.g. [7, 23, 28, 35, 36]), the full Burr–Erdős conjecture was finally resolved by Lee [39] in 2017.

*Department of Computer Science, ETH Zürich, Zürich, Switzerland. Email: pmorawski@student.ethz.ch.

†Institute for Theoretical Studies, ETH Zürich, Zürich, Switzerland. Email: yuval.wigderson@eth-its.ethz.ch. Supported by Dr. Max Rössler, the Walter Haefner Foundation, and the ETH Zürich Foundation.

¹In fact, they originally stated their conjecture in terms of the *arboricity* of G , but it is easy to verify that a graph has bounded degeneracy if and only if it has bounded arboricity. Most subsequent papers on this topic use degeneracy, which seems more well-suited for Ramsey-theoretic study.

For bounded-degree graphs, we have a fairly precise understanding of how large their Ramsey numbers are. Substantially improving on the early result of Chvatál, Rödl, Szemerédi, and Trotter [8], Graham, Rödl, and Ruciński [28] proved that every n -vertex graph G with maximum degree Δ satisfies $r(G) \leq C^{\Delta(\log \Delta)^2} n$, where C is an absolute constant. In a subsequent paper [29], they noted that their technique yields a better bound of $r(G) \leq C^{\Delta \log \Delta}$ in case G is bipartite, and also showed that this bound is close to best possible, in that there exist bipartite n -vertex graphs maximum degree Δ satisfying $r(G) \geq c^{\Delta} n$, for another absolute constant $c > 1$. Subsequent work by Conlon [10], Fox–Sudakov [22], and Conlon–Fox–Sudakov [12] removed one logarithmic factor from both upper bounds; in particular, it is now known that every n -vertex bipartite graph with maximum degree Δ satisfies $r(G) \leq C^{\Delta} n$, which is best possible up to the value of C .

We now turn our attention to directed graphs (*digraphs* for short), where similar questions can be asked, and which are the main focus of this paper. For an acyclic² digraph D , we define the *oriented Ramsey number* $\vec{r}(D)$ as the minimum number N such that every tournament on N vertices, that is, every edge-orientation of the complete graph K_N , contains a copy of D . The study of oriented Ramsey numbers was initiated in 1951 by Stearns [43], who showed that for a transitive tournament \vec{T}_n on n vertices we have $\vec{r}(\vec{T}_n) \leq 2^{n-1}$, which was complemented by a lower bound of $\vec{r}(\vec{T}_n) \geq 2^{\frac{n}{2}-1}$ by Erdős and Moser [20] in 1964. As in the undirected setting, for sparser digraphs D this number is in general much smaller than exponential in $|V(D)|$; for example, for $n > 8$ every orientation of the n -vertex path has oriented Ramsey number equal to n [33, 45]. For more general oriented trees, Sumner conjectured that $\vec{r}(T) \leq 2n - 2$ for any oriented tree T on n vertices. Sumner’s conjecture has attracted a great deal of interest over the years (e.g. [2, 3, 16, 17, 30, 31, 32, 37, 45]) and in 2011 it was proved for sufficiently large n by Kühn, Mycroft, and Osthus [38].

Motivated by these results, Bucić, Letzter and Sudakov [5] asked whether a natural analogue of the Burr–Erdős conjecture holds for acyclic digraphs, that is, whether for all acyclic digraphs D with maximum degree³ Δ we have that $\vec{r}(D) \leq c_{\Delta} |V(D)|$ for some constant c_{Δ} depending only on Δ . Quite surprisingly, Fox, He, and Wigderson [21] recently answered this question in the negative by showing that for any Δ and large enough n there exists an n -vertex digraph D with maximum degree Δ and $\vec{r}(D) \geq n^{\Omega(\Delta^{2/3}/\log^{5/3} \Delta)}$. In the other direction, they proved an upper bound of $\vec{r}(D) \leq n^{C_{\Delta} \log n}$ for any acyclic digraph D with maximum degree Δ . However, their results leave open the question of whether the worst case behavior for fixed Δ is always polynomial, or whether it can indeed be super-polynomial.

Why are the Ramsey numbers of acyclic digraphs so much larger than analogous Ramsey numbers in the undirected setting? There are a few reasons, but it seems that an important one has to do with chromatic numbers. In the undirected setting, a graph of bounded maximum degree also has bounded chromatic number, a fact which is often used in upper bounds on $r(G)$; this is also (roughly) why it is possible to prove stronger bounds if one assumes that G is bipartite. However, in the directed setting, the correct analogue of the chromatic number seems to be the *height* of D , which is defined as the number of vertices in the longest oriented path in D , or equivalently as the minimum number of parts in a partition $V(D) = V_1 \cup \dots \cup V_h$ with the property that all edges are directed from V_i to V_j for some $i < j$. As shown by simple examples such as oriented paths, acyclic digraphs can have bounded maximum degree but unbounded height, and this manifests itself in the fact that acyclic digraphs can have much larger Ramsey numbers than their undirected counterparts. In fact, Fox, He, and Wigderson [21] proved that if D has maximum degree Δ and height h , then $\vec{r}(D) \leq C_{\Delta, h} n$ for some constant $C_{\Delta, h}$ depending only on Δ and h , showing that an analogue of the Burr–Erdős conjecture is true for digraphs

²A digraph is *acyclic* if it contains no oriented cycles.

³By the maximum degree of a digraph, we mean the maximum degree of the underlying undirected graph.

of bounded height. However, it is not the case that D having large height necessarily implies that $\vec{r}(D)$ is large. Instead, it appears that what really matters is the *multiscale complexity* of D —roughly speaking, how many edge length scales D has (we refer to [21, Section 1] for more details). In particular, Fox, He and Wigderson [21] showed that even if D has unbounded height, its oriented Ramsey number is polynomial in $|V(D)|$ if it satisfies certain other restrictions on the distribution of edges. An important class of digraphs for which they proved such a result is the class of *graded* digraphs, which we now define.

Definition 1.1. We say that a digraph D is *graded* with a *graded partition* $V(D) = V_1 \cup \dots \cup V_h$ if every edge of D is directed from V_i to V_{i+1} for some $i \in [h-1]$.

We remark that graded digraphs are necessarily acyclic, and that the graded partition is unique assuming that the underlying graph of D is connected. In particular, the number h of parts in the graded partition is equal to the height of D .

There are many natural examples of graded digraphs. For example, the d -dimensional *grid digraph* $\vec{\Gamma}_{d,k}$ whose vertex set is $[k]^d$ and whose edges are all ordered pairs of the form

$$((x_1, \dots, x_i, \dots, x_d), (x_1, \dots, x_i + 1, \dots, x_d)) \quad \text{for some } i \in [d]$$

is graded; one obtains the graded partition by setting $V_i := \{(x_1, \dots, x_d) \in [k]^d : x_1 + \dots + x_d = i\}$. An important special case of this construction is the *oriented hypercube* \vec{Q}_d , which is obtained from the unoriented hypercube graph by directing all edges away towards the positive orthant. More generally, the Hasse diagram of any graded poset is a graded digraph.

Fox, He, and Wigderson [21] proved the following upper bound on oriented Ramsey numbers of graded digraphs.

Theorem 1.2 ([21, Theorem 1.5]). *If D is a graded digraph with n vertices, height h , and maximum degree Δ , then*

$$\vec{r}(D) \leq h^{10\Delta \log_2 \Delta} n.$$

In particular, as $h \leq n$, we have $\vec{r}(D) \leq n^{11\Delta \log_2 \Delta}$.

1.2 Main results

In this paper, we show a much stronger bound on the oriented Ramsey number of graded digraphs. In fact, we show that for any fixed Δ , any graded digraph with maximum degree Δ has an oriented Ramsey number that is linear in its order.

Theorem 1.3. *If D is a graded digraph on n vertices with maximum degree Δ , then*

$$\vec{r}(D) \leq 10^9 \Delta^3 2^{4\Delta} n.$$

More precisely, if D has maximum in-degree Δ^- and maximum out-degree Δ^+ , then

$$\vec{r}(D) \leq 10^9 \Delta^+ (\Delta^-)^2 2^{4\Delta^-} n.$$

We stress that the height of D does not affect the bound in Theorem 1.3 at all. This is somewhat surprising, given the intuition above that the height of a digraph should play an analogous role to the chromatic number of a graph, and should in turn affect the oriented Ramsey number. Nonetheless, an understanding that arises from our techniques is that graded digraphs “behave like” bipartite graphs, regardless of their height.

As an immediate corollary of Theorem 1.3, we obtain a linear upper bound on the oriented Ramsey numbers of grid digraphs in any fixed dimension, since $\vec{\Gamma}_{d,k}$ has maximum in- and out-degree equal to d .

Corollary 1.4. *For any $d \geq 1$, there exists a constant $C_d = 10^9 d^3 2^{4d}$ such that the d -dimensional grid digraph $\vec{\Gamma}_{d,k}$ satisfies $\vec{r}(\vec{\Gamma}_{d,k}) \leq C_d |V(\vec{\Gamma}_{d,k})|$.*

We remark that there has recently been a great deal of interest in Ramsey- and Turán-type questions involving grid graphs, see e.g. [4, 9, 14, 25, 26, 27, 34, 40]. In particular, it is proved in [40, Corollary 1.4] that the *undirected* Ramsey number of the two-dimensional $k \times k$ grid graph is $(\frac{3}{2} + o(1))k^2$.

At the other extreme, where k is fixed and d tends to infinity, we obtain a polynomial bound. For example, for the oriented hypercube \vec{Q}_d , Theorem 1.3 implies that $\vec{r}(\vec{Q}_d) \leq 2^{5d+o(d)} = |V(\vec{Q}_d)|^{5+o(1)}$. By optimizing our techniques, we are able to improve the exponent from 5 to $\log_2(17) \approx 4.09$.

Theorem 1.5. *There exists an absolute constant $C > 0$ such that $\vec{r}(\vec{Q}_d) \leq Cd^3 17^d$.*

In fact, motivated by the example of the hypercube, in Section 2 we prove a strengthening of Theorem 1.3, which gives a better bound for graded digraphs where the large in-degrees are only in the parts of the graded partition where the number of vertices is small. Such a result is useful for \vec{Q}_d , since in the hypercube, almost all vertices (and in particular those vertices lying in the very large parts of the graded partition) have in- and out-degree close to $d/2$.

In the undirected setting, it is a major open problem to determine $r(Q_d)$. A famous conjecture of Burr and Erdős [6] from 1975 is that the Ramsey number of the hypercube is linear in its order, i.e. that $r(Q_d) = O(2^d)$. This question has been intensively studied (see e.g. [1, 13, 22, 29, 41, 42]); the current best known bound is due to Tikhomirov [46], who proved that $r(Q_d) \leq 2^{(2-\varepsilon)d}$, where $\varepsilon > 0$ is some small absolute constant. However, the results in the undirected setting cannot be used directly to obtain upper bounds on $\vec{r}(\vec{Q}_d)$, and to the best of our knowledge Theorem 1.5 is the first known polynomial bound on $\vec{r}(\vec{Q}_d)$. Instead, the techniques from the undirected setting naturally yield polynomial upper bounds on the oriented Ramsey number of the *bipartite orientation* of the hypercube, where all edges are directed from vertices of even to odd Hamming weight, rather than the more natural orientation \vec{Q}_d .

Our second main result shows that Theorem 1.3 is close to best possible, in the sense that there exist graded digraphs with maximum degree Δ and oriented Ramsey number of at least $c^\Delta |V(D)|$ for some absolute constant $c > 1$. A very similar result was proved in the undirected setting by Graham, Rödl and Ruciński [29]; in fact, it is not hard to adapt their construction to show the existence of such a D which is *bipartite*, that is, of height 2. However, by modifying their construction appropriately, we are able to prove such a result for D of arbitrary height, where the lower bound again does not depend on the height.

Theorem 1.6. *There exist constants $c > 1$ and Δ_0 such that for all $\Delta \geq \Delta_0$, $n \geq \Delta$, and $h \geq 2$ there exists a graded digraph D with maximum degree Δ and with a graded partition $V(D) = V_1 \cup \dots \cup V_h$ such that $|V_i| \leq n$ for all $i \in [h]$, such that $\vec{r}(D) > c^\Delta hn \geq c^\Delta |V(D)|$.*

While our proof of Theorem 1.6 is fairly closely inspired by the techniques of Graham–Rödl–Ruciński [29], there are a number of techniques that go into the proof of Theorem 1.3. The three main ingredients are the dependent random choice technique, the median order of a tournament, and an embedding argument based on the Lovász local lemma. Dependent random choice is a very powerful technique which has been used in many of the recent breakthroughs in the Ramsey theory of sparse graphs; we refer to the survey [24] for more information. The median order is an elementary, yet surprisingly powerful, tool in the study of tournaments, and has been instrumental in many of the recent advances on Sumner’s conjecture and related questions; see e.g. [32] for more details. The idea of combining these two techniques⁴ was

⁴Strictly speaking, [15] does not use dependent random choice, but rather a simpler Kővári–Sós–Turán-type argument. Nonetheless, the high-level approach is closely related.

already used in [15], where an upper bound on the oriented Ramsey number of the power of an oriented path is proved; at a high level, this digraph is embedded in an arbitrary tournament by repeatedly embedding small pieces, and each small piece is embedded by combining dependent random choice with properties of the median order. We apply a similar technique, but as we are embedding an arbitrary graded digraph, more care is needed in the embedding step. To achieve this, we embed each part of D randomly, and ensure that the embedding is successful thanks to the Lovász local lemma; this is inspired by a similar argument in [13], upper-bounding the Ramsey numbers of bounded-degree bipartite graphs. However, as our digraphs may have arbitrary height, we need to ensure at every embedding step both that we are consistent with the past, and that we are well-equipped to apply the local lemma at the next step, which adds some further complications. We defer a more detailed proof sketch to Section 2.1.

The remainder of the paper is organized as follows. We prove (a strengthening of) Theorem 1.3 in Section 2, which begins with a detailed proof outline. Similarly, in Section 3, we first give a sketch of the proof of Theorem 1.6, then prove it for height-2 digraphs, and finally prove it in full generality. We end in Section 4 with some concluding remarks and open problems.

Notation: For a directed graph D , we use $V(D)$ to denote its vertex set and $E(D)$ to denote its edge set, which is a collection of ordered pairs of elements of $V(D)$. For a vertex $v \in V(D)$ we write $N_D^+(v)$ and $N_D^-(v)$ for its out- and in-neighborhood and $d_D^+(v)$ and $d_D^-(v)$ for its out- and in-degree, respectively. For a subset $U \subseteq V(D)$ we let $N_D^+(U)$ denote the common out-neighborhood of U and similarly $N_D^-(U)$ the common in-neighborhood of U . Additionally, for a vertex $v \in V(D)$ we let $d_D^+(v, U) = |N_D^+(v) \cap U|$ and $d_D^-(v, U) = |N_D^-(v) \cap U|$ be the out- and in-degree of v into U . Whenever the digraph D is clear from context, we omit the subscript and write $d^+(v)$ for $d_D^+(v)$, etc. Throughout the paper, we omit ceilings and floors whenever they are not crucial.

2 Proof of Theorem 1.3

As stated above, instead of proving Theorem 1.3, we will prove a stronger theorem that leverages the fact that the graded digraph could locally have different maximum in-degrees in different parts. This will allow us to use less overhead to embed the parts where the in-degree is small, and is particularly useful for graded digraphs in which the large in-degree only appears in parts of the graded partition whose sizes are small, as is the case for the oriented hypercube \vec{Q}_d . More formally, we prove the following theorem, which immediately implies Theorem 1.3 by upper-bounding each Δ_i^- by Δ^- .

Theorem 2.1. *Let D be a graded digraph on n vertices with a graded partition $V(D) = V_1 \cup \dots \cup V_h$ for some $h \in \mathbb{N}$ and maximum in- and out-degree Δ^- and Δ^+ respectively. Moreover, for each $i \in [h-1]$ let Δ_i^- be the maximum in-degree in the induced subgraph $D[V_i \cup V_{i+1}]$ and set $\Delta_0^- = \Delta_h^- = 0$. Then*

$$\vec{r}(D) \leq 10^9 (\Delta^-)^2 \Delta^+ \sum_{i=1}^h 2^{2(\Delta_{i-1}^- + \Delta_i^-)} |V_i|.$$

Before proceeding with the proof of Theorem 2.1, we note that it immediately yields Theorem 1.5, the upper bound on the oriented Ramsey number of the hypercube.

Proof of Theorem 1.5. Let $d \in \mathbb{N}$ and \vec{Q}_d be the d -dimensional hypercube on the vertex set $V = \{0, 1\}^d$. Moreover, for each $0 \leq i \leq d$, let $V_i := \{v \in V : \sum_j v_j = i\}$. Then $V = V_0 \cup \dots \cup V_d$

is a graded partition of \vec{Q}_d and for each $i = 0, \dots, d-1$ the maximum in-degree in the induced subgraph $\vec{Q}_d[V_i \cup V_{i+1}]$ is $i+1$. Therefore, by Theorem 2.1,

$$\vec{r}(\vec{Q}_d) \leq C' d^3 \sum_{i=0}^d \binom{d}{i} 4 \cdot 2^{4i} = C d^3 (16+1)^d = C d^3 17^d,$$

for some absolute constants C' and C . □

2.1 Proof outline

We now sketch our proof of Theorem 2.1, before proceeding with the technical details. As our goal is upper-bounding $\vec{r}(D)$, we wish to prove that any sufficiently large tournament contains a copy of D , so we fix such a tournament T . We will first find disjoint subsets A_1, \dots, A_h such that almost all (Δ_{i-1}^-) -subsets of A_{i-1} have many out-neighbors in A_i . To achieve this, we will use the notion of a median order and then apply the dependent random choice technique. A *median order* is an ordering v_1, \dots, v_N of vertices of T that maximizes the number of forward edges, that is, edges $v_i v_j$ with $i < j$. Given a median order v_1, \dots, v_N we write $[i, j]$ for $\{v_i, \dots, v_j\}$ and similarly $[i, j]$ for $\{v_i, \dots, v_j\}$. We note that in a median order for any v_i and $j < i$ at least half of the vertices $[j, i)$ are in-neighbors of v_i , since otherwise we could move v_i to the front of the interval to obtain an ordering with more forward edges.

To find our sets A_1, \dots, A_h , we will place A_h at the end of a median order of T and then inductively find A_{h-1}, A_{h-2} , and so on, all the way down to A_1 . At each step, we will make sure that the last A_i found is contained within some small interval I_i in the median order. To find A_{i-1} we will first find an interval I_{i-1} so that almost half of the edges go from I_{i-1} to A_i . Using properties of the median order, we can always find I_{i-1} within a bigger interval I'_{i-1} , which immediately precedes I_i and is only larger than I_i by a fixed multiplicative factor. However, we will also ensure that $|I_{i-1}| = |I_i|$, so that this multiplicative factor does not stack up as we proceed with the next layers. Given such I_{i-1} and A_i , a standard application of the dependent random choice technique allows us to construct $A_{i-1} \subseteq I_{i-1}$ with the property that most (Δ_{i-1}^-) -subsets of A_{i-1} have many common out-neighbors in A_i .

Having found all the sets A_1, \dots, A_h , we now want to embed each V_i into the corresponding A_i . We will start this time with V_1 and after having embedded the first $i-1$ layers we will use the Lovász Local Lemma to embed the i th layer. The added difficulty in this step is that we wish to maintain the ability to keep this process going; therefore, when embedding the i th layer, we will also make sure that each vertex of the $(i+1)$ st layer we will still have many possible embeddings in A_{i+1} . This allows us to then reuse the local lemma argument in the next layer as well.

We now proceed with the details of the proof, which is split across the next three subsections. The basic dependent random choice lemma, which we will iterate to construct the sets A_i , is stated and proved in Section 2.2. The basic embedding lemma, which allows us to embed V_i into A_i while being in the position to continue the process in step $i+1$, is given in Section 2.3. We combine these two tools in Section 2.4, where we prove Theorem 2.1 via two inductive arguments: first to find the sets A_1, \dots, A_h , and then the second to embed our digraph D into them. While this is essentially just a combination of the two main lemmas, some care is needed to ensure that the dependencies between the various parameters work out properly.

2.2 Dependent random choice

In this section, we prove a lemma that will allow us to find the subsets A_1, \dots, A_h as described above. While the statement may seem daunting because of all of the parameters, the basic idea

is the simple one described above: given a set B contained in an interval of the median ordering, we are able to find a set A , contained in a nearby interval of the same size, such that A is large and such that most (Δ^-) -subsets of A have many common out-neighbors in B .

Lemma 2.2. *Let $a, a', b, \ell, s, N, k, \Delta^- \in \mathbb{N}$ be integers with $2a' \geq a$, and let T be a tournament on N vertices with a median order v_1, \dots, v_N . Moreover, let $2ka' < j \leq N - a + 1$ and let $B \subseteq [j, j + a)$ be an arbitrary subset of size at least b . Then there exist $j - 2ka' \leq j' \leq j - a'$ and $A \subseteq [j', j' + a')$ such that*

- $|A| \geq \frac{a'}{2} \left(\frac{k-1}{2k}\right)^\ell$, and
- for all but at most $4 \binom{2k}{k-1}^\ell \binom{a'}{\Delta^-} \left(\frac{s}{b}\right)^\ell$ subsets $S \subseteq A$ with $|S| = \Delta^-$, we have $|N^+(S) \cap B| \geq s + 1$.

Proof. Note that we may assume $|B| = b$, because the conclusion for any larger value of $|B|$ is a strictly weaker statement. Let $J = [j - 2ka', j)$ and notice that since v_1, \dots, v_N is a median order of the vertices of T at least half of the vertices from $[j - 2ka', i)$ are in-neighbors of v_i for each v_i . Therefore, for every $v_i \in [j, j + a)$, we have

$$d^-(v_i, J) = d^-(v_i, [j - 2ka', i)) - d^-(v_i, [j, i)) \geq \frac{2ka' + i - j}{2} - (i - j) = \frac{2ka' - (i - j)}{2} \geq (k-1)a',$$

where the final step uses that $i - j \leq a \leq 2a'$. Therefore, since $B \subseteq [j, j + a)$, we have

$$\sum_{u \in J} d^+(u, B) = \sum_{v \in B} d^-(v, J) \geq b(k-1)a'.$$

Now, for $i \in [2k]$ let $I_i = [j + (i - 2k - 1)a', j + (i - 2k)a')$ and notice that $J = \bigcup_{i \in [2k]} I_i$. Therefore, by the pigeonhole principle there exists an i such that

$$\sum_{u \in I_i} d^+(u, B) \geq b \left(\frac{k-1}{2k}\right) a'.$$

Fix such an i , and let $I = I_i$. Let $j' = j + (i - 2k - 1)a'$ be the left endpoint of I .

Pick now a set L of ℓ vertices from B uniformly at random with repetitions, let $M = \{u \in I : L \subseteq N^+(u) \cap B\}$, and write $X = |M|$. Jensen's inequality then implies that

$$\mathbb{E}[X] = \sum_{u \in I} \left(\frac{d^+(u, B)}{b}\right)^\ell \geq a' \cdot \left(\frac{k-1}{2k}\right)^\ell.$$

Let Y be the random variable counting the number of subsets of M of size Δ^- with at most s common out-neighbors in B . For a given such $S \subseteq I$, the probability that it is a subset of M is at most $\left(\frac{|N^+(S) \cap B|}{b}\right)^\ell \leq \left(\frac{s}{b}\right)^\ell$ and therefore

$$\mathbb{E}[Y] \leq \binom{a'}{\Delta^-} \left(\frac{s}{b}\right)^\ell.$$

Let us suppose for the moment that $\mathbb{E}[Y] > 0$. By linearity of expectation we know that

$$\mathbb{E} \left[X - \frac{\mathbb{E}[X]}{2\mathbb{E}[Y]} Y - \frac{\mathbb{E}[X]}{2} \right] = 0.$$

Therefore, there exists a choice of L and the corresponding M for which this expression is non-negative. Fix such a choice. Then

$$|M| = X \geq \frac{\mathbb{E}[X]}{2} \geq \frac{a'}{2} \cdot \left(\frac{k-1}{2k}\right)^\ell,$$

and, since $X \leq |I| = a'$

$$Y \leq \frac{2X}{\mathbb{E}[X]} \mathbb{E}[Y] \leq 4 \left(\frac{2k}{k-1}\right)^\ell \left(\frac{a'}{\Delta^-}\right) \left(\frac{s}{b}\right)^\ell.$$

In particular, we may set $A = M$ and conclude the proof.

It remains to consider the case that $\mathbb{E}[Y] = 0$. In this case, as Y is a non-negative random variable, it must take on the value 0 with probability 1. Thus, if we select any L for which the corresponding X is at least $\mathbb{E}[X]$, we can again set $A = M$ and conclude the proof. \square

2.3 Finding a good embedding

In this section we prove a lemma that will allow us to embed each part V_i of the graded partition into the respective A_i . For this, we will use the Lovász local lemma [19], whose statement we now recall.

Lemma 2.3. *Let A_1, \dots, A_n be events in an arbitrary probability space and let $H = ([n], E)$ be a graph such that for each $i \in [n]$ the event A_i is mutually independent of the events $\{A_j : (i, j) \notin E\}$. Suppose moreover that $0 \leq x_1, \dots, x_n < 1$ are real numbers such that for all $i \in [n]$ we have $\Pr[A_i] \leq x_i \prod_{(i,j) \in E} (1 - x_j)$. Then*

$$\Pr \left[\bigwedge_{i=1}^n \overline{A_i} \right] \geq \prod_{i=1}^n (1 - x_i) > 0.$$

We are now ready to state and prove our embedding lemma. At a high level, it says the following. We wish to embed a bipartite digraph $D = (V_1 \cup V_2, E)$ into a tournament T . Eventually, the bipartite digraph will be the induced subdigraph on two consecutive layers of the graded partition. As such, we have target sets $A, B \subseteq V(T)$, and would like to embed the V_1 into A and V_2 into B . Eventually, these sets will come from our list A_1, \dots, A_h , so we may assume that most (Δ^-) -subsets of A have many common neighbors in B , which is of course convenient for successfully embedding D . However, because we wish to be consistent with the previous steps of the embedding, we cannot embed V_1 arbitrarily into A ; rather, we are given a list of options $f(v) \subseteq A$ for each vertex $v \in V_1$, and must embed v into one of the options from $f(v)$. Moreover, we want to ensure that, having embedded V_1 , there are actually many ways to embed each vertex of V_2 into B ; this is necessary to ensure that at the next stage of the embedding, the lists $f(v)$ remain sufficiently large to apply the same lemma again. The next lemma simply states that, as long as the relevant parameters are related in the appropriate ways, such an embedding is possible. It is proved by embedding each v into a uniformly random choice in $f(v)$, and using the local lemma to show that we succeed with positive probability.

Lemma 2.4. *Let $D = (V_1 \cup V_2, E)$ be a bipartite digraph with maximum in- and out-degree $\Delta^- \geq 1$ and Δ^+ , respectively, with all edges oriented from V_1 to V_2 . Let $a, b, c \in \mathbb{N}$ be integers and $\delta > 0$ be a parameter such that $a \geq b \geq 32|V_1|$ and $\delta \leq (2^{-1/2} \frac{b}{a})^{\Delta^-} / (4\Delta^+ \Delta^-)$. Let T be a tournament and $A, B \subseteq V(T)$ be subsets of its vertices such that $|A| = a$ and for all but at most a δ -fraction of subsets $S \subseteq A$ of size Δ^- we have $|N_T^+(S) \cap B| \geq c$.*

Let $f : V_1 \rightarrow 2^A$ be a function such that for each $v \in V_1$ we have $|f(v)| \geq b$. Then there exists an injective function $\phi : V_1 \rightarrow A$ such that for each $v \in V_1$ we have $\phi(v) \in f(v)$ and for each $u \in V_2$ we have $|N_T^+(\phi(N_D^-(u))) \cap B| \geq c$.

Proof. Let $s = |V_1|$ and let $\phi : V_1 \rightarrow A$ be a random mapping such that for each $v \in V_1$ the value of $\phi(v)$ is picked uniformly at random from the set $f(v)$, with all these choices made independently. For all distinct $v, w \in V_1$, let A_{vw} be the event that $\phi(v) = \phi(w)$. Moreover, for each $u \in V_2$ let us write $N_u = N_D^-(u)$ and let B_u be the event that $|\phi(N_u)| = N_u$ but $|N_T^+(\phi(N_u)) \cap B| < c$. Clearly, if none of the bad events A_{vw} and B_u hold, then ϕ satisfies the requirements of the lemma.

Let us now bound the probabilities for each of these events to hold. For $v, w \in V_1$ we have $\Pr[A_{vw}] \leq \frac{1}{b}$. To bound $\Pr[B_u]$, let us suppose that $|\phi(N_u)| = |N_u|$ but $|N_T^+(\phi(N_u)) \cap B| < c$. If we denote $\ell = |N_u|$, this implies that $|N_T^+(S) \cap B| < c$ for all $\binom{a-\ell}{\Delta^--\ell}$ (Δ^-)-subsets S of A containing $\phi(N_u)$. Moreover, the number of pairs consisting of a bad (Δ^-)-subset $S \subseteq A$ and an ℓ -subset of S is at most $\delta \binom{a}{\Delta^-} \binom{\Delta^-}{\ell}$. Therefore, the number of tuples $(\phi(v))_{v \in N_u} \in A^\ell$ for which the event B_u holds is at most

$$\frac{\delta \binom{a}{\Delta^-} \binom{\Delta^-}{\ell} \ell!}{\binom{a-\ell}{\Delta^- - \ell}} = \delta \binom{a}{\ell} \ell! \leq \delta a^\ell.$$

Since $(\phi(v))_{v \in N_u}$ is chosen uniformly at random from a subset of A^ℓ of size at least b^ℓ , this implies that $\Pr[B_u] \leq \frac{\delta a^\ell}{b^\ell} \leq \delta \left(\frac{a}{b}\right)^{\Delta^-}$.

Let us now consider the dependencies between the bad events. Note that since the random variables $\{\phi(v)\}_{v \in V_1}$ are mutually independent, the event A_{vw} is mutually independent from all $A_{v'w'}$ and B_u such that $\{v, w\} \cap \{v', w'\} = \emptyset$ and $\{v, w\} \cap N_u = \emptyset$, respectively. Thus, A_{vw} is dependent on at most $2(s-2) < 2s$ events $A_{v'w'}$ and at most $2\Delta^+$ events B_u . Similarly, the event B_u is mutually independent from all $A_{v,w}$ and $B_{u'}$ such that $N_u \cap \{v, w\} = \emptyset$ and $N_u \cap N_{u'} = \emptyset$, respectively, and thus it is dependent on at most $\binom{s}{2} - \binom{s-|N_u|}{2} \leq \frac{2s\Delta^- - (\Delta^-)^2 - \Delta^-}{2} < s\Delta^-$ events A_{vw} and at most $\Delta^+ \Delta^-$ events $B_{u'}$.

We now want to apply the local lemma. For each A_{vw} , we let the corresponding x_i be $x = \frac{8}{b}$ and for each B_u we let the corresponding x_i be $y = \frac{1}{2\Delta^+ \Delta^-}$. Since $\Delta^- \geq 1$ and $b \geq 32s$ we then get

$$x(1-x)^{2s}(1-y)^{2\Delta^+} \geq \frac{8}{b} 4^{-16s/b} 4^{-1/\Delta^-} \geq \frac{1}{b} \geq \Pr[A_{vw}],$$

where we use the inequality $1-z \geq 4^{-z}$, valid for all $0 \leq z \leq \frac{1}{2}$. Similarly, we have

$$y(1-y)^{\Delta^+ \Delta^-} (1-x)^{s\Delta^-} \geq \frac{1}{2\Delta^+ \Delta^-} 4^{-1/2} 4^{-8s\Delta^-/b} \geq \frac{1}{4\Delta^+ \Delta^-} 4^{-\Delta^-/4} = \left(\frac{a}{b}\right)^{\Delta^-} \delta \geq \Pr[B_u],$$

where the equality is by our choice of δ . Therefore, by Lemma 2.3, the probability that none of the bad events A_{vw} and B_u hold is positive and thus, there exists a choice of ϕ satisfying the desired properties. \square

2.4 Proof of Theorem 2.1

Given Lemmas 2.2 and 2.4, we are now ready to prove Theorem 2.1. We recall that the high-level idea of the proof is to inductively construct sets A_h, \dots, A_1 using Lemma 2.2, and then to embed D into these sets inductively using Lemma 2.4. Essentially all that remains is to define a large number of parameters, and to check that they satisfy certain inequalities so that Lemmas 2.2 and 2.4 can indeed be applied.

Proof of Theorem 2.1. Let D be a graded graph on n vertices with a graded partition $V(D) = V_1 \cup \dots \cup V_h$ for some $h \in \mathbb{N}$ and let Δ^+ and Δ^- be its maximum out- and in-degree, respectively. Moreover, for each $i \in [h-1]$ let Δ_i^- be the maximal in-degree in the induced subgraph $D[V_i \cup V_{i+1}]$ and set $\Delta_0^- = \Delta_h^- = 0$. Note that we may assume that $\Delta_i^- \geq 1$ for all $i \in [h-1]$, for otherwise the underlying graph of D is disconnected, and we obtain the desired bound on $\vec{r}(D)$ by summing up $\vec{r}(D')$ for every connected component D' of D .

We define $\varepsilon = 2/\Delta^-$ and $k = 4\Delta^- + 4$, and note that $\frac{2k}{k-1} \leq 2 + \varepsilon$. To find the subsets A_h, \dots, A_1 as described above we first need to fix the relevant parameters. We define integers c_s , c_b , and c_a by

$$c_s = 32, \quad c_b = 2000\Delta^+\Delta^-c_s, \quad \text{and} \quad c_a = 2c_b.$$

We further define

$$n_i = \sum_{j=i}^h \frac{(2+\varepsilon)^{2(\Delta_{j-1}^- + \Delta_j^-)} |V_j|}{2^{j-i}},$$

and let

$$a_i = c_a n_i \quad \text{and} \quad b_i = c_b (2+\varepsilon)^{-2\Delta_i^-} n_i.$$

In building the sets A_h, \dots, A_1 , we will ensure that $|A_i| \geq b_i$ and that A_i lies in an interval of length a_i . We now let

$$s_i = c_s (2+\varepsilon)^{-2(\Delta_i^- + \Delta_{i-1}^-)} n_i \quad \text{and} \quad \delta_i = \frac{1}{4\Delta^- \Delta^+} \left(2^{-1/2} \frac{s_{i+1}}{b_{i+1}} \right)^{\Delta_i^-}.$$

We will further guarantee that all but a δ_i -fraction of the (Δ_i^-) -subsets of A_i have at least s_{i+1} common out-neighbors in A_{i+1} . Finally, we let

$$o_i = 2k \sum_{j=i}^h a_j.$$

Throughout the process, we will ensure that A_i lies within the last o_i vertices of the median order. We note for future reference that

$$n_i = (2+\varepsilon)^{2(\Delta_{i-1}^- + \Delta_i^-)} |V_i| + \frac{1}{2} n_{i+1}.$$

In particular, this implies that $a_i \geq a_{i+1}/2$ and that $s_i \geq c_s |V_i| = 32|V_i|$.

Let $N = o_1$, let T be a tournament on N vertices, and fix a median order v_1, \dots, v_N of T . We now claim that we can find integers j_1, \dots, j_h and disjoint sets $A_1, \dots, A_h \subseteq V(T)$ satisfying the following properties.

- $j_i < j_{i+1}$,
- $j_i \geq N - o_i$,
- $A_i \subseteq [j_i, j_i + a_i)$,
- $|A_i| \geq b_i$, and
- for each $i \in [h-1]$, for all but at most a δ_i -fraction of subsets $S \subseteq A_i$ of size Δ_i^- , we have that $|N_T^+(S) \cap A_{i+1}| \geq s_{i+1}$.

We start by setting $j_h = a_h$ and $A_h = [N - a_h + 1, N]$, which clearly satisfy these properties. Suppose now that for some $i \in [h - 1]$ we have defined $A_{i'}$ and $j_{i'}$ for all $i < i' \leq h$. By applying Lemma 2.2 with $j = j_{i+1}$, $B = A_{i+1}$, $a = a_{i+1}$, $a' = a_i \geq a_{i+1}/2$, $b = b_{i+1}$, $\ell = 2\Delta_i^-$, $s = s_{i+1}$, $k = k$ and $\Delta^- = \Delta_i^-$ we can find $j_{i+1} > j_i \geq j_{i+1} - 2ka' \geq N - o_{i+1}$ and $A_i \subseteq [j_i, j_i + a_i]$ such that

- $|A_i| \geq \frac{a_i}{2}(2 + \varepsilon)^{-2\Delta_i^-} = b_i$, and
- at most $4 \cdot (2 + \varepsilon)^{2\Delta_i^-} \binom{a_i}{\Delta_i^-} \binom{s_{i+1}}{b_{i+1}}^{2\Delta_i^-}$ subsets $S \subseteq A_i$ of size Δ_i^- have fewer than s_{i+1} common out-neighbors in A_{i+1} .

We now note that, since $b_i \geq 2\Delta_i^-$, we have that

$$\frac{\binom{a_i}{\Delta_i^-}}{\binom{b_i}{\Delta_i^-}} \leq 2^{\Delta_i^-} \left(\frac{a_i}{b_i} \right)^{\Delta_i^-} = \left(2 \cdot \frac{c_a(2 + \varepsilon)^{2\Delta_i^-}}{c_b} \right)^{\Delta_i^-} = \left(4(2 + \varepsilon)^{2\Delta_i^-} \right)^{\Delta_i^-},$$

where we plug in our definitions of a_i, b_i, c_a , and c_b . Additionally, we have that

$$\left(\frac{s_{i+1}}{b_{i+1}} \right)^{\Delta_i^-} = \left(\frac{c_s(2 + \varepsilon)^{2\Delta_{i+1}^-}}{c_b(2 + \varepsilon)^{2(\Delta_{i+1}^- + \Delta_i^-)}} \right)^{\Delta_i^-} = \left(\frac{c_s}{c_b(2 + \varepsilon)^{2\Delta_i^-}} \right)^{\Delta_i^-} = \left(\frac{1}{2000\Delta + \Delta^- (2 + \varepsilon)^{2\Delta_i^-}} \right)^{\Delta_i^-}$$

by our choices of s_{i+1}, b_{i+1}, c_s , and c_b . Putting this all together, we see that the fraction of the (Δ_i^-) -subsets of A_i for which the common out-neighborhood is too small is at most

$$\begin{aligned} \frac{4 \cdot (2 + \varepsilon)^{2\Delta_i^-} \binom{a_i}{\Delta_i^-} \binom{s_{i+1}}{b_{i+1}}^{2\Delta_i^-}}{\binom{b_i}{\Delta_i^-}} &= \left[4(2 + \varepsilon)^{2\Delta_i^-} \left(\frac{s_{i+1}}{b_{i+1}} \right)^{\Delta_i^-} \right] \cdot \left[\frac{\binom{a_i}{\Delta_i^-}}{\binom{b_i}{\Delta_i^-}} \left(\frac{s_{i+1}}{b_{i+1}} \right)^{\Delta_i^-} \right] \\ &\leq \left[4(2 + \varepsilon)^{2\Delta_i^-} \left(\frac{s_{i+1}}{b_{i+1}} \right)^{\Delta_i^-} \right] \left(\frac{1}{500\Delta + \Delta^-} \right)^{\Delta_i^-} \\ &\leq \left(2^{-1/2} \frac{s_{i+1}}{b_{i+1}} \right)^{\Delta_i^-} \left(\frac{4 \cdot 2^{1/2}(2 + \varepsilon)^2}{500\Delta + \Delta^-} \right)^{\Delta_i^-} \\ &\leq \delta_i, \end{aligned}$$

where the final step uses that $\varepsilon \leq 2$, that $\Delta_i^- \geq 1$, and our definition of δ_i . This shows that the set A_i , as defined above, satisfies the desired properties. Continuing inductively in this way, we are able to find all the sets A_1, \dots, A_h .

Having found these sets we now want to embed each V_i into A_i , starting this time with V_1 . For $i \in [h]$, let $S_i = \bigcup_{j=1}^i V_j$. For each $i \in [h]$ we find a function $\phi_i : S_i \rightarrow (A_1 \cup \dots \cup A_i)$ such that

- ϕ_i is an embedding of $D[S_i]$ into $T[A_1 \cup \dots \cup A_i]$,
- $\phi(V_j) \subseteq A_j$ for each $1 \leq j \leq i$, and
- for each $i \in [h - 1]$ and $v \in V_{i+1}$, the vertices $\phi_i(N_D^-(v)) \subseteq A_i$ have at least s_{i+1} common out-neighbors in A_{i+1} .

For $i = 1$, we find ϕ_1 by applying Lemma 2.4 with $a = b = b_1$, $A = A_1$, $B = A_2$, $D = D[V_1 \cup V_2]$, $\delta = \delta_1$ and $f(v) = A_1$ for all $v \in V_1$. Suppose now that for some $i \in [h - 2]$ we have found ϕ_i satisfying the conditions above. For $v \in V_{i+1}$, let $f(v) = N_T^+(\phi_i(N_D^-(v))) \cap A_{i+1}$ so that $|f(v)| \geq s_{i+1}$ for all $v \in V_{i+1}$. We can now again apply Lemma 2.4 with $A = A_{i+1}$, $B = A_{i+2}$, $a = b_{i+1}$, $b = s_{i+1}$, $\delta = \delta_i$ and $D = D[V_{i+1} \cup V_{i+2}]$. We get a function ϕ such that

- for all $v \in V_{i+1}$ we have $\phi(v) \in N_T^+(\phi_i(N_D^-(v)))$, and
- for all $v \in V_{i+2}$ we have $|N_T^+(\phi(N_D^-(v)))| \geq s_{i+2}$.

Thus, the function

$$\phi_{i+1}(v) := \begin{cases} \phi(v), & v \in V_{i+1} \\ \phi_i(v), & v \in S_i \end{cases}$$

satisfies the conditions above.

Proceeding in this way inductively, we can therefore find ϕ_{h-1} satisfying the same properties. Now, since for all $v \in V_h$ we have $|N_T^+(\phi_{h-1}(N_D^-(v)))| \geq s_h \geq |V_h|$ we can greedily extend ϕ_{h-1} into ϕ_h satisfying the above conditions. In particular, ϕ_h is an embedding of D into T .

To complete the proof, it remains to estimate N , the number of vertices in T . Plugging in our choice of $\varepsilon = 2/\Delta^-$, as well as the estimates $1+x \leq e^x$ and $\Delta_j^- \leq \Delta^-$ for all j , we find that

$$(2+\varepsilon)^{2\Delta_{j-1}^-+2\Delta_j^-} = (2(1+\varepsilon/2))^{2\Delta_{j-1}^-+2\Delta_j^-} \leq e^4 2^{2\Delta_{j-1}^-+2\Delta_j^-}.$$

Additionally, we have that

$$\begin{aligned} \sum_{i=1}^h n_i &= \sum_{i=1}^h \sum_{j=i}^h \frac{(2+\varepsilon)^{2(\Delta_{j-1}^-+\Delta_j^-)} |V_j|}{2^{j-i}} \\ &= \sum_{j=1}^h (2+\varepsilon)^{2(\Delta_{j-1}^-+\Delta_j^-)} |V_j| \sum_{i=1}^j \frac{1}{2^{j-i}} \\ &\leq 2 \sum_{j=1}^h (2+\varepsilon)^{2(\Delta_{j-1}^-+\Delta_j^-)} |V_j|. \end{aligned}$$

Therefore,

$$o_1 = 2k \sum_{i=1}^h a_i = 2kc_a \sum_{i=1}^h n_i \leq 4kc_a \sum_{j=1}^h (2+\varepsilon)^{2\Delta_{j-1}^-+2\Delta_j^-} |V_j| \leq c(\Delta^-)^2 \Delta^+ \sum_{j=1}^h 2^{2(\Delta_{j-1}^-+\Delta_j^-)} |V_j|,$$

for some absolute constant $c = 4 \cdot 5 \cdot 32 \cdot 2000 \cdot 2 \cdot e^4 \leq 10^9$. Since T was an arbitrary tournament on $N = o_1$ vertices we have shown that

$$\vec{r}(D) \leq 10^9 (\Delta^-)^2 \Delta^+ \sum_{i=1}^h 2^{2(\Delta_{i-1}^-+\Delta_i^-)} |V_i|. \quad \square$$

3 Proof of Theorem 1.6

3.1 Proof outline

In this section, we prove Theorem 1.6, which states that there exists a graded digraph D with height h , maximum degree Δ , and equal number of vertices in each part of the graded partition, such that $\vec{r}(D) \geq c^\Delta |V(D)|$ for an absolute constant $c > 1$.

The main ingredient of the proof is the same statement in the case $h = 2$, from which the general statement will follow. In other words, we first show that exists a bipartite digraph D_0 with vertex classes of size n each and maximum degree at most Δ such that $\vec{r}(D_0) \geq c^\Delta \cdot (2n)$.

Having found such a bipartite digraph, we will generalize the construction for any height h , by taking a graded digraph D such that the induced subgraph between two neighboring parts

of the graded partition is a copy of D_0 . We will show that such a D is not contained in a tournament T obtained by replacing each vertex of a transitive tournament on $H = h/2 - 1$ vertices with a copy of R , a large tournament not containing D_0 .

Indeed, if we let D have the graded partition $V(D) = V_1 \cup \dots \cup V_h$ and let $V(T) = A_1 \cup \dots \cup A_H$ such that each $T[A_i]$ is a copy of R and all the edges go from A_i to A_j for $i < j$, then we show the following. If an embedding of D into T exists, then for any i , if we embedded most of V_i into $A_j \cup \dots \cup A_H$ for some j , then most of V_{i+2} must be embedded into $A_{j+1} \cup \dots \cup A_H$. This will give us a contradiction since there are only $H < h/2$ levels in the tournament T .

It remains to construct D_0 and R such that the argument above works. We will use a construction very similar to the one of Graham, Rödl and Ruciński [29]. Namely, we show that if we take R to be a blow-up of a random tournament and D_0 a sparse random bipartite digraph, then with positive probability R does not contain a copy of D_0 . To make the generalization for any height possible, we will in fact need to show a slightly stronger statement, namely that if we take any two large subsets A' and B' of the two vertex classes of D_0 , then R does not contain a copy of the induced subgraph $D_0[A' \cup B']$.

The remainder of this section contains the details of the argument. We first construct a suitable bipartite digraph D_0 and a suitable tournament R in Sections 3.2 and 3.3 respectively. Then in Section 3.4 we show that R indeed does not contain a copy of D . Finally, in Section 3.5 we give a proof of Theorem 1.6 by generalizing the construction to all heights.

Remark 3.1. As in [29], we quantify the respective sizes of the sets and other quantities with concrete numerical values. For example, “large” subsets of the vertex classes A and B of D_0 will mean $A' \subseteq A$ and $B' \subseteq B$ of sizes at least $0.98|A|$ and $0.98|B|$, respectively. However, we want to stress out that the actual numerical values are not of much importance; what matters is that the dependencies between them work out correctly.

3.2 Finding the guest graph for the bipartite case

We begin by constructing the bipartite digraph D_0 , which we will require to have bounded degree and satisfy certain pseudorandom properties. The following lemma, showing a graph with such properties can be constructed randomly, follows from a standard union bound argument, which we include for completeness. We use the notation $e_H(X, Y)$ to denote the number of pairs in $X \times Y$ that are edges of H .

Lemma 3.2. *There exist constants $c_0 > c_1 > 1$ and Δ_0 such that for each $\Delta \geq \Delta_0$ and $n \geq (c_0)^{2\Delta}$ there exists a bipartite graph H with vertex classes X and Y of size n and maximum degree at most Δ such that the following hold, where $k = (c_0)^\Delta$.*

1. *For all partitions $X = X_1 \cup \dots \cup X_k \cup D_X$ and $Y = Y_1 \cup \dots \cup Y_k \cup D_Y$ with $|X_i|, |Y_i| \leq (c_1/c_0)^\Delta n$ and $|D_X|, |D_Y| \leq 0.02n$, we have*

$$\sum_{i \neq j: e_H(X_i, Y_j) > 0} |X_i| |Y_j| > 0.55(0.98n)^2.$$

2. *For all $X' \subseteq X$ and $Y' \subseteq Y$ such that $|X'|, |Y'| \geq 0.01n$, we have $e_H(X', Y') > 0$.*

Proof. We take any c_0, c_1 satisfying $1 < c_1^2 < c_0 < (5/4)^{1/202}$ and choose Δ_0 so that $(c_1^2/c_0)^{\Delta_0} < 0.1$, $((0.8)^{1/101} c_0^2)^{\Delta_0} < 1/16$ and $(1 - 10^{-4})^{\Delta_0/101} < 1/8$. Note that we can choose such a Δ_0 since all three of these inequalities are satisfied for sufficiently large Δ_0 . Let moreover $\Delta \geq \Delta_0$, $d = \Delta/101$, and $m = 1.01n$.

To obtain our graph H , we will first draw a bipartite graph G uniformly at random from the set of all bipartite graphs with dm edges and with vertex classes V' and V'' of size m each.

Then we will remove the $n/100$ largest degree vertices on each side to obtain H . Since the number of vertices of degree larger than Δ in D is at most $\frac{dm}{\Delta+1} < \frac{m}{101} = \frac{n}{100}$, the maximum degree of H is at most Δ with probability 1. It thus suffices to show that H will also satisfy the other two properties with positive probability.

For the first one, let us bound the probability that there exist partitions $V' = V'_1 \cup \dots \cup V'_k \cup D_X \cup D'$ and $V'' = V''_1 \cup \dots \cup V''_k \cup D_Y \cup D''$ with $|D'| = |D''| = n/100$, $|D_X|, |D_Y| \leq n/50$, and $|V'_i|, |V''_i| \leq (c_1/c_0)^\Delta n$ for all $i \in [k]$, such that

$$\sum_{i \neq j: e_H(V'_i, V''_j) > 0} |V'_i| |V''_j| \leq 0.55(0.98n)^2.$$

To do that, notice that since

$$\sum_{i=1}^k |V'_i| |V''_i| \leq k \left(\left(\frac{c_1}{c_0} \right)^\Delta n \right)^2 = \left(\frac{c_1^2}{c_0} \right)^\Delta n^2 \leq \left(\frac{c_1}{c_0} \right)^{\Delta_0} n^2 < 0.1n^2,$$

such a partition must satisfy

$$\sum_{i \neq j: e_G(V'_i, V''_j) = 0} |V'_i| |V''_j| \geq (0.98n)^2 - 0.1n^2 - 0.55(0.98n)^2 \geq 0.2m^2.$$

Therefore, by the union bound, the probability that such a partition exists is at most

$$(k+2)^{2m} 2^{k^2} \frac{\binom{0.8m^2}{dm}}{\binom{m^2}{dm}} < (2k)^{2m} 2^{k^2} (0.8)^{dm} < 8^m ((0.8)^{1/101} c_0^2)^{\Delta_0 m} < \frac{1}{2},$$

where $(k+2)^{2m}$ is a bound on the number of partitions, 2^{k^2} bounds the number of possible choices of pairs (V'_i, V''_j) with no edges in between them and $\binom{0.8m^2}{dm} / \binom{m^2}{dm}$ is a bound on the probability that indeed no edges fall between them.

Similarly, the probability that there exist $X' \subseteq X$ and $Y' \subseteq Y$ of sizes at least $0.01n$ each such that $e_H(X', Y') = 0$ is at most

$$2^{2m} \frac{\binom{(1-10^4)m^2}{dm}}{\binom{m^2}{dm}} < 2^{2m} (1-10^4)^{dm} \leq 2^{2m} (1-10^4)^{\Delta_0 m / 101} < \frac{1}{2},$$

where 2^{2m} bounds the number of choices of X' and Y' and the fraction bounds the probability that there are no edges between them. Thus, by the union bound, there exists a choice of G such that both properties are satisfied, implying the existence of the desired H . \square

3.3 Finding the host tournament for the bipartite case

Our next lemma is of a similar flavor to Lemma 3.2, showing that a random object typically satisfies a certain pseudorandom property. In this case, we show that a random tournament typically has many directed edges from any large set to any other large set. We actually prove something slightly more general, which says the same not for sets, but for “weighted sets”, that is, for functions valued in $[0, 1]$. Here, and in the rest of the proof, all logarithms are to base e .

Lemma 3.3. *Let $k \geq 2$ and $x \geq (10^8 \log k)/2$. There exists a tournament R with vertex set $[k]$ such that for all pairs of weight functions $f, g : [k] \rightarrow [0, 1]$ with $f+g \leq 1$ and $\sum_{i=1}^k (f(i)+g(i)) = 2x$, we have*

$$W := \sum_{ij \in E(R)} f(i)g(j) \leq 0.51x^2.$$

Proof. Note that $2x = \sum_{i=1}^k (f(i) + g(i)) \leq k$, hence the statement is vacuous if $2x > k$, as in this case there exist no such functions f, g . Thus, we assume henceforth that $2x \leq k$, and in particular that $k > 10^8 \log 2$. We let R be a uniformly random tournament with vertex set $[k]$.

We first claim that we can assume that there exist $i_0 \neq j_0$ such that for all $i \neq i_0$ and all $j \neq j_0$ we have

$$f(i), g(j) \in \{0, 1\}.$$

Indeed, for any fixed outcome of R , suppose that f and g maximize W and that there exists an i such that $0 < f(i), g(i) < 1$. Now consider the sums $W_f(i) := \sum_{j:ij \in E(R)} g(j)$ and $W_g(i) := \sum_{j:ji \in E(R)} f(j)$. If $W_f(i) \geq W_g(i)$ then define new functions f' and g' such that there are equal to f and g except that $f'(i) = f(i) + g(i)$ and $g'(i) = 0$. Otherwise, we set $f'(i) = 0$ and $g'(i) = f(i) + g(i)$. In either case, we have $W' \geq W$ for the corresponding quantity W' .

Thus, we can assume that $\min\{f(i), g(i)\} = 0$ for all i . Now suppose that there exist $i \neq j$ such that $0 < g(i), g(j) < 1$. Again in case $W_g(i) \geq W_g(j)$, we let $\varepsilon_{ij} = \min\{g(j), 1 - g(i)\}$ and let $g'(i) = g(i) + \varepsilon_{ij}$ and $g'(j) = g(j) - \varepsilon_{ij}$. Otherwise, we do the same with i and j swapped and in both cases we get $W' \geq W$. By the same argument, we can also assume that for at most one i_0 we have $0 < f(i_0) < 1$.

Now, assuming f and g satisfy the property above, define $T = \{i : f(i) = 1\}$ and $S = \{j : g(j) = 1\}$ and let $t = |T|, s = |S|$. By our assumptions, we have that $t + s \leq 2x < t + s + 2$ and $2x \leq k$. Additionally, we have

$$W = \sum_{ij \in E(R)} f(i)g(j) = e_R(T, S) + g(j_0)W_g(j_0) + f(i_0)W_f(i_0) \leq e_R(T, S) + 2x.$$

In particular, if $W > 0.51x^2$ we find that

$$e_R(T, S) > 0.51x^2 - 2x \geq 0.501x^2 \geq 0.501 \frac{(s+t)^2}{4}.$$

We now claim that with positive probability (over the randomness in R), there exist no sets $S, T \subseteq V(R)$ satisfying this inequality. Suppose first that $t \leq s$. Note that we must have $t > s/7$ since otherwise $e_R(T, S) \leq ts < 0.5(t+s)^2/4$. Similarly, we must have $s > s_0 := 2 \cdot 10^7 \log k$. For any fixed disjoint T, S we have that $e_R(T, S) \sim \text{Bin}(ts, \frac{1}{2})$ and thus by Chernoff's inequality, and using $(t+s)^2/4 \geq ts \geq s^2/7$, we get

$$\Pr[e_R(T, S) > 0.501(t+s)^2/4] \leq \Pr[e_R(T, S) > 0.501ts] < e^{-10^{-7}s^2}.$$

Moreover, for $s > s_0$, we have

$$e^{-10^{-7}s^2} < e^{-10^{-7}s \cdot s_0} = k^{-2s}.$$

Therefore, the probability that such T and S exist is at most

$$\sum_{s=s_0}^k \sum_{t=s/7}^s \binom{k}{s} \binom{k}{t} e^{-10^{-7}s^2} \leq \sum_{s=s_0}^k \sum_{t=s/7}^s \left(\frac{ek}{s}\right)^{2s} k^{-2s} \leq k \sum_{s=s_0}^k \left(\frac{e^2}{s^2}\right)^s < k^2 \left(\frac{e^2}{s_0^2}\right)^{s_0} < \frac{1}{2},$$

where in the first inequality we use that since $t \leq s$, we have that $\binom{k}{s} \binom{k}{t} \leq \binom{k}{s}^2 < (ek/s)^{2s}$. By interchanging the roles of s and t , we obtain the same bound in case $t \geq s$. Thus, we find that R satisfies the desired property with positive probability. \square

3.4 The bipartite case

We are now ready to prove Theorem 1.6 in the case $h = 2$, that is, when D is bipartite. As discussed above, this step is actually the heart of the proof, as the proof for arbitrary h will essentially be a reduction to this case.

Lemma 3.4. *There exist constants $c' > 1$ and Δ_0 such that for all $\Delta \geq \Delta_0$ and $n \geq \Delta$ the following holds. There exists a bipartite digraph $D_0 = (A \cup B, E)$ with $|A| = |B| \leq n$ and maximum degree Δ such that all its edges are directed from A to B , as well as a tournament R on $(c')^\Delta n$ vertices such that for any $A' \subseteq A$ and $B' \subseteq B$, the following hold.*

1. If $|A'|, |B'| \geq 0.98|A|$ then there is no copy of $D_0[A' \cup B']$ in R , and
2. if $|A'|, |B'| \geq 0.01|A|$ then $e_{D_0}(A', B') > 0$.

Proof. Let c_0, c_1 and Δ_0 be the constants from Lemma 3.2. By potentially increasing Δ_0 further, we may also assume that $(c_0/c_1)^\Delta > 10^9 \Delta \log c_0$ for all $\Delta \geq \Delta_0$. We let $1 < c' = \min\{c_1, 2^{0.3}/c_0\}$.

If $\Delta \geq \Delta_0$ and $\Delta \leq n < \frac{1}{0.98} 2c_0^\Delta$, let D_0 be the oriented complete bipartite graph $\vec{K}_{\Delta, \Delta}$ with vertex classes A and B and where all the edges are oriented from A to B . Clearly, for any non-empty $A' \subseteq A$ and $B' \subseteq B$ we have $e_D(A', B') > 0$. To show that the first property also holds, note that if $|A'|, |B'| > 0.98\Delta$ then $D_0[A' \cup B']$ has the complete bipartite graph $\vec{K}_{[0.98\Delta], [0.98\Delta]}$ as a subgraph. Therefore, since the probability that a uniformly random tournament R on $N = 2^{\lfloor 0.98\Delta \rfloor / 2} \geq (c')^\Delta n$ vertices contains a copy of $\vec{K}_{[0.49\Delta], [0.49\Delta]}$ is at most $\binom{N}{[0.98\Delta]} \binom{N}{[0.98\Delta]} 2^{-[0.98\Delta]^2} < N^{2[0.98\Delta]} 2^{-[0.98\Delta]^2} \leq 1$, there is a choice of R such that the conditions of the lemma are satisfied.

Otherwise, we have $0.98n \geq 2c_0^\Delta$. Let then D_0 be a digraph obtained by taking the graph H from Lemma 3.2 and orienting every edge from A to B , and let R' be the tournament from Lemma 3.3. We obtain a tournament R by taking the blow-up of R' in the following way. Let $N = c_1^\Delta n \geq (c')^\Delta n$, $k = c_0^\Delta$ and partition $[N] = U_1 \cup \dots \cup U_k$ such that $|U_i| = N/k$. Let R be an arbitrary tournament on the vertex set $[N]$ such that for all $ij \in E(R')$ we have $U_i \times U_j \subseteq E(R)$. That is, the edges between distinct U_i, U_j form a blown-up copy of R' , and the edges inside any U_i are oriented arbitrarily.

D_0 satisfies the second condition of the lemma by Lemma 3.2. Suppose now for contradiction that for some $A' \subseteq A$ and $B' \subseteq B$ with $|A'|, |B'| \geq 0.98n$ there is a copy of $D_0[A' \cup B']$ in R and let X and Y be its two vertex classes. By Lemma 3.2 we have that for $X_i := X \cap U_i$ and $Y_i := Y \cap U_i$ it holds

$$\sum_{ij \in E(R')} |X_i||Y_j| \geq \sum_{i \neq j: e_D(X_i, X_j) > 0} |X_i||Y_j| > 0.55(0.98n)^2.$$

For $i \in [k]$, let now $f(i) = \frac{|X_i|k}{N}$ and $g(i) = \frac{|Y_i|k}{N}$. We have $0 \leq f + g \leq 1$ and

$$2x := \sum_i (f(i) + g(i)) = \frac{k}{N} (|X| + |Y|) \geq 2 \cdot 0.98 \frac{nk}{N} = 1.96 \left(\frac{c_0}{c_1} \right)^\Delta \geq 10^9 \Delta \log c_0 > 10^8 \log k.$$

Therefore, by Lemma 3.3,

$$\sum_{ij \in E(R')} |X_i||Y_j| = \frac{N^2}{k^2} \sum_{ij \in E(R)} f(i)g(j) < \frac{N^2}{k^2} 0.51x^2 \leq 0.51(0.98n)^2,$$

a contradiction. This shows that there is no copy of $D_0[A' \cup B']$ in R for any such A' and B' . \square

3.5 Proof of Theorem 1.6

With the ingredients above, we are ready to prove Theorem 1.6.

Proof of Theorem 1.6. Let $c' > 1$ and Δ'_0 be the constants from Theorem 3.4 and set Δ_0 to be a constant such that $\Delta_0 \geq \Delta'_0$ and $(c')^{\Delta_0/2} > 4$. Moreover, let $c > 1$ be a constant such that $c^{\Delta_0} = (c')^{\Delta_0/2}/4$. Let $\Delta \geq 2\Delta_0$ and $n \geq \Delta$. Let now $D_0 = (A \cup B, E)$ and R be respectively the bipartite digraph and tournament from Lemma 3.4, applied with the parameters n and $\Delta/2$.

If $h = 2$, then we take $D = D_0$ and since R doesn't contain a copy of D_0 we get $\vec{r}(D) \geq (c')^{\Delta/2}n \geq c^{\Delta}n$. Otherwise, we define a graded digraph D on nh vertices and with a graded partition $V_1 \cup \dots \cup V_h$, where $|V_i| = n$ for all i , by declaring that for all $i \in [h-1]$, the induced subgraph $D[V_i \cup V_{i+1}]$ is a copy of D_0 such that V_i plays the role of A and V_{i+1} plays the role of B . Note that the maximum degree in D is at most Δ , since the maximum in-degree and maximum out-degree are both at most $\Delta/2$.

Now let $H = \lceil \frac{h}{2} \rceil - 1 > 1$ and define a tournament T on $(c')^{\Delta/2}Hn$ vertices with a vertex partition $V(T) = A_1 \cup \dots \cup A_H$, where $|A_i| = (c')^{\Delta/2}n$ for each $i \in [H]$, as follows. For each i , we let $T[A_i]$ be a copy of R , and for all $1 \leq i < j \leq H$, we direct all edges from A_i to A_j . We claim that there is no copy of D in T .

Indeed, suppose for contradiction that ϕ is an embedding of D into T . For $i \in [h]$ and $A' \subseteq V(T)$, let $f_i(A') = \frac{|\phi(V_i) \cap A'|}{|V_i|}$ be the fraction of vertices of V_i embedded into A' . Additionally, let $U_i = \bigcup_{j=i}^H A_j$, and let j_i be the largest index $1 \leq j \leq H$ such that $f_i(U_j) \geq 0.99$. Note that this is well-defined since $U_1 = V(T)$, and hence $f_i(U_1) = 1$.

The key observation is that if $f_i(U_j) \geq 0.01$, then $f_{i+1}(U_j) \geq 0.99$. Indeed, if this is not the case, then there exist $X_i \subseteq V_i, X_{i+1} \subseteq V_{i+1}$ with $|X_i|, |X_{i+1}| \geq 0.01n$, with the properties that $\phi(X_i) \subseteq U_j$ and $\phi(X_{i+1}) \subseteq V(T) \setminus U_j$. But all edges in T are directed from $V(T) \setminus U_j$ to U_j , hence the second condition in Lemma 3.4 implies that if such X_i, X_{i+1} exist, then ϕ is not a valid embedding.

In particular, applying this observation with $j = j_i$, we conclude that $f_{i+1}(U_{j_i}) \geq 0.99$. This implies that $j_{i+1} \geq j_i$ for all $i \in [h-1]$, that is, that the indices j_i are monotonically non-decreasing. We now claim that for each $i \in [h-2]$, we have that $j_{i+2} > j_i$.

Indeed, if $j_{i+1} > j_i$, then we are done by the monotonicity property $j_{i+2} \geq j_{i+1}$. Hence we may assume that $j_i = j_{i+1}$, which in particular implies that $f_i(U_{j_{i+1}}) < 0.01$ by the key observation. If $f_{i+1}(U_{j_{i+1}}) \geq 0.01$, then we are again done by the key observation. Therefore, we may assume that $f_{i+1}(U_{j_{i+1}}) < 0.01$. Together with the fact that $j_{i+1} = j_i$, we conclude that $f_i(A_{j_i}), f_{i+1}(A_{j_i}) \geq 0.98$.

In other words, there exist $X_i \subseteq V_i, X_{i+1} \subseteq V_{i+1}$ with $|X_i|, |X_{i+1}| \geq 0.98n$ such that $\phi(X_i), \phi(X_{i+1}) \subseteq A_{j_i}$. But this is a contradiction to Lemma 3.4, since $T[A_{j_i}]$ is a copy of R' . We conclude that, as claimed, $j_{i+2} > j_i$ for all $i \in [h-2]$.

Since $j_1 \geq 1$ and $j_{i+2} \geq j_i + 1$ for all i , we find that $j_i \geq i/2$ for all i . In particular, $j_h \geq h/2 > H$. But this is a contradiction as there are only H parts in T , implying that there is no copy of D in T . Therefore,

$$\vec{r}(D) > (c')^{\Delta/2}Hn \geq \frac{1}{4}(c')^{\Delta/2}hn \geq c^{\Delta}hn. \quad \square$$

4 Concluding remarks

While Theorem 1.3 is roughly best possible in general, it is reasonable to expect that one could improve it in certain cases. In particular, for the oriented hypercube \vec{Q}_d , we expect that the bound in Theorem 1.5 could be significantly improved. In fact, our techniques are

already sufficient show that the induced subgraph of \vec{Q}_d obtained by taking the vertices with at most $d/2$ non-zero coordinates has oriented Ramsey number at most $2^{3d+o(d)}$. As this digraph consists of simply the first half of the graded partition of \vec{Q}_d , this suggests to us that the bound in Theorem 1.5 is not particularly close to best possible. Concretely, we make the following conjecture, which is a directed analogue of the Burr–Erdős conjecture [6] that $r(Q_d) = O(2^d)$.

Conjecture 4.1. *There is an absolute constant $C > 0$ such that $\vec{r}(\vec{Q}_d) \leq C2^d$ for all $d \geq 1$.*

We also recall that Fox, He, and Wigderson [21] proved that any bounded-height, bounded-degree acyclic digraph has linear oriented Ramsey number, and Theorem 1.3 implies that the same is true for graded digraphs even of unbounded height. It would be very interesting to extend this result, and identify other classes of unbounded-height digraphs where an analogue of the Burr–Erdős conjecture holds.

Finally, we reiterate the main question left open from [21], which we consider to be of central importance in the study of oriented Ramsey numbers.

Question 4.2. *Given $\Delta \geq 1$, does there exist some $C > 0$ such that every n -vertex acyclic digraph D with maximum degree Δ satisfies $\vec{r}(D) \leq n^C$?*

Acknowledgments: We are grateful to Domagoj Bradač and Benny Sudakov for helpful discussions, and for suggesting a simplification to the proof of Lemma 2.2. We would also like to thank Xioayu He for constructive comments on an earlier draft of this paper.

References

- [1] J. Beck, An upper bound for diagonal Ramsey numbers, *Studia Sci. Math. Hungar.* **18** (1983), 401–406.
- [2] A. Benford and R. Montgomery, Trees with few leaves in tournaments, *J. Combin. Theory Ser. B* **155** (2022), 141–170.
- [3] A. Benford and R. Montgomery, Trees with many leaves in tournaments, 2022. Preprint available at arXiv:2207.06384.
- [4] D. Bradač, O. Janzer, B. Sudakov, and I. Tomon, The Turán number of the grid, *Bull. Lond. Math. Soc.* **55** (2023), 194–204.
- [5] M. Bucić, S. Letzter, and B. Sudakov, Directed Ramsey number for trees, *J. Combin. Theory Ser. B* **137** (2019), 145–177.
- [6] S. A. Burr and P. Erdős, On the magnitude of generalized Ramsey numbers for graphs, in *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday)*, Vols. I, II, III, *Colloq. Math. Soc. János Bolyai*, vol. Vol. 10, North-Holland, Amsterdam-London, 1975, 215–240.
- [7] G. Chen and R. H. Schelp, Graphs with linearly bounded Ramsey numbers, *J. Combin. Theory Ser. B* **57** (1993), 138–149.
- [8] C. Chvatál, V. Rödl, E. Szemerédi, and W. T. Trotter, Jr., The Ramsey number of a graph with bounded maximum degree, *J. Combin. Theory Ser. B* **34** (1983), 239–243.
- [9] D. Clemens, M. Miralaei, D. Reding, M. Schacht, and A. Taraz, On the size-Ramsey number of grid graphs, *Combin. Probab. Comput.* **30** (2021), 670–685.

- [10] D. Conlon, Hypergraph packing and sparse bipartite Ramsey numbers, *Combin. Probab. Comput.* **18** (2009), 913–923.
- [11] D. Conlon, The Ramsey number of dense graphs, *Bull. Lond. Math. Soc.* **45** (2013), 483–496.
- [12] D. Conlon, J. Fox, and B. Sudakov, On two problems in graph Ramsey theory, *Combinatorica* **32** (2012), 513–535.
- [13] D. Conlon, J. Fox, and B. Sudakov, Short proofs of some extremal results II, *J. Combin. Theory Ser. B* **121** (2016), 173–196.
- [14] D. Conlon, R. Nenadov, and M. Trujić, On the size-Ramsey number of grids, *Combin. Probab. Comput.* **32** (2023), 874–880.
- [15] N. Draganić, F. Dross, J. Fox, A. Girão, F. Havet, D. Korándi, W. Lochet, D. Munhá Correia, A. Scott, and B. Sudakov, Powers of paths in tournaments, *Combin. Probab. Comput.* **30** (2021), 894–898.
- [16] F. Dross and F. Havet, On the unavoidability of oriented trees, *Journal of Combinatorial Theory, Series B* **151** (2021), 83–110.
- [17] A. El Sahili, Trees in tournaments, *J. Combin. Theory Ser. B* **92** (2004), 183–187.
- [18] P. Erdős and R. L. Graham, On partition theorems for finite graphs, in *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vols. I, II, III, Colloq. Math. Soc. János Bolyai*, vol. Vol. 10, North-Holland, Amsterdam-London, 1975, 515–527.
- [19] P. Erdős and L. Lovász, Problems and results on 3-chromatic hypergraphs and some related questions, in *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vols. I, II, III, Colloq. Math. Soc. János Bolyai*, vol. Vol. 10, North-Holland, Amsterdam-London, 1975, 609–627.
- [20] P. Erdős and L. Moser, On the representation of directed graphs as unions of orderings, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **9** (1964), 125–132.
- [21] J. Fox, X. He, and Y. Wigderson, Ramsey numbers of sparse digraphs, *Israel J. Math.* (2024), to appear. Preprint available at arXiv:2105.02383.
- [22] J. Fox and B. Sudakov, Density theorems for bipartite graphs and related Ramsey-type results, *Combinatorica* **29** (2009), 153–196.
- [23] J. Fox and B. Sudakov, Two remarks on the Burr-Erdős conjecture, *European J. Combin.* **30** (2009), 1630–1645.
- [24] J. Fox and B. Sudakov, Dependent random choice, *Random Structures Algorithms* **38** (2011), 68–99.
- [25] Z. Füredi and M. Ruszinkó, Uniform hypergraphs containing no grids, *Adv. Math.* **240** (2013), 302–324.
- [26] J. Gao, O. Janzer, H. Liu, and Z. Xu, Extremal number of graphs from geometric shapes, *Israel J. Math.* (2024), to appear. Preprint available at arXiv:2303.13380.

- [27] L. Gishboliner and A. Shapira, Constructing dense grid-free linear 3-graphs, *Proc. Amer. Math. Soc.* **150** (2022), 69–74.
- [28] R. L. Graham, V. Rödl, and A. Ruciński, On graphs with linear Ramsey numbers, *J. Graph Theory* **35** (2000), 176–192.
- [29] R. L. Graham, V. Rödl, and A. Ruciński, On bipartite graphs with linear Ramsey numbers, *Combinatorica* **21** (2001), 199–209.
- [30] R. Häggkvist and A. Thomason, Trees in tournaments, *Combinatorica* **11** (1991), 123–130.
- [31] F. Havet, Trees in tournaments, *Discrete Math.* **243** (2002), 121–134.
- [32] F. Havet and S. Thomassé, Median orders of tournaments: a tool for the second neighborhood problem and Sumner’s conjecture, *J. Graph Theory* **35** (2000), 244–256.
- [33] F. Havet and S. Thomassé, Oriented Hamiltonian paths in tournaments: a proof of Rosenfeld’s conjecture, *J. Combin. Theory Ser. B* **78** (2000), 243–273.
- [34] J. H. Kim, C. Lee, and J. Lee, Two approaches to Sidorenko’s conjecture, *Trans. Amer. Math. Soc.* **368** (2016), 5057–5074.
- [35] A. Kostochka and B. Sudakov, On Ramsey numbers of sparse graphs, *Combin. Probab. Comput.* **12** (2003), 627–641.
- [36] A. V. Kostochka and V. Rödl, On graphs with small Ramsey numbers, *J. Graph Theory* **37** (2001), 198–204.
- [37] D. Kühn, R. Mycroft, and D. Osthus, An approximate version of Sumner’s universal tournament conjecture, *J. Combin. Theory Ser. B* **101** (2011), 415–447.
- [38] D. Kühn, R. Mycroft, and D. Osthus, A proof of Sumner’s universal tournament conjecture for large tournaments, *Proc. Lond. Math. Soc. (3)* **102** (2011), 731–766.
- [39] C. Lee, Ramsey numbers of degenerate graphs, *Ann. of Math. (2)* **185** (2017), 791–829.
- [40] G. O. Mota, G. N. Sárközy, M. Schacht, and A. Taraz, Ramsey numbers for bipartite graphs with small bandwidth, *European J. Combin.* **48** (2015), 165–176.
- [41] L. Shi, Cube Ramsey numbers are polynomial, *Random Structures Algorithms* **19** (2001), 99–101.
- [42] L. Shi, The tail is cut for Ramsey numbers of cubes, *Discrete Math.* **307** (2007), 290–292.
- [43] R. Stearns, The voting problem, *Amer. Math. Monthly* **66** (1959), 761–763.
- [44] B. Sudakov, A conjecture of Erdős on graph Ramsey numbers, *Adv. Math.* **227** (2011), 601–609.
- [45] A. Thomason, Paths and cycles in tournaments, *Trans. Amer. Math. Soc.* **296** (1986), 167–180.
- [46] K. Tikhomirov, A remark on the Ramsey number of the hypercube, *European J. Combin.* **120** (2024), Paper No. 103954.